A joint Initiative of Ludwig-Maximilians-Universität and Ifo Institute for Economic Research

CESifo Working Papers

PUBLIC HEALTH CARE WITH WAITING TIME: THE ROLE OF SUPPLEMENTARY PRIVATE HEALTH CARE

Michael Hoel
Erik Magnus Sæther

CESifo Working Paper No. 562

September 2001

CESifo
Center for Economic Studies & Ifo Institute for Economic Research
Poschingerstr. 5, 81679 Munich, Germany
Phone: +49 (89) 9224-1410 - Fax: +49 (89) 9224-1409
e-mail: office@CESifo.de
ISSN 1617-9595

CESifo *Working Paper No. 562*
September 2001

PUBLIC HEALTH CARE WITH WAITING TIME: THE ROLE OF SUPPLEMENTARY PRIVATE HEALTH CARE

Abstract

We consider an economy where most of the health care is publicly provided, and where there is waiting time for several types of treatments. Private health care without waiting time is an option for the patients in the public health queue. We show that although patients with low waiting costs will choose public treatment, they may be better off with waiting time than without. The reason is that waiting time induces patients with high waiting costs to choose private treatment, thus reducing the cost of public health care that everyone pays for. Even if higher quality (i.e. zero waiting time) can be achieved at no cost, the self-selection induced redistribution may imply that it is socially optimal to provide health care publicly and at an inferior quality level. We give a detailed discussion of the circumstances in which it is optimal to have waiting time for public health treatment. Moreover, we study the interaction between this quality decision and the optimal tax/subsidy on private health care.

JEL Classification: H42, H51, I111, I118.

Keywords: public health care, private health care, waiting time, health queues.

Michael Hoel
Department of Economics
University of Oslo
P.O.Box 1095
0317 Blindern
Norway

Erik Magnus Sæther
Ragnar Frisch Centre for Economic Research
Gaustadalléen 21
0349 Oslo
Norway
e.m.sather@frisch.uio.no

1 **Introduction**

In several countries with dominantly public health care, there are often queues for some types of treatments. Waiting time in the public health system is often explained by some referral to limited public resources. It is however not clear why a system with a queue should cost less than a system without. One obvious explanation is that demand for most types of health services fluctuates over time. If one were to dimension the capacity of the health system such that there never was any waiting time, there would be periods of idle capacity.[1] This would be more costly than a system in which there always was full capacity utilization, and with a waiting time during periods of high demand[2]. However, if this were the only reason for having waiting time, one would expect the waiting time to fluctuate between something close to zero and to, say, a month or two. The waiting times observed for many types of treatments are however considerably larger. More importantly, although they fluctuate, they are always bounded well away from zero. For instance, in England there were almost 2 million patients waiting for outpatient services, and more than 1 million patients registered for ordinary (inpatient) or day case admissions in the National Health Service in September 1999. Of the latter, 49% had been waiting for 3 months or more, and 26% had been waiting for more than 6 months. Similarly, in Norway the average waiting time for non-prioritized patients varied from about 3 months (outpatients) to about 4 months (day case and inpatients). [3] Clearly, cost savings due to better capacity utilization cannot explain waiting times of this length.

One way aggregate costs could be held down through queues is to let the treatment per unit of time be lower than the flow of new cases per unit of time. If all new cases were added to the queue, this would make waiting times increase without any limit. Although e.g. NHS waiting lists have grown over time, they have not grown as such rates. The reason we do not see such rapid growth in waiting times is that the queue itself has an effect on how the flow of cases translates into a flow demand for treatment in the public sector. One possibility is that the queue causes some people to exit from the queue before being treated.

---

[1] Worthington (1987) gives an analysis of waiting time that is caused by stochastic demand for health services.
[2] Mobley and Magnussen (1998) present the need for excess capacity to ensure availability in private sector as an explanation of why they found no support to the hypothesis that private American hospitals in an competitive environment are more efficient than Norwegian public hospitals.
[3] See Department of Health, UK (1999) and Norwegian Patient Register (1999). A further discussion of waiting lists in England and Norway is given in Appendix D of Hoel and Sæther (2000).

The most drastic form of exit would be that patients die while waiting for treatment. Even though there are surveys confirming such deaths[4], the longest queues typically are for medical cases that are not life threatening. A more positive possibility is that the illness heals on its own while waiting for treatment. To the extent that this occurs, the patient's cost of the queue is the postponement in recovery. A related possibility is that after experiencing a particular health defect for some time, a patient finds it less unpleasant than they initially find it. If medical treatment (e.g. an operation) has some risk of actually making the condition worse, this may imply that after a period of waiting the patient prefers to exit from the queue and accept the health defect. An alternative explanation is given by Lindsay and Feigenbaum (1984): They assume that waiting time induces some people to never join the queue for treatment, due to the existence of costs incurred at the time one joins the queue. When this is the case, it can be better to never receive treatment than to receive treatment after a waiting period.

Several of the possibilities above probably are relevant explanations of how the existence of a queue might reduce the flow demand for health treatment in the public sector. We shall ignore all of these possibilities in the present paper, and instead focus on what perhaps is a more important effect of waiting time in the public sector. Patients who enter into such a queue for medical treatment sometimes have the option of using a private alternative to the public health care, thus avoiding the queue (see e.g. Cullis and Jones (1985), Iversen (1997)). However, by doing this they often incur larger costs, as they have to pay for the private treatment (directly or through a private supplementary insurance they previously have purchased), while the treatment in the public system would have been free or almost free. The longer the waiting time, the more people choose the private alternative. The waiting time is thus an equilibrating mechanism making the demand for public treatment equal the supply, which is politically determined.

