DIRITTI, REGOLE, MERCATO
Economia pubblica ed analisi economica del diritto

Pavia, Università, 3 - 4 ottobre 2003

Società italiana di
economia pubblica

XV
CONFERENZA

Horizontal and vertical cream skimming in the health care market

Rosella Levaggi * Marcello Montefiori**

*Dipartimento di Scienze Economiche, Università di Brescia
**DISEFIN - Sezione Economia Pubblica, Università di Genova

1. Introduction

In western countries a substantial proportion of expenditure for health care is financed by the public sector. Since the first oil crisis in 1975, and more recently from the '90's the objective to rationalise and control expenditure has become a priority for any effective government policy.

Health care systems have been widely reformed and in most countries a separation between purchasing and delivering health care has been put forth to mimic the structure of a competitive market. However, this market presents peculiarities that prevent reaching a first-best solution.

Purchasing at government level presents problems from the point of view of the incentives for both parties to act in a competitive way. The structure proposed for the internal market is only virtually competitive because the separation of functions does not imply a separation of interests as in the competitive market. Laffont and Tirole (1995) show that procurement contracts might not be incentive compatible because of collusion since these agencies are judged on their performances, but they do not receive the surplus of their cost minimisation actions. Levaggi (1999) shows that in procurement contracts, like those for health care, where a binding budget constraint exists, technical problems prevent the implementation of an optimal incentive structure.

Even if we do not take into account the possibility of collusion between the purchaser and the provider, achieving an optimal allocation is particularly difficult.

In the market for health care all the actors do not necessarily share the same objectives. Levaggi (2003) shows that in the internal market for health care several actors coexists:

- private organisation working for the public sector;
- public hospitals independent form the purchaser; they usually are non profit institutions with their own management. They have the right to withhold information from the purchaser as concerns their running costs and can pursue their own objectives;
- hospitals that are directly controlled by the purchaser; they are public firms with a high degree of independence as far as their organisation is concerned, but the purchaser can impose them to reach some common objectives;
- hospitals that are directly managed by the purchaser; they do not have the ability to withhold information from the purchaser that can make them to pursue its own objectives;

The outcome in terms of costs and quality is rather different according to the type of providers[1].

Furthermore, the provision of health care services is characterised by uncertainty and asymmetry of information in the cost of the treatment. Asymmetry of information exists because while at the time the contract is set both the purchaser and the provider share the same information set, the purchaser can observe some relevant information (freely or at a cost) before making his effort.

The ability to observe patients' severity can be used for self-interest-advantage by the hospital (principal-agent) through two alternative behaviours:

- they can choose to treat only patient with specific diseases("horizontal" cream skimming);
- they can affect the state-of-the-world probability distribution opting for specific "patient type" within the same ailment group ("vertical" cream skimming).

These behaviours will be defined as "market cream skimming" and they alter the competition among hospitals causing relevant effects in the whole market system. The literature has studied this problem in the context of a private market (Ellis, 1997, Lewis and Sappington,1999) but only few authors have approached the problem from the standpoint of a public provider[2].

Horizontal cream skimming is a legal practice, although it goes against the principle of universal access to health care and it might have perverse effects on the finance of the whole health care system. It is however interesting to note that it arises form a regulatory problem, i.e. the regulator has not set prices correctly and the hospital finds it convenient to specialise in some outputs. This behaviour in fact arises from the inability of the regulator to observe the reaction function to a predetermined payment scheme.

Vertical cream skimming is instead an illegal behaviour that consists in offering health care only to the patients that have a low cost. It arises from the inability of the purchaser to observe the patient type and it might be solved through control and sanctions rather than incentives.

All these elements determine the structure of the market and its degree of competitiveness. In this article we want to concentrate on asymmetry of information and in particular on the problems arising from cream skimming. The model presented here assumes that private and independent hospitals share the same objectives so that their behaviour can be modelled in the same way.

The objective of the paper is to design contracts that allow to reduce the problems arising from cream skimming. The design of the scheme is made from the standpoint of

---

[1] See Levaggi (2003)

[2] See Levaggi (2002 a,b) for a review.

a benevolent regulator that appoints a purchaser (P) to provide hospital care to his population at the least possible cost. To do so, P has to employ two hospitals that have fixed locations. One of the two hospitals is independent from P while the other one is directly controlled. Both the quality and the cost cannot be observed, but the former can be inferred through the choice of patients. In this paper it will be assumed that the two hospitals compete for patients according to the rules of Hotelling competition and they react to the payment system offered by P by fixing quality to the level that allows them to maximise their objective function. The paper will be organised as follows: in section two the model is presented; in section three horizontal cream skimming is discussed and in the following section vertical cream skimming is presented. Section five discusses the results and section six presents the conclusions of the analysis.

2. The model

The model presented here draws on Levaggi (2003) and Montefiori (2003). It develops a three stage game from the standpoint a purchaser (P) to buy health care from two providers that are located at the extremes of a line whose distance has been normalised to one. One of the two providers is an independent structure (a trust or a private hospital) while the other is a directly controlled one[3]. The patients lie within this line and are uniformly distributed and can be affected by one of two ailments (1 and 2); their number is normalised to one, so that demand can be interpreted in terms of market shares. The distribution of the ailments is uniform, i.e. each disease will account to ½. In the first stage of the game, the effort of the management is defined through cost minimisation and the payment is defined for a set quality level; in the second stage, the two hospitals compete for patients through quality using the rules of the spatial competition à la Hotelling. The results of the second stage allows P to define a relationship between reimbursed and delivered quality which will be used in the third stage to define the payment scheme that allows to provide health care to all the patients on the line.

