Società italiana di economia pubblica

XV Conferenza

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

Pavia, Università, 3 - 4 ottobre 2003

THE INSTITUTIONAL SETTING OF INTERNAL THE MARKET FOR HEALTH CARE

ROSELLA LEVAGGI

Dipartimento di Scienze Economiche, Università di Brescia

società italiana di economia pubblica

dipartimento di economia pubblica e territoriale – università di Pavia

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

XV Conferenza SIEP - Pavia, Università, 3 - 4 ottobre 2003

pubblicazione internet realizzata con contributo della

COMPAGNIA di San Paolo

società italiana di economia pubblica

dipartimento di economia pubblica e territoriale – università di Pavia

The institutional setting of the internal market for health care

Rosella Levaggi

Abstract

Internal health care market are only virtually competitive because the separation of functions might not be supported by a separation of interests as in the competitive market. The cost of the health care is determined by quality, by personal characteristics of the patient and by the effort of the medical staff, but information is asymmetric. In this article the cost minimising properties of alternative systems to reimburse hospital treatments will be discussed. The design of the scheme is made from the standpoint of a benevolent Health Authority that wants to provide services of a given quality at the least possible cost using two hospitals that have fixed locations. The paper shows that the organisation and financial setting depend on the relative importance of three main elements: the slacks in production caused by lack of incentives to cost minimisation, the information rent deriving from the separation of purchasing and providing functions and finally a position rent each hospital enjoys and that can be extracted only in some cases. The model shows that when the competitors on the market do not share the same objectives price discrimination should be used whenever the environment allows to do so. Furthermore, pure prospective payments systems induce cost inflation if the hospital can get full information on all the decision variables before defining his strategy.

J.E.L. I11, I18, D82

Keywords: Internal markets, Asymmetry of Information,

Introduction

In most countries, a substantial proportion of health care is financed by the public sector and since the first oil crisis[1], the objective to rationalise and control expenditure has become a priority for any effective government policy.

Health care systems have been widely reformed and a separation between purchasing and delivering the service has been enforced to mimic the structure of a competitive market.

The organisation of the reformed internal market can vary among countries and, in some cases, even among regions of the same country[2]. Purchasers in this market can either be agencies directly controlled by Central Government as in the U.K. or local decisions makers; their role might be confined to buying products from providers or they might be allowed to produce some themselves. As for hospitals, several models have been developed. They range from independent entities that pursue similar objectives to private hospital to bodies that are directly run by the purchasers. The contract that regulates the provision of health care varies according to the organisation of the market; their distinguishing features being the method chosen to finance the provision and the rules for competition. Financial aspects are defined by output-related (prospective such as DRG, stage contingent or mixed payment) or cost-reimbursement schemes while the rules for competition are more complicated, especially when private hospitals enter in the picture.

The internal markets for health care have mixed performances as cost containment and health gains and it is difficult to say if some of the forms used are more effective than others in reaching a good balance between cost containment and health outcomes. Enthoven (2002) points out that competition in this market is not likely to work very well and that "*not anything that sounds like "competition or markets or private sector will necessarily improve economic performance"*.

The cost of the health care services is determined by the quality of care, by the ability of the patient to take advantage of health care and by the effort of the medical staff. The relationship between quality and health gains is unobservable since it depends on personal characteristics of the patients that are often unpredictable. This makes it very difficult to define the structure of the market for health care and the incentives. However, as Enthoven points out "*Engineering economic incentives is crucial to a good long-term outcome. In health care, it is not possible to make them perfect, but they can be made roughly right, and that, in turn, can be improved upon.*"

---

[1] and more recently from the '90's

[2] After the reform in 1995, the internal market for health care in Italy can be organised according to three different models.

In this article the cost minimising properties of alternative systems to organise and reimburse hospital treatments in an internal market will be discussed. The design of the scheme is made from the standpoint of a benevolent regulator that appoints a purchaser (HA) to provide hospital care to his population at the least possible cost. To do so, HA has to employ two hospitals that have fixed locations. 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, i.e. they react to the payment system offered by HA by fixing quality to the level that allows them to maximise their objective function.