In Section 2 we set up a model describing the choice between treatment in the public and private sector. In Section 3 we use this model to discuss health queues within a framework of standard

[4] Plump et al. (1999) examined the circumstances of death regarding patients who died in 1994 and 1995 while on waiting list for cardiac surgery in the Netherlands. They found that waiting lists for cardiac surgery engender high risks for the patients involved and approximately 100 deaths per year in this patient group was waiting list related. At least half of the deaths occur within the first six weeks.

welfare theory. In particular, we wish to see what type of considerations might make waiting time for treatment in the public sector part of a welfare maximizing policy.

An important result in our analysis is that if distributional objectives are sufficiently strong, it may be optimal to have waiting time for public treatment. This result has some similarities to the literature on public provision of private goods. An important contribution to this literature is Besley and Coate (1991). They present a model in which the quantity demanded for a good is given (as in our model), but that if privately purchased, there is a choice of quality. They show that if sufficient weight is given to equity, welfare can be improved if this good of a particular quality is offered for free by the public sector, financed through a head tax. The favorable distributional effect is caused by the high-income persons choosing to buy the good privately, in a higher quality version, instead of accepting the publicly offered version of the good. Just like in our model, there is thus a self-selection mechanism that gives the desired result. Similar results focusing on the self-selection mechanism have been given by Blackorby and Donaldson (1988), Boadway and Marchand (1995) and Blomquist and Christiansen (1995). The two latter articles also demonstrate that there is a role for publicly provided private goods even if an optimal non-linear tax is available for redistribution.

Our model has some similarities to the models used in the literature above. In our model the quality dimension of the good is waiting time for public treatment. However, unlike the models above, there is no reduction in production costs associated with a reduction in the quality level from the best (no waiting time) to an inferior quality level (positive waiting time). We show that even when higher quality can be achieved at no cost, the self-selection induced redistribution may imply that it is socially optimal to provide the good publicly and at an inferior quality level.

Unlike most of the literature discussed above, we give a relatively detailed discussion of the circumstances in which it is optimal to have an inferior quality of the publicly provided good (i.e. waiting time for public health treatment). Moreover, we study the interaction between this quality decision and the optimal tax/subsidy on the privately supplied version of the good.

An important issue in a system with predominantly public health care is how the government should treat alternative private treatment. It has been argued that a private alternative may undermine the public system, so that the government ought to discourage any private alternative. The most drastic form of "discouragement" would be to forbid various types of private treatment. Norway can serve as an interesting example, where the private-for-profit health care providers face a prohibitive tax in the form of legal regulation prohibiting new inpatient facilities (some beds were accepted before the law came into practice in 1986).[5] A less drastic form of discouragement would be to impose a tax on private treatment. One could however also argue that those who choose the private alternative should be subsidized by the public health insurance. One argument for such subsidization is that in a public system, everyone has paid his or her mandatory insurance premium. Therefore, everyone should be entitled to compensation if they become ill. In particular, a person choosing the private system should be entitled to whatever it would have cost to treat this person in the public system.

The argument above for subsidizing private health care was based on fairness. However, even disregarding the issue of fairness, one could make an argument for such subsidization. By subsidizing the private alternative, the cost of this alternative will be lowered. Therefore more people will choose this alternative. If the subsidy is sufficiently below the cost of treatment in the public sector, there may be a net cost saving for the public sector. This cost saving could be used to expand the treatment capacity in the public health care system, and thus reduce queues for those who don't choose the private alternative. In other words, even if we give no weight to the interests of those who choose to use the private alternative, it might be sensible to partially subsidize treatment in the private sector. This reason for subsidizing private treatment is briefly discussed in Cullis and Jones (1985).

[5] However, the policy in Norway is not very consistent: The local governments and the National Insurance scheme are the key purchasers of private services to reduce the public waiting lists. During the last years there have been several initiatives to purchase privately provided services, also for inpatients. The Norwegian National Insurance scheme finances private health care services for employed on sick leave, restricted to those with a prognosis for a rapid return to work. Some counties in addition offer the whole population a choice between a free private or public treatment. There are also municipalities that provide their community with a free private health insurance scheme.

In Sections 4-6 we use our basic model to see what the optimal tax/subsidy of private sector treatment is. Section 4 discusses the case when waiting time and the tax/subsidy are chosen simultaneously without any further restictions. In Section 5, we add the restriction that the capacity of the public health treatment is given. Finally, Section 6 considers the case where the waiting time is given exogenously by mechanisms not explained by the model.

Section 7 offers some concluding remarks.