2.1 The cost for health care: risk sharing properties of the contracts

The cost incurred by the hospital to produce health care is assumed to be a linear function of quality, patients' characteristics and the effort of the medical staff. The unit cost function can be written as:

$$C_{ij} = \beta_{ij} + q_j - e_{ij} \tag{1}$$

where

---

[3] In Italy it would correspond to aziende ospedaliere.

j=1, 2 (ailment) and i=low cost, high cost (i.e. the patient type), q^* is the quality level, e_i is the effort of the medical staff[4] and β_i is a patient-related cost.

For the second disease β_{i2} can assume two values, L and H(*L=low cost; H=high cost*) with $\beta_{L2} < \beta_{H2}$; both events have a known probabilities equal to *p* and *(1-p)* respectively. For disease 1, there is no variance in the cost of provision, i.e. $\beta_{L1} = \beta_{H1} = \overline{\beta_1}$.

Quality is a multidimensional vector that includes medical and non-medical variables that affects the outcome of health care, i.e. prevention, treatment and aftercare[5]. Both medical and non-medical quality are extremely relevant to determine patient utility. At this stage we assume quality to be a scalar including all the relevant aspects of interest; when asymmetry of information is introduced we will explicitly model the characteristics of quality[6].

The cost for the two diseases can then be written as:

$$C_1 = \overline{\beta_1} + q_1 - e_1$$
$$C_{i2} = \beta_{i2} + q_2 - e_{2i} \qquad\qquad i = L, H$$

(2)

The effort produces a disutility that is linear in the number of patients, but increasing in the effort , i.e.

$$f(e,n) > 0; \qquad f'_e(e,n) > 0; \qquad f''_e(e,n) > 0; \qquad f'_n(e,n) > 0; \qquad f''_n(e,n) = 0$$

The hospital management participates to the production process only if the reward received, net of the cost of production, produces a positive utility. We assume that this condition is binding for each ailment separately:

$$U_1(t_1 - C_1 - f(e_1)) \geq 0$$
$$EU_2(t_{2i} - C_{2i} - f(e_{2i})) \geq 0$$

(3)

The hospital is assumed to be risk neutral. This means that the utility function is linear in U, i.e. $U'_i > 0; U''_i = 0$

For a given level of quality, the regulator has to set t_1 and t_2 so to minimise cost and make hospital participate to the production process. In this paper we will assume that Central Government shifts all the risk implicit in the uncertainty on the cost of provision on the provider by foreseeing a prospective payment scheme. The problem can be written as:

---

[4] q^* is not the actual quality offered by the hospital, but there is a strict and observable relationship between the quality reimbursed and the one that is actually delivered which derives from the second stage of the game. For the cost minimising stage this problem in not important since it does not alter the results of the analysis.
[5] The same approach is used by Chalkley and Malcomson (1998).
[6] In next section of this paper we will explicitly model the two most important characteristics of quality: medical and non-medical.

$$Min \quad t_1 + t_2$$
$$C_1 = q_1 + \overline{\beta_1} - e_1$$
$$C_{2i} = q_2 + \beta_{2i} - e_i \qquad\qquad i = L, H \qquad\qquad\qquad (4)$$
$$U_1(t_1 - C_1 - f(e_1)) \geq 0$$
$$EU_2(t_2 - C_{2i} - f(e_{2i})) \geq 0$$

The solution is presented in appendix 1 and recorded in table one.

Tab. 1: Optimal contract: symmetry case and q*=0.

	Ailment 1	Ailment 2
Condition on effort	$f'(e_1) = 1$	$f'(e_{2i}) = 1$
Condition on utility	$U_1(t_1^* - C_1^* - f(e_1^*)) = 0$	$pU_2(t_2^* - C_{2L}^* - f(e_{2L}^*)) +$ $(1-p)U_2(t_2^* - C_{2H}^* - f(e_{2H}^*)) = 0$
Payment scheme	$t_1 = f(e^*) + C_1.$ [7]	$t = E(C_2) + E[f(e)]$

For the first ailment, since there is no uncertainty, the hospital is reimbursed for the cost that it bears. For the second one, he is reimbursed using a prospective payment scheme that was first advocated by Ellis and McGuire (1986) and other authors who widely criticised cost reimbursement for its cost inflation properties and for not being able to make hospitals produce an efficient level of care for their patients. Prospective payment solutions have however a number of drawbacks.

Prospective payments can be used only for large contracts where the average actual cost will be close to its expected value. When these conditions are not met, the risk of using a prospective payment system is that the hospital might go bankrupt even if it is efficient[8]; Chalkley and Malcomson (2002) show that prospective contracts have in fact higher cost than cost reimbursement one. An explanation for this apparently incongruous result can be found in Levaggi (2003a): the hospital observes before treating the patient his severity degree and accepts a prospective payment only if the reimbursement allows the to get a non negative utility also in the worst possible state of nature. Prospective payment systems are also one of the causes for opportunist behaviour as shown in the following sections.

3. Quality determination under horizontal cream skimming
Once the reimbursement scheme has been defined, the private hospital maximises his surplus through the number of cases they admit and are reimbursed for. The number of

---

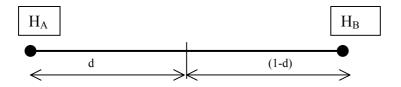[7] It is assumed that $U_i = 1$ to simplify the algebra.
[8] Or it might have a surplus and being inefficient.

cases to be treated for each ailment depends on the behaviour of patients that can choose where to receive their treatment.

3.1 The choice of the patients

We start to describe the process leading to quality determination by presenting the choice of the patients.