The paper shows that the organisation and financial setting depend on the relative importance of three main elements: the slacks in production caused by lack of incentives to cost minimisation, the information rent deriving from the separation of purchasing and providing functions and finally a position rent each hospital enjoys and that can be extracted only in some cases. The model shows that when the competitors on the market do not share the same objectives price discrimination should be used whenever the environment allows to do so. Furthermore, pure prospective payments systems induce cost inflation if the hospital can get full information on all the decision variables before defining his strategy.

The paper will be organised as follows: in section two the market and its possible organisation is presented; section 3 develops the model used to determine the cost minimising properties of the different contracts that are discussed in section four. Finally section 5 draws the conclusions of the analysis.

2. The internal market for health care

The internal market for health care derives from the separations of the functions of purchasing and providing services for the citizens-taxpayers within a public health care system. The former organisation of health care, usually vertically integrated, led to inefficiencies in delivering the services that made policymakers decide to introduce the structure of a competitive market. From this idea, several models have been developed to solve the problems that the creation of an internal market produces. The main questions relate to the nature of the bodies that the reform has created, their independence in decision making, the contract rules for the provision of services and the regulation of the competition with private suppliers. In this paper we will concentrate on hospitals by assuming that purchasers are perfect agents for the community they represent and in this light they want to allow access to any patient needing hospital care at the minimum total cost.

Hospitals have different degrees of independence from the public sector, but in any case they are not private firm. A true form of privatisation is not possible for several reasons. The hospital cannot go bankrupt and be closed nor the management can change its destination of use: in both cases in fact the first to suffer would be the patients that, in the short run, would not receive health care[3]. Profits should also be avoided or reinvested in health care since they derive from compulsory contribution form the public rather than willingness to pay like in the private market. Taxation causes deadweight losses and administrative costs; to minimise these side effects tax should be kept to the minimum. Furthermore, taxation is an instrument to share the cost for public services and to improve the distribution of income. Its use to increase the profit of some industry would be against the rules of good public administration.

Some systems have solved this problem by imposing several restrictions; usually hospitals should balance their accounts and if a profit exists, it has to be re-invested to accrue the equipment of the hospital[4].

The degree of autonomy of the hospital is an important element in determining its objectives, the incentives to cost minimisation and the quantity of information that can be extracted at no cost.

For the purpose of our analysis we have grouped the institutional arrangements for hospitals in three different categories that are presented in table 1.

Tab. 1: The taxonomy of hospitals competing on the market

	Independent (S)	Directly controlled (DC)	Directly managed (DM)
Objectives	Their own	Share with HA	Share with HA
Reimbursement	Output-related	Output-related	Costs or output-related
Costs	Not observed	Not observed	Can be observed
Slack	None	Some slack – x_{DC}	Maximum slack - x_{DM}

Independent hospitals (S) are separated entities from the purchasers; 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. A contract has to be made in order to get their services and they can refuse any clause that limits their independence. The incentive to cost minimisation are maximum since these hospitals are run as a private firm.

[3] See Dawson and Howart (1994)

[4] The profit made by privately-run hospitals that work for the public sectors is still an open problem in mixed health care systems; the solutions will not be dealt with in this paper since its discussion goes beyond the scope of the present paper.

Hospitals that are directly controlled (DC) are public firms with a high degree of independence as far as their organisation is concerned. The control from HA is concerned with the imposition to reach some common objectives, but they are not constrained to reveal any private information on their running costs. These restrictions create slacks in the productive process and inflate their cost by an x-inefficiency parameter x_{DC}. Finally, the hospitals that are directly managed have a limited degree of autonomy; they do not have the ability to withhold information from HA that can make them to pursue its own objectives. These constraints means more slacks in the productive process an cost inflation by an x-inefficiency parameter $x_{DM}>x_{DC}$.

3. The model

The model presented here draws on Levaggi (2003). It develops a four-stage game from the standpoint of a benevolent regulator (R) that wants to minimise the cost to provide health care. R delegates a collective purchaser (HA) to buy health care from two providers whose objectives and costs of production depend on the institutional arrangement chosen to organise the internal market; they are located at the extremes of a line whose distance has been normalised to one. The N patients lie within this line and are uniformly distributed. HA is a perfect agent for R that shares with the latter the objective to provide care at the minimum cost. 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 HA 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. Finally in the fourth stage R compares the costs deriving from different institutional arrangements and chooses the one that allows to minimise costs.