2. A model of public and private health care

Consider the simple case in which an exogenously given (and non-stochastic) number of cases requiring medical treatment of a particular type occur each year. Denote this number of cases by x. Moreover, assume that in the public sector there is a waiting time T before treatment is performed. Once treatment is given, it is free.[6] The unit cost of treatment is assumed to be constant, denoted by q, in the public health system.[7]

The private sector gives the same type of treatment, but without any waiting time, at a positive price p. Obviously, if there were no costs associated with waiting for treatment, everyone would prefer public to private treatment, since the former is free and the latter is not. There are, however, costs associated with waiting for treatment. One such cost could be that the medical condition deteriorates during the waiting time. The cost of this deterioration would either be a more severe treatment once the patient gets it, and/or a worse condition after treatment than the condition would have been after immediate treatment.[8] Another type of waiting cost is that patients suffer a welfare loss during the waiting period. This welfare loss could be either outright pain or various types of discomfort.[9] For instance, a person waiting for a knee operation would

[6] In the end of this Section we demonstrate that in this model a fee for public treatment is equivalent to subsidizing private treatment.
[7] We thus completely ignore all issues related to the labour market for physicians. Iversen (1997), Brekke and Sørgård (2000) and Marchand and Schroyen (2000) are among the few studies of these issues and their relationship to the interaction between the public and private health sector.
[8] In a study of patients admitted to hospital for elective orthopedic surgery in Norway, Rossvoll et.al. (1993) found that the probability of returning to work after surgery is strongly influenced by the length of time on the waiting list. A high proportion of the patients with a chronic orthopedic disorder were not able to work while waiting.
[9] Hamilton et al. (1996) investigated the effect of waiting time for hip fracture surgery in Canada on post-surgery length of stay in hospital and inpatient mortality. They found no evidence of a detrimental impact caused by pre-

have to abstain from physical activities he/she otherwise would have undertaken. Another example could be a couple that does not wish to have more children, so that one of the persons wishes to be sterilized. During the waiting phase, the couple either must risk pregnancy or at least one of the persons would have to bear the inconveniences of preventive measures.[10] Additional health care cost may also be invoked in the form of care while waiting or the need for new tests and diagnosis.[11] Patients that are not able to work while waiting for treatment normally face a income loss. This loss is often (partially) compensated thrugh a sickness benefit scheme. This incurs a cost for the society at large in the form of increased tax distortions to fund the benefit schemes, in addition to the production loss of absenteeism[12].

Whatever the background for the waiting costs, we shall assume that they are proportional to the waiting time.[13] The cost per unit of waiting time is assumed to vary among the population. We would expect this variation to be correlated to income variations, as a higher income typically will imply a higher willingness to pay to avoid waiting. However, waiting costs are also likely to vary among individuals for other reasons: An active skier or runner is likely to have considerably higher waiting costs for a knee operation than a person with a less active life style.

Denote the waiting cost per unit of waiting time for a particular person by θ, so that the total waiting cost for this person is θT. The distribution of waiting costs across the population is given by the distribution function $F(\theta)$. The lowest and highest values of θ are α and β, respectively, so that $F(\alpha)=0$ and $F(\beta)=1$.

---

surgery delay, but that surgery delay may lead to greater pre-surgery inpatient costs and more patient discomfort. Roy and Hunter (1996) studied 97 orthopaedic patients awaiting lower-limb surgery. 90 had pain, 44 significant night pains. Psychological and social problems were common. Only 11 were employed full-time. 68 required help with daily activities and 48 patients walked less than 120 metres in 12 minutes. The study also revealed that the planned procedure was no longer appropriate for 12 of the 97 patients.

[10] Using Norwegian data, Hørding et al. (1982) showed that the rate of abortions among women on waiting lists for sterilization was 3.4 times the rate in the normal population.

[11] Stern & Brown (1994) establish a significant relationship between failure to attend initial appointments and the length of time between referral and appointments in a child and family clinic.

[12] With a long waiting time the size of these costs may by far outnumber the direct medical expenses. Hagen and Østtveiten (1999) report in an evaluation of a Norwegian project that uses sickness benefits to procure non-complicated health services that the average medical expenses equals 14 days of sickness benefit payments.

[13] Notice that this assumption implies that the analysis of waiting lists by e.g. Lindsay and Feigenbaum (1984) does not apply to the present case, as a crucial assumption in their analysis is that there is a positive fixed cost of joining the waiting list.

From the assumptions above, it is straightforward to derive the demand for private treatment. A person will choose private treatment if and only if the waiting cost for public treatment (θT) exceeds the price of private treatment (p). This gives the demand for private treatment, denoted by y, as

$$y = x(1 - F(\frac{p}{T})) = y(\pi) \tag{1}$$

where π=p/T is the ratio between the private sector price and the public sector waiting time. If this ratio is sufficiently low, everyone will choose private treatment, while if the ratio π is sufficiently high, no one will choose private treatment. Formally, it follows from (1) that

$$y(\pi) = x \qquad for \quad \pi \leq \alpha \tag{2}$$

$$y(\pi) = 0 \qquad for \quad \pi \geq \beta \tag{3}$$

The most interesting case is the when α<π<β, implying 0<y<x.

Let the price of private treatment in the absence of a tax or subsidy be equal to mq. We assume that the parameter m≥1, although this is not obvious. There are at least two reasons why we may expect to find m>1. One reason is that the private sector is assumed to have no waiting time, which implies that it must have a lower capacity utilization, since the need for treatment in reality will fluctuate over time. The second reason for m>1 is that in a health system where the private sector is only a supplement, there is reason to believe that competition will be less than perfect, thus making the equilibrium price exceed the unit cost. On the other hand, the private sector could be more efficient than the public sector. If this were true and the efficiency difference was sufficiently large, this could outweigh the two factors mentioned above, so that the net result was m<1. The reason why we nevertheless assume that m≥1 is that if m<1, the public sector could purchase health services from the private sector instead of producing them. By doing this, the unit cost of publicly provided health services would be brought down to the price of privately produced services, thus making m=1.[14]