Since the two diseases are independent, we can study the behaviour of the patient just with reference to one of them. The two hospitals that serve the market have a fixed location and are placed at the two extremes of a line of length one:

The service is free at the point of use so that patient's utility depends on the quality and on travelling costs which are linear in the provider's distance from patient's location.

$$U_{pz} = \begin{cases} \alpha q_a - \gamma d \\ \alpha q_b - \gamma(1-d) \end{cases} \tag{5}$$

where α denotes how quality enters the utility function and γ denotes the disutility because of the distance (d). Patients will be indifferent between hospital A and hospital B when: $\alpha q_a - \gamma d = \alpha q_b - \gamma(1-d)$.

Solving for d we get:

$$d = \frac{\alpha}{2\gamma}(q_a - q_b) + \frac{1}{2} \tag{6}$$

The distance d represents the demand for hospital A and, at the same time, the location of the marginal consumer. The demand for hospital j can be obtained multiplying the distance by the density which, given the unit length of the line is equal to ½. The demand for each hospital can then be written as:

$$D_j = \left[\frac{\alpha(q_a - q_b)}{2\gamma} + \frac{1}{2} \right] \frac{1}{2} \tag{7}$$

3. 2 The reaction function of the hospital to quality reimbursed

The hospital that is directly controlled (A in our case) sets quality to the level reimbursed by the provider, i.e. $q_{1A} = q_1^*$; $q_{2A} = q_2^*$

Hospital B is the independent structure that chooses q_{1b} and q_{2b} in order to maximise the following surplus function:

$$\Pi_B = (q_1^* - q_{1B})[\frac{\alpha}{4\gamma}(q_{1B} - q_1^*) + \frac{1}{4}] + \left\{ \begin{array}{l} p[(q_2^* - q_{2B}) + E(\beta) - \beta_l - \sigma_\beta^2] \\ + (1-p)[(q_2^* - q_{2B}) + E(\beta) - \beta_h - \sigma_\beta^2] \end{array} \right\} * [\frac{\alpha}{4\gamma}(q_{2B} - q_2^*) + \frac{1}{4}]$$

(8)

Hospital surplus is increasing in the average reimbursement [$E(\beta)$] and decreasing in its variance (σ_β^2). This assumption allows us to model a trade off between average reimbursement and its variance. Both ailments considered allows to break even for a set quality level, but one of the two is more risky than the other one. Even if the hospital is risk neutral, his behaviour will be affected by the uncertainty of the payment and it is reasonable to assume that a trade off exists between expected reimbursement and its variance.

The F.O.C. for the problem can be written as:

$$\frac{\partial \Pi_B}{\partial q_{1B}} \qquad q_{1B} = q_1^* - \frac{\gamma}{2\alpha}$$

(9a)

$$\frac{\partial \Pi_B}{\partial q_{2B}} \qquad q_2 = q_{2B}^* - \frac{\gamma}{2\alpha} - \frac{\sigma_\beta^2}{2}$$

(9b)

The first order conditions shows that private hospital uses his position rent to increase the surplus. Equation (9b) shows that only part of the payment made by P for quality is in fact passed onto the consumer. The fraction $\gamma/2\alpha$ kept by the hospital is his position rent. For the second ailment, the hospital keeps the variance of β as well. This term represents the premium for the risk that the hospital has to bear.

3.2 The quality level to be reimbursed by the purchaser

The purchaser chooses the level of q_i^* for which all the patients are treated at the minimum cost. This condition requires that the marginal patient gets zero utility when admitted.

Starting from patient demand and solving the following equation system, we can derive the quality condition, with reference to disease 1, in order to provide non negative utility to the marginal patient.

$$\begin{cases} q_1^* = \dfrac{\gamma}{\alpha} d_a \\[2mm] q_{1b} = q^* - \dfrac{\gamma}{2\alpha} = \dfrac{\gamma}{\alpha} d_b \\[2mm] d_a + d_b = 1 \end{cases} \rightarrow \begin{cases} q_1^* = \dfrac{3}{4}\dfrac{\gamma}{\alpha} \\[2mm] q_{1b} = \dfrac{1}{4}\dfrac{\gamma}{\alpha} \end{cases} \tag{10}$$

The hospital directly controlled will cover ¾ of the unit length market while the residual ¼ will be served by the private one.

The same procedure can be used in order to define quality and market shares with respect to the second disease

$$q_2^* = \frac{1}{4}\left(\frac{3\gamma}{\alpha} + \sigma_\beta^2\right)$$

$$\tag{11}$$

$$q_{2b} = \frac{1}{4}\left(\frac{\gamma}{\alpha} - \sigma_\beta^2\right)$$

Market shares will turn out to be $\dfrac{1}{4}\left(3 + \dfrac{\alpha}{\gamma}\sigma_\beta^2\right)$ and $\dfrac{1}{4}\left(1 - \dfrac{\alpha}{\gamma}\sigma_\beta^2\right)$ respectively for hospital

A (directly controlled) and B (private).

The private hospital has a preference for ailment 1, i.e for the disease with no uncertainty on patient type. The quality level for ailment 2 takes into account the uncertainty concerning the treatment cost and, as a result, an additional market share, proportional to β variance, will be given up for public hospital. The private hospital exploiting the horizontal product differentiation, earns positive total surplus equal to:

$$\Pi_B = \frac{1}{8}\left[\frac{\gamma}{\alpha} + \sigma_\beta^2\left(\frac{\alpha}{2\gamma}\sigma_\beta^2 - 1\right)\right] \tag{12}$$

Equation (12) shows that the surplus increases in the variance of β as one might expect. γ represents a good proxy of hospital monopolistic rent given to product horizontal differentiation. For a larger γ value, the private hospital quality increases, i.e. a higher quality level is required to meet patient participation constraint.