3.1 Stage one: the payment scheme

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:

$$C_{ij} = \beta_i + q^* + x_j - e_i \qquad i = l,h \qquad j = S,DC,DM$$

where β_i is a patient-related cost, q^* is the quality level, e_i is the effort of the medical staff[5] and x_j is an x-inefficiency parameter related to the organisation of the hospital with $x_S=0$ and $x_{DC}<x_{DM}$. β_i is a random variable that can take only two values, β_l for patients with low morbidity and β_h for patients with high morbidity[6]. Both events have a known probability equal to p and $(1-p)$ respectively. β is specific to each hospital, but there is a degree of correlation r among its realisation, so that the observation of a parameter in one hospital makes the realisation of the same level more probable in the other one. In particular, we can define the joint probabilities of the event β_l β_h as:

$$\pi_{h^j,h^i} = (1-p)(1-p(1-r))$$

$$\pi_{l^j,l^i} = p[(1-(1-p)(1-r)]$$

$$\pi_{l^j,h^i} = \pi_{h^j,l^i} = p(1-p)(1-r)$$

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 environment is characterised by asymmetry of information as shown in figure 1.

Fig.1: The timing of information

HA sets contract $\longrightarrow$ realisation of β $\longrightarrow$ hospital observes β $\longrightarrow$ effort by hospital $\longrightarrow$ outcome $\longrightarrow$

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.

Asymmetry of information of the type described in figure one means that the hospital management participates to the production process only if the reward received, net of the cost of production, produces a positive utility:

$$t_i - C_i - f(e_i) \geq 0$$

The choice of the reimbursement scheme depends on the institutional setting and the type of market where the hospitals are competing. The two basic methods are output or cost related schemes. The former has to be used in the presence of asymmetry of information due to the well known principle that contract can only be made on

---

[5] 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.

[6] It is assumed that morbidity is correlated with the recovery speed of the patient and hence with cost. If morbidity is low, the recovery rate is high and hence cost is low.

observable characteristics[7]. In this context, output-related schemes are less flexible than cost reimbursement because they do not allow any form of price discrimination. An output-related scheme, in a market where hospitals have a different level of independence compete, must foresee uniform pricing rules meaning that rents deriving from asymmetry of information and x inefficiencies have to be paid to both competitors. When the market setting allows to do so, a tailor made payment scheme reduces the total cost.

3.1.1 Direct management of both hospitals

If both hospitals are directly managed by HA the latter can observe ex post the realisation of β and the hospital cannot cheat on the effort. The problem can be written as[8]:

$$Min \quad pt_l + (1-p)t_h$$
$$s.t.$$
$$C_i = q_i + \beta_i + x_{DM} - e_i \qquad i = l,h$$
$$t_i - C_i - f(e_i) \geq 0$$

and the optimal solution can be written as:

$$f'(e_i) = 1$$
$$t_i^* = C_i^* + f(e_i^*) \quad = \quad q_i + \beta_i + x_{DM} - e_i^* + f(e_i^*) \quad = q^* + C_i^{\min} + x_{DM}$$

where $C_i^{\min}$ is the minimum cost to offer a service with quality equal to zero. The hospital receives its reservation utility in each state of the world since the information on the type of patient to be treated is known before making the effort. In this environment, in fact, the hospital will never accept a contract that allows to receive the reservation utility only in expected terms[9].

3.1.2 Independent hospitals

If hospitals are independent, costs cannot be observed and an incentive compatible mechanism has to be devised to make them reveal their private information. HA's objective in this new environment can be written as:

$$Min \quad pt_l + (1-p)t_h$$
$$s.t. \qquad C_{ij} = q^* + \beta_i + x_j - e_i \qquad i = l,h \qquad j = S,DC$$
$$t_i - C_i - f(e_i) \geq 0$$
$$t_i - C_i - f(e_i) \geq t_j - C_j - f(e_{ij}) \qquad i = l,h \qquad j = l,h \qquad j \neq j$$