Assume that the public sector taxes or subsidizes treatment in the private sector at a rate t (i.e. t>0 is a tax and t<0 is a subsidy), so that the net price paid by users of the private system is p=mq+t. The total costs for the public sector related to the medical care under consideration consists of treatment costs plus the costs of subsidizing the private sector, or minus the revenue from taxing the private sector. Denoting the total costs by C we thus have

$$C = q(x - y(\pi)) - ty(\pi) \tag{4}$$

Nothing is lost by normalizing units so x=1. With this normalization we may rewrite (4) as

$$C = q - (q + t)y(\pi) \tag{5}$$

The total expected costs of the health system for a person of type θ consists of two terms. The first term is this person's contribution to the expenditures of the public system. Assume that that total expenditures of the public health system are shared equally between everyone, so this first term is equal to C/N where N is the size of the population.[15] The second term is the expected costs of treatment should the person become ill. The probability of becoming ill is x/N, and if this event occurs the cost of treatment is the lowest of waiting costs (=θT) and the cost of treatment in the private sector (=p(t)=mq+t). Denoting total expected cost for a person of type θ by B we thus have

$$B = \frac{C}{N} + \frac{x}{N}\min\left[\theta T, p(t)\right] \tag{6}$$

[14] In spite of this argument, data from Norway indicate that for some types of treatment the price charged by private hospitals is considerably lower than the costs in public hospitals. A further discussion of the costs of private and public health services in Norway is given in Appendix E of Hoel and Sæther (2000).

[15] As we shall soon see, a small change in taxes of the type "equal absolute change for all" does not contradict an optimal design of the tax system, provided the initial tax system is optimally designed.

Since both N and x are given, costs per person and costs per medical case are strictly proportional. It is slightly more conveniently to work with the latter cost, which we denote b(θ,T,s) = BN/x. Inserting from (8) we thus have (using our normalization x=1):

$$b(\theta,T,t) = q-(q+t)y(\frac{p(t)}{T})+\min\left[\theta T, p(t)\right] \tag{7}$$

The objective of the government is to minimize the sum of costs for all persons. However, it is not reasonable to assume that this sum is unweighted. If the government cares about equity, it is reasonable to assume that low-income persons have a higher weight in the sum of costs than high-income persons. In the hypothetical case of the government being perfectly informed about everyone's income earning abilities, lump-sum taxation could be used to obtain whatever distributional objectives the government had. In this case everyone would have the same welfare weight (on the margin) after the lump-sum taxation had been applied. In practice the government does not have perfect information about individuals' abilities. Distributional considerations must therefore be achieved through distortionary taxation. In this case it is not optimal to redistribute so much that marginal welfare weights are equalized across the population. The optimal design of a distortionary tax system implies that redistribution has been taken to the point where the social gain from further redistribution is exactly offset by the incremental distortion of higher rates of taxation. Notice that for an optimally designed tax system of this type, social welfare cannot be increased by increasing or reducing a tax component which is equal for all (and thus non-distortionary) and adjusting the distortionary part of the tax system so that total revenue is unchanged. On the margin, it thus makes no difference whether an increase in health expenditures is financed by a non-distortionary tax increase i.e. a tax increase that is the same for everyone (as assumed in (6)), or by an increase in distortionary tax rates.

We denote the welfare weights used in the social objective function by w(θ). As argued previously, a higher income typically will imply a higher willingness to pay to avoid waiting. We therefore assume that the welfare weights w(θ) are lower the higher is θ, i.e. w'(θ)<0 (with w'=0 as a limiting case). We also normalize the level of the welfare weights so that the average welfare weight over the population is one, i.e.

$$\int_{\alpha}^{\beta} w(\theta) f(\theta) d\theta = 1 \tag{8}$$

where f(θ) is the density function for the distribution of θ (i.e. f(θ) ≡ F'(θ)). For the subsequent analysis, it is convenient to introduce the notation

$$h(\theta) = w(\theta) f(\theta) \tag{9}$$

and

$$H(\theta) = \int_{\alpha}^{\theta} h(i) di \tag{10}$$

Notice that H(α)=0 and H(β)=1. Moreover, if w'(θ)<0 for all θ, then H(θ)>F(θ) for α<θ<β.[16]

The social objective function is given by

$$S(T,t) = \int_{\alpha}^{\beta} w(\theta) f(\theta) b(\theta, T, t) d\theta \tag{11}$$

where b(θ,T,t) is given by (7). Using (8)-(10) we can rewrite this as

$$S(T,t) = T\int_{\alpha}^{\pi} h(\theta)\theta d\theta + p\int_{\pi}^{\beta} h(\theta) d\theta + \left[q - (q+t) y(\pi)\right] \tag{12}$$

Since p=mq+t and π=p/T, S is given once T and t are given. The optimal waiting time and tax/subsidy is the choice of (T,t) that minimizes (12).

---

[16] Proof: assume H(γ)≤F(γ) for some γ∈(α,β). Since H(β)=F(β)=1, this means that H'-F'=h-f=(w-1)f must be non-negative for some θ>γ. But since w'<0 this in turn implies w(γ)>1. From (9) and (10) w'<0 implies H(γ)>w(γ)F(γ). Together with w(γ)>1 we thus have H(γ)>F(γ), which contradicts our assumption H(γ)≤F(γ). We therefore must have H(θ)>F(θ) for all θ∈(α,β).