The public hospital is forced to provide the quality q^* and its demand is residually determined by private-hospital quality decision. The absolute value of its surplus is directly correlated to the market share they serve. The strategic behaviour of the private hospital is to reduce its market share in presence of uncertainty because it shirks risk. The public hospital cannot act in the same way and it is forced, given its "social welfare rule", to cover the residual part of the unit length market line. The payment system secures that the hospital will break even in expected terms.

The analysis of this scenario shows further interesting features. Firstly, given the model structure, q^* represent the reimbursement the purchaser sets in order to meet the marginal-consumer-participation-constraint given a reservation utility equal to zero. Secondly, the reimbursement scheme has to be varied depending on the type of ailment. The purchaser pays an additional price because of uncertainty. The private hospital shirks risk. The bigger results the β variance, the smaller turns out to be the market share the hospital chooses to cover. The intuition of it is straightforward: quality level is the instrument to control the demand, but quality is expensive. Thus a lower quality, given reimbursement uncertainty because of the variance in patient-type, reduces risks of treatment cost above the average. The private hospital can opt for this behaviour because of a la Hotelling spatial differentiation: even through a low quality, it can attract a portion of the market. The bigger the variance, the lower the market portion will be, because quality is directly connected to cost and variance to reimbursement risk.

4. Asymmetry of information

In this section we consider the effects of the introduction of asymmetry of information on cost to provide health care. The cost of health care provision, as much as the outcome in terms of improved health, depends on characteristics of each patient that are observed only by the hospital and are used by this agent to pursue the maximisation of his objective function [9]. The timing of information in the contract can be summarised as follows:

Fig.1: The timing of information

HA sets contract   realisation of β outcome

hospital observes β effort by hospital

When the contract is stipulated both parties have the same information on β, but the hospital can observe it before setting his effort and can hide it from the purchaser. If the hospital is under direct management this parameter can be observed by both agents and the game becomes of symmetric information.

When the hospital (agent) gets better information with respect to the purchaser (principal), then the asymmetry in information plays a crucial role determining the well known principal-agent problem.

In the usual approach to this problem, the hospital would have an incentive to an opportunistic behaviour by cheating on the state of the world that is private information.

---

[9] See, for example, Levaggi (2002 a,b)

In this case, an incentive compatible scheme of the type described in Levaggi (2003a, 2003b) could be sufficient to solve the problem. However, in health care the issue is more complicated; the hospital can in fact:

- offer a different quality level to patient of a different type (cream skimming),
- choose patients according to their severity degree (cream skimming or dumping)

If the hospital knows the patient-type before the treatment, the risk of vertical cream skimming (choosing patients whose recovery rate is expected to be higher) takes place. Consequently high risk patients will be rejected by the private hospital and will have to be treated by the public hospital. Since the payment scheme is designed to break even on the expected average costs, this means that the private hospital will make a surplus while the public hospital will have a deficit.

Choosing patients is however an illegal behaviour and the hospital that wants to pursue this strategy has to be very careful in the way it undertakes it. To introduce this aspect in our model, we explicitly model the multidimensional aspect of quality that was pointed out in the previous paragraph. The medical quality typically includes aspects like appropriateness, health, nursing, aftercare, etc., while the non-medical quality includes comfort, information, kindness, catering service and so on. We will assume the two quality as scalars resulting from the weighted sum of the afore-mentioned variables. In particular we assume that the quality delivered by the hospital is made up by three elements:

- a minimum quality level that is implicit in the definition of β and that determines the minimum level of medical resources that are necessary to make the treatment effective
- hotel services $(\hat{q})$ which comprises all those services that are not strictly medical, but that can improve patient's stay in hospital. They are services such as the number of beds per room, hours of visits, private telephones, nurses per ward, ect.
- health related services $(\tilde{q})$ that might improve the quality of the care delivered.

q is then the sum of the two last components:
$$q_i = \hat{q}_i + \tilde{q}_i$$

They are assumed to be perfect substitutes and their cost of provision is constant and equal to one to simplify the exposition. With this new hypothesis the cost for hospital services can be written as:

$$C_1 = \overline{\beta_1} + \hat{q}_1 + \tilde{q}_1 - e_1$$
$$C_{i2} = \beta_{i2} + \hat{q}_2 + \tilde{q}_2 - e_{2i} \qquad\qquad i = L, H \tag{13}$$

For the first type of ailment, given that the cost for health care is uniform across patients, the hospital is indifferent about the choice of hotel and health quality; for this reason the quality will still be denoted by q.

As per the second type of ailment, we assume that the patients with a different cost are not uniform with reference to their preferences for quality. In particular, we assume that those that have a higher degree of severity (hence of costs) are more interested in health related quality. The perceived level of quality for the second ailment by the two groups can then be written as:

$$q_{2L} = w\hat{q}_2 + (1-w)\tilde{q}_2$$
$$q_{2H} = (1-w)\hat{q}_2 + w\tilde{q}_2$$

(14)

where ½<w<1

Public hospitals are assumed to be unable to discriminate among patients. It is also reasonable to assume that they will invest only in health related quality so that for the public hospital the following relation holds:

$$q_{2A} = \tilde{q}_{2A} = q_2^*$$

The second equality means as before that public hospital share his objectives with the purchaser and will then pass on to the patient any quality level that is reimbursed by the provider. The private hospital wants instead to maximise his surplus. The problem in this new scenario can be written as:

$$Max \qquad \Pi = (q_1^* - q_{1B})[\frac{\alpha}{4\gamma}(q_{1B} - q_1^*) + \frac{1}{4}]$$

$$+ [q_2^* - \tilde{q}_{2B} - \hat{q}_{2B} + E(\beta) - \beta^L]p[\frac{\alpha}{4\gamma}[w\hat{q}_{2B} + (1-w)(\tilde{q}_{2B} - q_2^*)] + \frac{1}{4}] +$$

$$+ [q_2^* - \tilde{q}_{2B} - \hat{q}_{2B} + E(\beta) - \beta^H](1-p)[\frac{\alpha}{4\gamma}[(1-w)\hat{q}_{2B} + w(\tilde{q}_{2B} - q_2^*) + \frac{1}{4}]$$

s.t

(15)

$$q_{1B} \geq 0$$
$$\hat{q}_{2B} \geq 0$$
$$\tilde{q}_{2B} \geq 0$$

The problem can be solved using Kuhn Tucker conditions. To simplify the algebra and without loss of generality we assume that high and low cost patients follow a uniform distribution, i.e. p=½ [10]. In appendix 2 it is shown that if the following condition holds

---

[10] The more general solution is presented in appendix 2.