---

[7] See Rees (1985)
[8] See Levaggi (2003) for a formal proof.
[9] See Levaggi (2003) and Sappington (1983)

where $f(e_{ij})$ represents the effort compatible with declaring C_j when the true state of the world is i . This new constraint, which is also called the incentive compatibility constraint, means that the hospital has the incentive to reveal truthfully the state of the world that has occurred and to do its effort accordingly. The solution is characterised by the following conditions whose derivation is presented in appendix 1:

$$f'(e_l) = 1 \qquad e_l = e_l^*$$

$$f'(e_h) = 1 - p[1 - f'(q^* + x_j + \beta_l - C_h)] < 1 \qquad e_h = e_h^{IC} < e_h^*$$

$$t_l^{IC} = C_l^* + f(e_l) + [f(e_h^{IC}) - f(e_{lh}^{IC})] \qquad t_l^{IC} = q^* + C_l^{\min} + x_j + \Delta^{IC}$$

$$t_h^{IC} = C_h^{IC} + f(e_h^{IC}) = \qquad t_h^{IC} = q^* + C_h^{\min} + x_j + \Delta C^{IC}$$

$$U_l = f(e_h^{IC}) - f(e_{lh}^{IC})$$

$$U_h = 0$$

In this scheme cost is inflated by an inefficient use of the effort in the worst period that is equal to $\Delta C^{IC} = \left[(e_h^* - e_h^{IC}) - ((f(e_h^*) - f(e_h^{IC}))\right]$ and by the information rent $\Delta^{IC} = [f(e_h^{IC}) - f(e_{lh}^{IC})]$ which corresponds to the utility above the reservation level that is received for each patient that is low morbidity.

Since both hospitals are independent and costs cannot be observed, only output-based, uniform pricing schemes can be used. This means that when hospitals with a different degree of independence and x-inefficiency compete, the payment will have to be tailored to the hospital with higher costs:

$$t_l^{IC} = q^* + C_l^{\min} + \max\{x_i; x_j\} + \Delta^{IC}$$
$$t_h^{IC} = q^* + C_h^{\min} + \max\{x_i; x_j\} + \Delta C^{IC} \qquad \text{i=S,DC}$$

The formula presented here is different from a pure output-DRG based system where hospitals are reimbursed on a pure prospective basis. Levaggi (2003) analyses the reasons why a prospective payment system, one of the most popular way to reduce the costs arising from asymmetry of information, might not be an optimal choice for this contract. Timing of information is crucial; in this context it is reasonable to assume that the type of patient is known before setting the effort and this means that the hospital will never accept to bear any risk in terms of his utility. A prospective payment in this context would have to allow the management to get his reservation utility in the worst state of the world; in the terms of our model it would mean $t = t_h$. Chalckley and Malcomson (2002) show that cost savings ranging from 7% to more than 60% could be achieved by using cost sharing contracts. Pure cost reimbursement contracts inflate costs (Weisbrod, 1991, Newhouse, 1992, Ma, 1994), but when they are corrected for asymmetry of information they perform better than prospective payment schemes.

3.1.3 One hospital is managed directly by HA

When HA manages one of the hospitals and observes the realisation of β, it can use this information to reduce the information rent of the hospital that is not directly controlled. The incentive compatible scheme can be made through a Bayesian update of the probability of β after this parameter has been observed for the hospital under direct management. The problem for HA can then be written as:

$$Min \quad z_{DM=i}t_l + (1 - z_{DM=i})t_h \qquad i = l,h$$

$$s.t.$$

$$C_i = q^* + x_j + \beta_i - e_i \qquad i = l,h \qquad \mathrm{j = S, DC}$$

$$t_i - C_i - f(e_i) \geq 0$$

$$t_i - C_i - f(e_i) \geq t_j - C_j - f(e_{ij})$$

where $z_{DM=i}$ is the conditional probability of a high/low recovery parameter for the other hospital given that in DM the observed parameter was high/low. The solution for the game is similar to the one presented in appendix 1 for the classical incentive compatible problem and can be written as:

$$f'(e_l) = 1 \qquad e_l = e_l^*$$

$$f'(e_h) = 1 - z_{A=i}[1 - f'(q^* + x_j + \beta_l - C_h)] < 1 \qquad e_h = e_h^{DM}$$