Before analyzing the optimal waiting time and tax/subsidy, it is useful to consider an obvious third policy instrument in addition to waiting time and a tax/subsidy on private treatment. This third alternative is to charge a price r for treatment in the public sector. With this alternative, each individual will compare r+θT with p when deciding whether to be treated by the public or by the private sector. In the demand function (1), p/T is thus replaced by (p-r)/T. Moreover, in the expression (5) for the cost covered by taxation we must replace q with q-r. With these changes, it is straightforward to see that the total expected costs of the health system for a person of type θ is no longer given by (7), but by

$$b(\theta,T,t,r) = (q-r) - ((q-r)+t)y(\frac{mq+t-r}{T}) + \min\left[\theta T + r, mq+t\right] \tag{13}$$

Introducing the notation τ=t-r, we can rewrite this as

$$b(\theta,T,\tau) = q - (q+\tau)y(\frac{mq+\tau}{T}) + \min\left[\theta T, mq+\tau\right] \tag{14}$$

This is exactly the same as (7) except that t is replaced by τ. The social objective function S, and the analysis of it, thus remains valid, when t is replaced by τ. Since t now is replaces by t-r, this means that a subsidy to the private sector in our model is equivalent to a fee for treatment in the public sector. From now on we ignore the possibility of a fee for public treatment, but simply note that if a subsidy of private is optimal, we can achieve exactly the same by instead introducing a fee for public treatment. Obviously, this equivalence would no longer hold if we relaxed our assumption that total demand for health services was given independently of any prices.

3. The optimal waiting time

In this Section we shall let p be given (e.g. equal to mq, which will be the case when there is no tax or subsidy on private treatment). We return to the issue of a tax or subsidy in the next section.

Using (1) and (10), some manipulation of (12) yields

$$S = q + (m-1)qy(\pi) - p\left[H(\pi) - F(\pi)\right] + TI(\pi) \qquad (15)$$

where

$$I(\pi) = \int_{\alpha}^{\pi} h(\theta)\theta d\theta \qquad (16)$$

The term TI(π) is equal to zero if either T is zero or π≤α, i.e. so small that everyone chooses private treatment. Otherwise the term TI(π) is positive.

It is useful to first consider the case in which everyone has the same welfare weight. In this case the functions H and F are identical, so that the term H-F in (15) is zero. It is clear that in this case the best we can do if m>1 is to choose T=0 (i.e. π=”∞”). This gives y=0 and S=q. If m=1 it remains true that we cannot do better than choose T=0. In this case one could do equally well by setting π≤α, i.e. T so high that y=1, giving B=q also in this case. This is equivalent to letting the health treatment under consideration be fully privatized. We thus have the following proposition:

Proposition 1: For the case when everyone has equal welfare weights, it is never optimal to have a positive waiting time for public health treatment.

From the discussion in the Introduction it is clear that “positive waiting time” means waiting time beyond what is necessary for full capacity utilization.

Consider next the case with welfare weights that differ across types. Starting with the case of no waiting time, we initially have TI(π)=0 and H=F. An increase in T (i.e. a reduction in π) will as before increase the term TI(π). However, as some persons now choose private treatment, H-F becomes positive. If the growth in p(H-F) is larger than the growth in TI(π) (plus the growth in (m-1)qy if m>1), the introduction of waiting time will reduce aggregate social costs S. In this case it will therefore be optimal to have a positive waiting time for treatment in the public sector.

In order for aggregate costs S to be lower with waiting time than without, it obviously must be the case that the costs b(θ,T,t) for some types must be lower with than without waiting time. From (7) we know that b is increasing in θ, and that b=q for all types when T=0. If b(θ,T,t)<q for some positive value of T for some types, this inequality must certainly hold for the type with the lowest waiting cost, i.e. for θ=α. Since waiting time for public treatment is only relevant if p>αT, it follows from (7) that

$$b(\alpha,T,t) = q - (q+t)y(\frac{p(t)}{T}) + \alpha T \tag{17}$$

Inserting p/T=π and p=mq+t gives us

$$b(\alpha,T,t) = q - \left[\frac{q+t}{mq+t}\pi y(\pi) - \alpha\right]T \tag{18}$$

Whether or not b<q for some value of T depends on the sign of the term in square brackets. If this term is negative for all π it must be true that b>q if there is waiting time. If on the other hand this term is positive for some π, there will exist a positive waiting time giving b<q.

From this discussion, and remembering that m≥1 the following proposition follows:

Proposition 2: If πy(π)≤α for all π, it is not optimal to have a positive waiting time for public health treatment no matter what the welfare weights are.

Since with appropriately chosen welfare weights we can give as much weight to the type θ=α as we want, the following proposition must also hold:

Proposition 3: *If πy(π)>α for some π and m=1, it is optimal to have a positive waiting time for public health treatment if the welfare function gives sufficiently high weights to the types with the lowest waiting costs.*

Notice that if the tax rate t is chosen freely, the condition m=1 is not necessary for the result in proposition 3. The reason is that even if m>1, the term (q+t)/(mq+t) can be set as close to 1 as one wants by choosing a sufficiently high t. The choice of t and T for a given value of π is discussed in more detail in the next Section.