$\frac{\gamma}{\alpha} - (\beta^H - \beta^L)(\frac{2w-1}{4}) \le q_2^* < \frac{\gamma}{\alpha} + (\beta^H - \beta^L)(\frac{2w-1}{4})$ we will have a corner solution, i.e. the health quality $\tilde{q}$ will be zero[11]. In this case, the solution will be equal to:

$$\tilde{q}_{2B} = 0 \tag{16}$$

$$\hat{q}_{2B} = q^* - \frac{\gamma}{\alpha} + (\beta^h - \beta^l)(\frac{2w-1}{4}) \tag{17}$$

For hotel quality, the non negativity constraint is satisfied. It interesting to note that the third term on the RHS of (17) is always positive, provided that w> ½ as it was assumed from the start.

4.1 Quality determination for vertical cream skimming

In the third stage of the game the provider determines the level of quality q* to be reimbursed in order to have the market fully covered. The purchaser has however to set this parameter in a context of information asymmetry. In the previous paragraph we noted that the private hospital can skim on patients by the choice of an optimal mix for the two qualities; this action goes however undetected because the provider cannot separate the two types of quality. In other words, we assume that the provider can predict the reaction of hospitals to the quality level reimbursed, but it cannot monitor the split between hotel and health quality. To show this, let us define the optimal quality to be reimbursed as:

$$q_2^* = \frac{\gamma}{\alpha} - \frac{\vartheta}{2} \tag{18}$$

where:

$$\vartheta = (\beta^H - \beta^L)(\frac{2w-1}{4})$$

If we substitute in (17) we get:

$$\hat{q}_{2B} = \frac{\vartheta}{2} \tag{19}$$

The private hospital demand for low and high cost patients can be written as:

$$d_{2B}^L = \frac{1}{8}\left\{\frac{\alpha\vartheta}{2\gamma} + w\right\} \tag{20}$$

$$d_{2B}^H = \frac{1}{8} + \frac{1}{8}\left\{\frac{\alpha\vartheta}{2\gamma} - w\right\} \tag{21}$$

---

[11] The interval of q* has been computed setting, alternatively, hotel and health quality to zero. For values greater than zero, the afore-mentioned interval would increase.

The analogous demand for the public hospital can be written as:

$$d_{2A}^{L} = \frac{1}{4} - \frac{1}{8}\left\{\frac{\alpha\vartheta}{2\gamma} + w\right\} \tag{22}$$

$$d_{2A}^{H} = \frac{1}{8} - \frac{1}{8}\left\{\frac{\alpha\vartheta}{2\gamma} - w\right\} \tag{23}$$

The share of the patients among the two hospital depend on the weight w the patients assign to hotel quality. If patients were indifferent between the two types of quality (w=1/2) we would in fact get[12] $d_{2A}^{L} = d_{2A}^{H} = \frac{3}{16}$ and $d_{2B}^{L} = d_{2B}^{H} = \frac{1}{16}$.

The scope for cream skimming is directly related to the value of w. For w=1, the market shares for low and high cost patients treated by the private hospital will be equal to:

$$d_{2B}^{L} = \frac{1}{8}\left\{\frac{\alpha}{2\gamma}\frac{(\beta^{H} - \beta^{L})}{4} + 1\right\} \tag{24}$$

$$d_{2B}^{H} = \frac{1}{8}\left\{\frac{\alpha}{2\gamma}\frac{(\beta^{H} - \beta^{L})}{4}\right\} \tag{25}$$

In this case, the private hospital gets a further 25% of the market share for patients that are low cost.

Substituting back in the surplus function the equilibrium values for $q_2^*; \hat{q}_{2B}; \tilde{q}_{2B}; d_{2B}^{L}; d_{2B}^{H}$, we can find the surplus with reference to ailment two:

$$\Pi_{B} = \frac{1}{8}\left(\frac{\gamma}{\alpha} + 2\vartheta - \frac{\alpha}{\gamma}\vartheta^{2}\right) \tag{26}$$

For w==> 1/2==> ϑ=0 ==> $\Pi = \frac{\gamma}{8\alpha}$

For w=1 ===> $\vartheta = \frac{\beta^{H} - \beta^{L}}{4}$ ===> $\Pi = \frac{\gamma}{8\alpha} + \frac{(\beta^{H} - \beta^{L})}{16}[1 - \frac{\alpha}{8\gamma}(\beta^{H} - \beta^{L})]$

In this case, cream skimming produces his best result in terms of increase in surplus, provided that $(\beta^{H} - \beta^{L}) < \frac{8\gamma}{\alpha}$. The surplus for the private hospital is maximum for $(\beta^{H} - \beta^{L}) = \frac{4\gamma}{\alpha}$.