$$t_l^{DM} = C_l^* + f(e_l^*) + [f(e_h^{DM}) - f(e_{lh}^{DM})] \qquad t_l^{DM} = q^* + C_l^{\min} + x_j + \Delta^{DM}$$

$$t_h^{DM} = C_h^{DM} + f(e_h^{DM}) \qquad t_h^{DM} = q^* + C_h^{\min} + x_j + \Delta C^{DM}$$

$$U_l = f(e_h^{DM}) - f(e_{lh}^{DM})$$

$$U_h = 0$$

In this case, it is interesting to note the role played by z in determining the two incentive schemes. z in fact assumes the following two values:

$$z_{DM=l} = p + (1 - p)r$$

$$z_{DM=h} = p(1 - r)$$

where r is the correlation between β's in the two hospitals.

In the first case, having observed β_l for DM, this event is more probable than β_h and for this reason the effort in this occurrence is reduced, C_h^{DM*} increases, but the incentive given to the hospital for his information rent is reduced. In the second occurrence, β_h increases its probability of occurring and for this reason, the effort in this state is increased, hence increasing the efficiency of the game. The use of information on DM has the advantage to tailor the incentive to the other hospital, but do not avoid the problem of cheating. $\Delta C^{DM} = \left[(e_h^* - e_h^{DM}) - ((f(e_h^*) - f(e_h^{DM}))\right]$ is the cost deriving from the non-optimal use of the effort and $\Delta^{DM} = p[f(e_h^{DM}) - f(e_{lh}^{DM})]$ is the information rent of the hospital.

In this context, HA could use price discrimination (cost-per-case reimbursement for DM and an output-related payment for the other one) or uniform pricing (an output-related scheme for both hospitals).

The price discrimination scheme can be written as:

$$t_l^* = q^* + C_l^{\min} + x_{DM}$$

$$t_h^* = q^* + C_h^{\min} + x_{DM}$$

$$t_l^{DM} = q^* + C_l^{\min} + x_j + \Delta^{DM} \qquad \mathrm{j = S, DC}$$

$$t_h^{DM} = q^* + C_h^{\min} + x_j + \Delta C^{DM}$$

while uniform pricing would imply:

$$t_l = q^* + C_l^{\min} + \max\{x_{DM}; \Delta^{DM} + x_j\} \qquad \mathrm{j = S, DC}$$

$$t_h = q^* + C_h^{\min} + \max\{x_{DM}; \Delta C^{DM} + x_j\}$$

Uniform pricing produces cost inflation. For a fixed level of q*, the second formulation foresees the payment of the highest cost. At this stage we cannot however conclude that uniform pricing should be avoided as far as possible. In the second stage, when hospitals compete for quality, they use some of the extra resources received to increase their level of quality, hence q* to get the market fully covered using cost reimbursement is different from the one needed to achieve the same objective under uniform pricing.

3.2 Second stage: quality determination

In the second stage of the game, given the cost reimbursement scheme, the two hospitals compete for patients using the rules of spatial competition.

The population needing health care consists of N patients, uniformly distributed on a unit line while the two hospitals are located at the two extremes (0 and 1). The hospitals have the same size and technology, but might differ in their degree of independence from HA.

The service is free at the point of use while travelling costs have to be borne by the patient. Each patient is indexed by $d \in [0,1]$, so that *d* represents the patient located at point *d* from the origin. Patients observe quality and incur the same marginal distance cost *s*; they choose to go to the hospital that maximises the difference between quality and travelling cost. The utility function of a patient located at point *d* depends on the hospital he gets admitted to and it can be written as:

$$V_d = \begin{cases} \varphi q_0 - sd \\ \varphi q_1 - s(1-d) \end{cases} \qquad ()$$

where φq_i is the monetary evaluation of hospital services of quality q from the hospital located at i (0 or 1 in this model); *sd* and *s(1-d)* are travel costs. The patient will use hospital services only if $V_d \geq 0$