We conclude this Section with an example demonstrating the possibility of waiting time being optimal. Let $F(\theta)=\theta$ (implying $\alpha=0$ and $\beta=1$, and $f(\theta)=1$)[17], and $w(\theta)=h(\theta)=(1-\sigma)\theta^{-\sigma}$ where $0<\sigma<1$, implying $H(\theta)=\theta^{1-\sigma}$. It is then straightforward to see that

$$TI(\pi) = p\frac{I(\pi)}{\pi} = p\frac{1-\sigma}{2-\sigma}\pi^{1-\sigma} \tag{19}$$

and

$$H(\pi) - F(\pi) = \pi^{1-\sigma} - \pi \tag{20}$$

Inserting (19) and (20) into (15) and rearranging gives

$$S = q + (m-1)qy(\pi) + \frac{p\pi}{2-\sigma}\left[2-\sigma-\pi^{-\sigma}\right] \tag{21}$$

It is easily verified that for $0<\sigma<1$the term in square brackets is negative for some values of π (smaller than 1). If e.g. m=1, this means that S is lower with waiting time than without.

4. Waiting time and a tax or subsidy on private treatment

If the optimal waiting time is zero, everyone chooses public treatment, and the issue of a tax or subsidy on private treatment is irrelevant. Similarly if the optimal outcome is to completely privatize the health treatment under consideration. In the present model it makes no difference whether or not the private treatment is subsidized in this case, as the costs in either case are divided equally among everyone.

[17] Notice that since $\alpha=0$, the condition in Proposition 3 holds in this example.

The interesting case to consider is the case in which it is optimal to have a waiting time that makes some but not all persons choose private treatment. In the previous Section we saw that this could be the case when individuals are given different welfare weights in the social objective function.

To address the issue of a tax or subsidy on private treatment, it is useful to rewrite (15) as

$$S = q + (m-1)qy(\pi) + p\left[\frac{I(\pi)}{\pi} - (H(\pi) - F(\pi))\right] \qquad (22)$$

We showed in the previous Section that it is only if the term in square brackets is negative for some π that it is optimal to have waiting time in the public sector. In the opposite case the best we can do is have T=0 (and y=0) or y=1.

Assume that the term in brackets is negative for some π, and that there exists a value of $\pi \in (\alpha,\beta)$ that minimizes S for any given p. However, when we minimize S with respect to both T and t, there is no solution to the minimization problem. To see this, consider an arbitrary value π* that makes the square brackets in (22) negative. It then follows from (22) that for this value of π, S is lower the higher is p. We thus have the following proposition:

Proposition 4: Assume that it is optimal to have waiting time for public treatment. Start with an initial situation where the waiting time has a length implying lower social costs that the situation with no waiting time. From this initial situation, increasing the waiting time and a tax on private treatment simultaneously while keeping the ratio π=(mq+t)/T constant will always be welfare improving.

The reason we get this somewhat surprising result is that the demand for private treatment in this simple model remains unchanged as the price of this treatment increases, as long as π=p/T remains constant. In reality, a large increase in the price of taxed private treatment would lead to a reduction in demand. One form of such a demand reduction would be substitution towards untaxed private treatment, for instance treatment abroad.

5. Capacity limit on public treatment

Till now, we have assumed that the amount of treatment in the public sector was determined implicitly through the choice of waiting time in the public sector and the tax/subsidy of private treatment. In this Section we shall briefly consider the case in which the capacity of the public sector is given. We ignore the possibility of increasing this capacity through the direct purchase of private capacity by the public health authorities, cf. the discussion in footnote 5.

A given capacity in the public sector less than total demand means that a given part of the demand must be covered by the private sector. In our model this means that y is given by the capacity limitation of the public sector. Since the demand for private treatment is given by $y(\pi)$, this means that π is given. The government's decision problem is thus to choose T and t subject to $(mq+t)/T$ equal to the market clearing value of π. The government wants as before to make S given by (22) as small as possible. Since y now is given, this simply means making the last term in (22), i.e. the square brackets multiplied by p, as small as possible for the given value of π. The solution to this problem is straightforward: If the square brackets are negative, S is lower the higher t and T are for the given ratio $(mq+t)/T$. As mentioned in the end of the previous Section, one cannot take this result to literally, due to limitations of the model.

If the square brackets in (22) are positive (as they will be for e.g. equal welfare weights, implying H=F), the price p should be as small as possible. In practice this means that best the government can do is to subsidize the private sector by so much that the price of private treatment is zero. The waiting time in the public sector should also be set equal to zero. In this situation everyone is indifferent between the two alternatives, and in equilibrium the number of people treated in the public sector is equal to the capacity there.

As mentioned in Section 3, the discussion of T versus t applies also to whether or not one should pay for public treatment. If the choice is between a fee for public treatment and rationing by waiting time, our analysis shows that this depends on the sign of the square brackets in (22). If distributional considerations are ignored (or are small), this term will be positive. In this case it is better to use a fee for public treatment than to ration the public treatment by waiting time. However, if distributional considerations are sufficiently strong, the square brackets in (22) will be negative. In this case the best policy is to provide public treatment free of charge, but with waiting time for the treatment.

The discussion above may be summarized by the following proposition:

Proposition 5: For a given value of π, implied by a given capacity limit on public treatment, social costs will be lower the higher is the waiting time for public treatment and the tax on private treatment, provided $I(\pi)<\pi[H(\pi)-F(\pi)]$. If $I(\pi)>\pi[H(\pi)-F(\pi)]$ for the given value of π, the optimal waiting time is zero, and the optimal tax/price policy is to have an equal price for public and private treatment.

6. The optimal tax/subsidy when waiting time is given.

In this section it is assumed that the waiting time is determined by mechanisms outside the present analysis. We therefore consider waiting time as exogenous, and ask the question of what the optimal tax or subsidy is in this case. The answer to this question will of course depend on what the waiting time is. We assume that the waiting time is such that without any tax or subsidy, some but not all persons will choose private treatment.