For the first ailment, the reaction function of the private hospital is the same as in the previous section so that its total profit can be written as:

$$\Pi = \frac{1}{8}[\frac{3\gamma}{2\alpha} + 2\vartheta - \frac{\alpha}{\gamma}\vartheta^{2}] \tag{27}$$

---

[12] This result is valid only for $\lim_{1/2} w$

As for the public hospital, he will get the residual part of the market. For the first ailment he will break even as in the previous case. For the second ailments, he gets a higher proportion of patients that are severely ill and this means that the expected surplus will be negative and equal to $-\dfrac{\vartheta}{4}$.

5. Discussion

The model presented in this paper shows that there are many areas in contracting for health care where opportunistic behaviour might play a fundamental role. In this context we have examined just two of the possible areas where problems might arise. The first type of behaviour we have described is the choice of patients according to their ailment type, something that we have defined horizontal cream skimming. This type of opportunistic behaviour might be determined in the actual world by several reasons: some payment schemes might be more generous than others, it might be the case that the technology involved is not readily available (either because of rationing in human capital or in funding). In this paper we have considered another interesting cause: the variance in the cost to provide health care because we think it is particular relevant in hospital management. While the hospital is ready to accept contracts that provide a remuneration only in expected terms, if they cannot choose how much risk to take without affecting their rate of profit, they will naturally choose to specialise in those treatments that have a lower variance. This is the economic interpretation of the mean variance model presented here that we think might reflect the real world. The most obvious way to reduce this opportunistic behaviour is to take account of the variance (hence the risk) of the payment in the reimbursement scheme. In this way a trade off is created between the expected value of the surplus and its variance and hospitals might be less ready to specialise. Another and less costly alternative would be to impose the hospital a range in the case mix or a ceiling to the number of treatments of the same ailments that are to be reimbursed.

It is interesting to note that in this case it is the purchaser that does not foresee the full implications of hospital behaviour. In fact it observes the attitudes towards risk of the other agent but he is not able to predict that the agent will choose a strategy that takes account both of the surplus and of its variance.

The second opportunistic behaviour we have examined consists of choosing patients according to their expected cost. It arises from an asymmetry of information between the purchaser and the provider which means that the provider can observe the cost of the patient before treating him. It goes against the rules of the contract, but it is quite hard to

be detected since it is usually carried out using very subtle devices. In our model we have assumed that this selection is done through the quality mix: patient that are healthier are more interested in hotel services than those that need more health care and their preferences can be exploited by the provider of the service to choose. The extent of cream skimming depends on a range of parameters such as the distribution of low and high cost cases and the preference bias of the low cost group. Vertical cream skimming has important consequences:

- public hospitals will usually have a deficit since they will treat a higher proportion of patients with higher cost;
- some might not be treated;
- welfare is reduced;
- private hospital makes a surplus that is not related to a higher degree of efficiency;
- the cost to provide health care is higher than in first-best.

In some countries, such as Italy cream skimming has perverse effects on the whole cost minimisation incentive structure of the reformed health care market. Prospective payments were mainly introduced to force hospitals to reduce slacks in the production process. If the payment is tailored to break even, a persistent deficit of the hospital can be interpreted as a signal for inefficiency. However, in the presence of cream skimming, this is no longer the case. The public hospital have higher cost because of the patients it has to treat. While it might be reasonable in this case to allow extra funds to public hospitals, it has however to be recognised that the cost minimising properties of the prospective payment scheme are completely undermined.

This behaviour is quite difficult to be reduced or corrected for. Most of the formulas presented in the literature for asymmetry on cost revelation do not allow to reduce it. The traditional incentive compatible schemes uses effort levels to define a structure where the gain in cheating on the effort is optimal. Most of these models however foresee a positive surplus in the better state of the world and this means that selection is still incentivated.

Lewis and Sappington (1999) propose a double payment system to avoid the problem a prospective payment system for low cost patients and a partial cost reimbursement for high cost ones. This allows the hospital to choose the scheme that makes them better off and secures that both patients types are treated.

Chalkley and Khalil (2001) develop a framework to show that, with asymmetric information, the choice of the payment scheme may depend on the responsiveness of demand by consumers. The analysis indicates that there is an important role for the information that consumers have regarding the nature of medical interventions. They suggest payment based on treatment cost in order to contain costs in cases of emergency

care and life threatening illness and payment schemes conditioned on health outcomes for elective procedures and less severe illness.

Another way to reduce cream skimming might be to introduce very severe penalties if hospitals are found cheating. The process of controlling is however quite costly since it implies having a panel of doctors to evaluate all the single cases treated by the hospital. In any case, cream skimming should be curbed because of its undesirable characteristics.

6. Conclusions

In this paper we have examined some of the possible reasons that lead hospital to opportunistic behaviour in the choice of the patients to treat.

In particular we have concentrated on two practices that we have called horizontal and vertical cream skimming.

Horizontal cream skimming is a legal practice, although it goes against the principle of universal access to health care; it arises form a regulatory problem, i.e. the regulator has not set prices correctly and the hospital finds it convenient to specialise in some outputs.

Vertical cream skimming is instead an illegal behaviour that consists in offering health care only to the patients that have a low cost. It arises from the inability of the purchaser to observe the patient type and it might be solved through control and sanctions rather than incentives.

For horizontal cream skimming the behaviour of the market is somehow able to correct for the mistake made by purchaser in setting the price. The difference between the quality that is reimbursed and the one that is supplied to the patient is inflated and in this way the private hospital gets his premium for the risk he has to bear.

Vertical cream skimming has a more pervasive effect and it might endanger the entire market structure. The ability of the hospital to observe patients type enables the private organisation that operate on the market to choose low cost patients. This means that they will make a profit that does not depend from their ability to cost minimisation. On the other hand the public hospital, having to treat all the high cost will have a deficit that the purchaser or Central Government will have to eventually bear. This system has perverse incentive effects because both private and public hospitals might have an interest in the existence of vertical cream skimming: the private one since are allowed a surplus on top of the normal remuneration for their own effort and the public one because they might claim that cream skimming is the cause of their apparent inefficiency, even when it depends on slacks in the production process.