Without loss of generality we set T=1, so that π=p and (22) may be rewritten as

$$S(p) = q + (m-1)qy(p) + I(p) - p\left[H(p) - F(p)\right] \qquad (23)$$

Since p=mq+t, choosing the optimal tax is equivalent to choosing the optimal price for private treatment.

Differentiating (23) and using y=1-F, F'=f and H'=h gives, after some manipulation:

$$S'(p) = (q+t)f(p) - \left[H(p) - F(p)\right] \quad (24)$$

From the definition of the function H (see equations (9)-(10)) we know that H(p)-F(p) approaches zero as p approaches its upper limit β. Moreover, since we have assumed that without a tax or subsidy y is positive, p can only be close to β if t>0. Since f(p)=F'(p)=-y'(p) it therefore follows from (24) that

$$S'(p) \approx -(q+t)y'(p) > 0 \qquad \textit{for p below, but close to } \beta \quad (25)$$

This inequality implies that if we start by considering a prohibitively high tax, making p=β, social cost S can be reduced by reducing the tax rate. We therefore have the following proposition:

Proposition 6: It can never be optimal to have a tax rate that is so high that no one wants private health care. Equivalently, it cannot be optimal to use legal regulation to prohibit private health care.

The result above is valid no matter what one assumes about the welfare weights. For further results, it is useful to distinguish between the cases where welfare weights are equal for all types, and where the weights are declining in θ.

Consider first the case of equal welfare weights for everyone, implying H=F for all p. In this case it follows from (24) that

$$S'(p) = -(q+t)y'(p) \qquad \textit{if } w(\theta) \textit{ is independent of } \theta \quad (26)$$

It is clear from (26) that S' is positive for t>-q and p∈(α,β). Remembering that p=mq+t, the following proposition therefore follows

Proposition 7: For the case when everyone has the same welfare weights, it is optimal to subsidize private treatment. The optimal subsidy, denoted –t, is given by –t*=min[α,q].*

In other words, the optimal subsidy is at the most equal to q, but lower than q if a lower subsidy is sufficient to induce everyone to choose private health care (implying p so low that y=x and y'=0). Since we have assumed that mq>α (i.e. some persons choose public treatment if there no subsidy on private treatment), it is clear that –t*=α if m=1.

Consider next the case of unequal welfare weights. In this case H(p)>F(p) for t=0. The sign of S'(mq) is therefore generally ambiguous. Moreover, as there is no reason to believe that S(p) is convex, we would not know the sign of the optimal t even if we knew the sign of S'(mq). The sign of the optimal t depends on S' in the following way:

Remark 1: If S'(p) is positive (negative) for all p larger than (smaller than) or equal to mq, it is optimal to subsidize (tax) private treatment.

It is useful to consider S'(p) in some more detail. Rewrite H-F as

$$H-F=\frac{(1-F)-(1-H)}{1-F}y=(1-\sigma)y \qquad (27)$$

where

$$\sigma=\frac{1-H}{1-F}=\frac{\int_p^\beta w(\theta)f(\theta)d\theta}{\int_p^\beta f(\theta)d\theta} \qquad (28)$$

In words, σ is the weighted average of the welfare weights of those who choose private treatment. Since it is the types with high θ, i.e. low $w(\theta)$, who choose private treatment, σ must be less than one. Notice that σ depends on p, and thus on t.

Using (28), we can rewrite (24) as

$$S'(p) = -(q+t)y'(p) - (1-\sigma)y \tag{29}$$

For y>0 (which is always optimal) it is useful to introduce the demand elasticity (measured positively) for private treatment, denoted by ε(t):

$$\varepsilon(t) = -y'(mq+t)\frac{mq+t}{y(mq+t)} \tag{30}$$

Using this elasticity, (29) can be rewritten as

$$S'(p) = \left[\varepsilon(t) - \frac{mq+t}{q+t}(1-\sigma(t))\right]\frac{(q+t)y(mq+t)}{mq+t} \tag{31}$$

Since m≥1, the following proposition follows immediately from Remark 1 and eq. (31):

Proposition 8: if ε(t)<1-σ(t) for all –q<t≤0, it is optimal to have a positive tax on private treatment. For m=1 the "opposite" is also true: If ε(t)>1-σ(t) for all t≥0, it is optimal to subsidize private treatment.

Also for the general case of unequal welfare weights it is possible for the corner solution with no public treatment to be optimal:

Proposition 9: If S'(p) given by (31) is positive for all p∈(α,β), the optimal solution is characterized by p=α, i.e. the optimal subsidy is given by –t=mq-α.*

From (31) we also find the following condition for the optimal tax/subsidy when we have an interior solution, i.e. mq+t*∈(α,β)):

Proposition 10: If it is optimal to have some public health care, the optimal value of t, denoted by t, is given implicitly by*

$$\varepsilon(t^*) = 1 - \sigma(t^*) \qquad \textit{for } m=1 \tag{32}$$

and by

$$t^* = -q\frac{m(1-\sigma(t^*)) - \varepsilon(t^*)}{(1-\sigma(t^*)) - \varepsilon(t^*)} \quad \text{for } m>1 \tag{33}$$

7. Concluding remarks

Our analysis gives several important results regarding the optimal size of waiting time for public health treatment (in particular, zero or positive) and of a tax/subsidy of private health treatment. Several of the results depend strongly on the properties of the welfare weights in the social welfare function. The properties of these welfare weights obviously depend on how strong the concern for equity is. With no such concern the welfare weights will be the same for all, and for this case we conclude that it is not optimal to have any waiting time for public treatment (proposition 1 and 5). In this case everyone will choose public treatment if there is no capacity limit on this treatment. If there is a capacity limit, one should subsidize private treatment and/or have a fee on public treatment such that the consumer price of the two options is identical (proposition 5). Similarly, if waiting time for public treatment is given exogenously by mechanisms not included in our model, one should subsidize private treatment (proposition 7).