The model presented in this paper is the first step towards a comprehensive assessment of cream skimming. It needs to be developed in at least two direction: from a theoretical point of view, it is necessary to compare the properties of incentive systems such as the one described in Lewis and Sappington (1999) with control and penalties methods; from an empirical point of view it is necessary to determine the extent of the problem, something that so far almost no study has attempted to estimate.

Appendix 1

The problem faced by P can be written as:

$Min \quad t_1 + t_2$

$C_1 = q_1 + \overline{\beta_1} - e_1$

$C_{2i} = q_2 + \beta_{2i} - e_i \qquad\qquad i = L, H$

$U_1(t_1 - C_1 - f(e_1)) \geq 0$

$EU_2(t_2 - C_{2i} - f(e_{2i})) \geq 0$

The inequality for the reservation utility of the hospital can be taken as an equality. The first constraints can then be substituted in the second constraints. The problem can be solved using the Lagrange approach:

$Min \quad Tr = t_1 + t_2 - \lambda_1[U_1(t_1 - C_1 - f(e_1))] - \lambda_2[pU(t_2 - C_{2L} - f(e_{2L})) + (1-p)U_2(t_2 - C_{2H} - f(e_{2H}))]$

The F.O.C. for the problem can be written as:

$\dfrac{\partial TR}{\partial e_1} \qquad \lambda_1[f'(e_1) - 1] = 0$

$\dfrac{\partial TR}{\partial e_{2L}} \qquad \lambda p[f'(e_{2L}) - 1] = 0$

$\dfrac{\partial TR}{\partial e_{2H}} \qquad \lambda(1-p)[f'(e_{2H}) - 1] = 0$

$\dfrac{\partial TR}{\partial t_1} \qquad 1 - \lambda_1 U_1' = 0$

$\dfrac{\partial TR}{\partial t_2} \qquad 1 - \lambda_2(pU_{2L}' + (1-p)U_{2H}') = 1 - \lambda_2 E(U_2')$

giving:

$f'(e_1) = f'(e_{2L}) = f'(e_{2H}) = 1$

$pU(t_2 - C_{2L} - f(e_{2l})) + (1-p)U(t_2 - C_{2H} - f(e_H)) = 0$

where U' is the marginal utility of the reimbursement scheme in any state of the world. Among the contracts that are optimal when the hospital is risk neutral, let us choose the prospective payment where U=k=1

In this case t_2 can be obtained solving the following equation:

$[t_2 - [pC_l + (1-p)C_h] - [pf(e_l) + (1-p)f(e_h)] = 0$

$t_2 = E(C_i) + E[f(e_h)]$

Appendix two:

The utility function that the hospital wants to maximise can be written as:

$$\text{MAX} \quad (q_1^* - q_1)\left[\frac{\alpha}{2\gamma}(q_1 - q_1^*) + \frac{1}{2}\right] + \{[(q_2^* - \tilde{q}_2 - \hat{q}_2) + E(\beta) - \beta^l]p\{\frac{\alpha}{2\gamma}[w\hat{q}_2 + (1-w)(\tilde{q}_2 - q_2^*)] + \frac{1}{2}\} +$$

$$+ \{[(q_2^* - \tilde{q}_2 - \hat{q}_2) + E(\beta) - \beta^h](1-p)\{\frac{\alpha}{2\gamma}[(1-w)\hat{q}_2 + w(\tilde{q}_2 - q_2^*)] + \frac{1}{2}\}$$

s.t.

$$q_1; \hat{q}_2; \tilde{q}_2 \geq 0$$

and the optimal choice for is determined by the value of q*,p,w,β_L and β_H. Some insights can however be gained by studying special cases. Let us start by assuming as in the text that p=1/2.

The F.O.C. for the problem can be written as:

1) $\frac{\partial \Pi}{\partial q_1} = \frac{\alpha}{\gamma}q^* - \frac{\alpha}{\gamma}q_1 - \frac{1}{2} \leq 0; q_1 \geq 0; q_1(\frac{\partial \Pi}{\partial q_1}) = 0 \rightarrow q_1 = q^* - \frac{\gamma}{2\alpha}$

2) $\frac{\partial \Pi}{\partial \tilde{q}_2} = \frac{\alpha\beta^L w}{4\gamma} - \frac{\alpha\beta^H w}{4\gamma} + \frac{\alpha q^*}{2\gamma} - \frac{\alpha\tilde{q}_2}{2\gamma} - \frac{\alpha\hat{q}_2}{2\gamma} - \frac{\alpha\beta^L}{8\gamma} + \frac{\alpha\beta^H}{8\gamma} - \frac{1}{2} \leq 0; \tilde{q}_2 \geq 0; \tilde{q}_2(\frac{\partial \Pi}{\partial \tilde{q}_2}) = 0$

3) $\frac{\partial \Pi}{\partial \hat{q}_2} = -\frac{\alpha\beta^l w}{4\gamma} + \frac{\alpha\beta^h w}{4\gamma} + \frac{\alpha q^*}{2\gamma} - \frac{\alpha\tilde{q}_2}{2\gamma} - \frac{\alpha\hat{q}_2}{2\gamma} + \frac{\alpha\beta^L}{8\gamma} - \frac{\alpha\beta^H}{8\gamma} - \frac{1}{2} \leq 0; \hat{q}_2 \geq 0; \hat{q}_2(\frac{\partial \Pi}{\partial \hat{q}_2}) = 0$

from which we can derive the following expressions:

(2) $\tilde{q}_2 \geq q^* - \hat{q}_2 - \frac{\gamma}{\alpha} - (\beta^H - \beta^L)(\frac{2w-1}{4}); \hat{q}_2 \geq 0; \hat{q}_2(\frac{\partial \Pi}{\partial \hat{q}_2}) = 0$

(3) $\hat{q}_2 \geq q^* - \tilde{q}_2 - \frac{\gamma}{\alpha} + (\beta^H - \beta^L)(\frac{2w-1}{4}); \hat{q}_2 \geq 0; \hat{q}_2(\frac{\partial \Pi}{\partial \hat{q}_2}) = 0$

from which we can observe that health quality will be positive if the following constraint is satisfied:

$$q_2^* - \frac{\gamma}{\alpha} \geq (\beta^H - \beta^L)(\frac{2w-1}{4})$$

while for hotel quality it is sufficient that:

$$q_2^* - \frac{\gamma}{\alpha} \geq -(\beta^H - \beta^L)(\frac{2w-1}{4}).$$

If p≠1/2, the solution depends on the relative value of p and w. To start with, let us assume that patients care either for hotel services (the low cost one) or for health services (the high cost ones). In this case w=1.

The F.O.C can be written as:

(1) $\dfrac{\partial L}{\partial \tilde{q}} \; -\!-\!> \tilde{q} = (q^* p + \dfrac{E(\beta)p}{2} - \dfrac{\beta^H p}{2} - q^* + \dfrac{\hat{q}}{2} - \dfrac{E(\beta)}{2} + \dfrac{\beta^H}{2} + \dfrac{\gamma}{2\alpha}) \dfrac{1}{(p-1)}$

(2) $\dfrac{\partial L}{\partial \hat{q}} \; -\!-\!> \hat{q} = \dfrac{E(\beta)}{2} - \dfrac{\beta^L}{2} + \dfrac{q^*}{2p} - \dfrac{\tilde{q}}{2p} - \dfrac{\gamma}{2\alpha p}$

From (1) we can see that heath related quality ($\tilde{q}$) is non negative if:

$$q^* \geq -[E(\beta) - \beta^H] + \dfrac{\hat{q}}{2(1-p)} + \dfrac{\gamma}{2\alpha(1-p)}$$

i.e. q* has to be high enough to compensate for the loss in terms of cost deriving from a high cost patient and has to secure the monopoly rent to the provider. If p=1, q* should be infinitely high to have a positive $\tilde{q}$.

For p=0 we will have instead that: $q^* > \dfrac{\gamma}{2\alpha}$ [13] and this condition can be secured in the third stage of the maximisation process. In general, we can then say that:

for 0<p<1 e $q^* \leq \dfrac{\gamma}{2\alpha} - \dfrac{1}{2}[E(\beta) - \beta^H]$ the optimal choice for the hospital would be a negative quality. In this case the Khun Tucker constraints are binding and the optimal choice will be 0.

Hotel quality will then be equal to:

$$\hat{q} = \dfrac{1}{2p}[p[E(\beta) - \beta^L] + q^* - \dfrac{\gamma}{2\alpha}]$$

For the first type of ailment, the quality level q^* required to grant a non-negative utility is the same as in the previous case. For the second type of ailment, the optimal combination ($\tilde{q}; \hat{q}$) depends on the value of p and w and cannot be in general defined in general terms.

---

[13] For p=0, the following expression holds: $E(\beta) = \beta^H$

References

Biglaiser, G and C.A. Ma (2001) "Price and Quality Competition under Adverse Selection: Market Organization and Efficiency, mimeo

Chalkley, M. and Malcomson, J.M. (1998), Contractiong for health services with unmonitored quality, *The Economic Journal,* 108, pp.1093-1110

Chalkley, M. and Malcomson, J.M. (2000), Government Purchasing of Health Services, in Culyer, A.J. and Newhouse, J.P. (eds.), *Handbook of Health Economics*, North-Holland, 461-536.

Chalkley, M. and Khalil, F. (2001), Third party purchasing and incentives: the "outcomes movement" and contracts for health services, *Dundee Discussion Papers in Economics*, No.125, June 2001

Chalkley, M. and Malcomson, J.M. (2002), Cost sharing in health service provision: an empirical assessment of cost saving, *Journal of Health Economics,* 84, pp.219-249

Ellis R (1997) Creaming, skimping and dumping: provider competition on intensive and extensive margins, *Journal of Health Economics*, vol.17, 537-555

Ellis, R. and T. Mc Guire (1986) Provider behaviour under prospective reimbursement, *Journal of Health Economics*, 5(2), 129-51

Levaggi, R. (1998) Optimal procurement contracts under a binding budget constraint, Public Choice, vol. 101, n.1-2 pp.23-37

Levaggi, R. (2002a) Asimmetria informativa sul mercato delle prestazioni sanitarie: il problema del cream skimming, *Tendenze Nuove*, n.2 (marzo-aprile) pp.221-238

Levaggi, R. (2002b) Cream skimming in the provision of health care, *Rivista di Scienza delle Finanze e Diritto Finanziario*, vol. LXI, n.3(1) pp.405-423

Levaggi, R. (2003) Asymmetry of information in a mixed oligopoly market for health care, *mimeo*.

Lewis, T. R. and Sappington, D.E.M. (1999) Cream skimming, cost sharing, and subjective risk adjusting in the health industry*, Journal of Economics and Management*, vol.8(3), pp.351-382.

Montefiori, M. (2003) Patient behaviour in the market for hospital services, *working paper DISEFIN n.4/2003, Genova, Maggio 2003*.

Salmon, P (1987) Decentralisation as an incentive scheme, Oxford Review of Economic Policy, 2, 24-43

Salop, S.C. (1979) Monopolistic competition with outside goods, *Bell Journal of Economics*, 10, 141-156.