The more interesting case is the one where there is a concern for equity. In this case the welfare weights will depend on the how strong this concern is. However, the properties of the welfare weights will also depend on what the possibilities and efficiency costs are for redistributing income through the tax system. In the hypothetical case where perfect lump-sum taxation is feasible, equity concerns can be fully taken care of through the tax system. In this case welfare weights in our analysis will be the same for everyone, and we get the same results as we found for the case of no concern for equity. In the more realistic case where any redistributive taxes are distortionary, the welfare weights will be given as a byproduct of the solution to the problem of designing an optimal tax system. The higher are the distortionary costs of redistributing income through the tax system, the more disperse will the equilibrium welfare weights in our analysis be. If welfare weights are sufficiently disperse, it may be optimal to have a positive waiting time for public treatment (proposition 3 and 5). In this case it is also optimal to have a positive tax on private treatment (proposition 4 and 5). If waiting time is given exogenously, the sign of the optimal tax/subsidy is ambiguous for the case where welfare weights differ across types (propositions 8-10).

In the discussion above, we have implicitly argued as if differences in types, measured by waiting costs, are perfectly correlated with ability, and thus income, in the optimal tax problem. As discussed in Section 2, we would expect the variation in waiting costs to be correlated to income variations, but waiting costs are also likely to vary among individuals for other reasons. It is not unlikely that "other reasons" are more important for health problems that prevent persons from doing specific activities (work or leisure related) than health problems that to a larger extent simply involve pain or discomfort that is more or less the same for everyone. If this is the case the types θ in our analysis will be less correlated with income in the former type of health problem than in the latter. It therefore seems reasonable that the dispersion of the size of the welfare weights should be less for the former type of treatment than for the latter. An implication of this is that it can be rational to let policies regarding waiting time and taxes/subsidies to private treatment differ between different types of health problems.

References:

Besley, T., Coate, S., 1991. Public provision of private goods and the redistribution of income. American Economic Review 18 (4), 979-984.

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

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

Boadway, R., Marchand, M., 1995. The use of public expenditures for redistributive purposes. Oxford Economic Papers 47 (1), 45-59.

Brekke, K., Sørgård, L., 2000. Where are all the physicians? Private versus public health care. Mimeo, University of Bergen.

Cullis, J.G., Jones, P.R., 1985. National Health Service waiting list: A discussion of competing explanations and a policy proposal. Journal of Health Economics 4, 119-135.

Department of Health, UK, NHS, 1999. http://www.doh.gov.uk/waitingtimes/booklist.htm.

Hagen, T.P., Østtveiten, H.S., 1999. Social security benefits under the knife. NIBR Report 16.

Hamilton, B.H., Hamilton, V.H., Mayo, N.E., 1996. What are the costs of queuing for hip fracture in Canada? Journal of Health Economics 15 (2) 161-185.

Hoel, M., Sæther, E.M., 2000. Private health care as a supplement to a public health system with waiting time for treatment. Memorandum from the Department of Economics, no 41/2000.

Hørding, U., Pederson, K.H., Holsteen, V., 1982. Frequency of unwanted pregnancies among women on waiting lists for sterilization. Ugeskrift for læger 144, 875-876.

Iversen, T., 1997. The effect of private sector on the waiting time in a national health service. Journal of Health Economics 16, 381-396.

Lindsay, C.M., Feigenbaum, B., 1984. Rationing by waiting list. American Economic Review 74, 404-417.

Marchand, M., Schroyen, F., 2000. Markets for public and private health care: redistribution arguments for a mixed system. Discussion paper 20/00, Norwegian School of Economics and Business Administration.

Mobley, L., Magnussen, J., 1998. "An internal comparison of hospital efficiency: does institutional environment matter? Applied Economics 30, 1089-1100.

Norwegian Patient Register, SINTEF UNIMED. http://www.npr.no/

Plump, J., Redekop, W.K., Dekker, F.W, vanGeldrop, T.R, Haalebos, M.M.P., Kingma, G.J., Zijlstra F., Tijssen, J.G.P., 1999. Death on the waiting list for cardiac surgery in the Netherlands in 1994 and 1995. HEART 81, 593-597.

Rossvoll, I., Benum, P., Bredland, T.R., Solstad, K., Arntzen, E., Jorgensen, S. 1993. Incapacity for work in elective orthopedic-surgery – A study of occurrence and the probability of returning to work after treatment. Journal of Epidemiology and Community Health 47, 388-394.

Roy, C.W., Hunter, J., 1996. What happens to patients awaiting arthritis surgery? Disability and Rehabilitation 18 (2), 101-105.

Stern, G., Brown, R., 1994. The effect of a waiting list on attendance at initial appointments in a child and family clinic. Child Care Health and Development 20 (4), 219-230.

Worthington, D., 1987. Queueing models for hospital waiting lists. Journal of Operations Research Society 38, 413-22.