Patients have the same valuation of quality characteristics and incur the same marginal distance cost *d*; they will be indifferent between the two hospitals when $\varphi q_0 - sd = \varphi q_1 - s(1-d)$. This expression can be solved for the location of the marginal consumer:

$$d = \frac{\varphi(q_0 - q_1)}{2s} + \frac{1}{2}$$

The demand for hospital i is obtained by multiplying the distance by the density which, given the unit length of the line, is equal to N. The demand for each hospital can then be written as:

$$D_i = \left[\frac{\varphi(q_i - q_j)}{2s} + \frac{1}{2}\right] N$$

The hospitals that are directly managed or controlled will choose $q_i = q^*$; the one that are independent will instead choose the level of q_i that maximises their total utility:

$$Max \quad U^{tot} = \left[p(t_l - C_l^* - f(e_l^*)) + (1-p)(t_h - C_h^* - f(e_h^*))\right] D_i \qquad ()$$

where t depends on the reimbursement scheme adopted. Given that the rules for cost reimbursement have already been defined, we can rewrite () as:

$$Max \quad (q^* - q_i + (1-\delta^R) p\Delta^K + \delta^R \Delta E^j) * D_i$$

where

$$\Delta^K = \begin{cases} \Delta^{IC} \\ \Delta^{IC} \\ \Delta^{DM} \end{cases} \qquad \Delta E^j = \begin{cases} p\Delta^{IC} & \text{if S competes with S} \\ x_{DC} + p\Delta^{IC} & \text{if S competes with DC} \\ (1-p)\left[Max\{x_{DM}, \Delta C^{DM}\} - \Delta C^{DM}\right] + pMax\{x_{DM}, \Delta^{DM}\} & \text{if S competes with DM} \end{cases}$$

and $\delta^R=1$ for uniform and $\delta^R=0$ for tailor made payments.

The F.O.C. can be written as:

$$-D_i + (q^* - q_i + (1-\delta^R) p\Delta^K + \delta^R \Delta E^j) * \frac{\varphi N}{2s}$$

Solving for *q* we can write that:

$$q_i = \frac{1}{2}(q^* + q_j + (1-\delta^R) p\Delta^K + \delta^R \Delta E^j - \frac{s}{\varphi})$$

This expression shows that there is a trade off between the payment made in excess of cost, the quality offered by the hospital and the position rent. This becomes quite important in the evaluation of the minimum cost to deliver health care; the type of competition developed might allow HA to get back in the form of better quality some of rents it has to pay for asymmetry of information and local monopoly.

With three hospital structures, six are the possible market combinations for delivering health care, each of them implying a different total cost.

3.2.1 Both hospitals are independent (S/S)

With identical hospitals, it seems reasonable to assume that a symmetric Nash equilibrium will exists in which firms choose the same quality [10]. The quality offered by each competitor can then be written as:

$$q_i = q^* + p\Delta^{IC} - \frac{s}{\varphi} \qquad ()$$

where s/φ is the spatial monopoly rent enjoyed by the hospital. Equation () shows the trade off between quality and cheating; the information rent is in fact fully used to compete for quality with the other hospital in order to attract patients.

3.2.2 Hospitals share objectives with HA (PS/PS DM/DM PS/DM).

The hospital shares with HA the objective of delivering the same quality level HA reimburses. In this case, the quality set by both hospitals will be equal to $q_i = q^*$. Since both hospitals are equal they will have the same market share. The costs implied by these models are however different as it will be shown in the following section. In general the cost structure depends on the type of scheme chosen (cost reimbursement or output related) and on hospital type.

3.2.3 One of the hospital shares quality objectives with HA (PS/S DM/S)

In this case, the hospital that shares its quality objectives with HA will set its quality level to $q_i = q^*$. The reaction of the other hospital to this level of quality can be written as:

$$q_i = q^* + \frac{1}{2}((1-\delta^R)p\Delta^K + \delta^R \Delta E^j) - \frac{s}{\varphi})$$

Just half of the extra payment made to the hospital because of private information and greater efficiency is paid back in the form of extra quality; however, only half of the position rent the hospital enjoys can be used to decrease quality. This means that the introduction of the market of competitors that do not share the same objectives allows to reduce the monopoly power of those pursuing straight utility maximisation, but has costs in terms of incentives that have to be paid to increase the quality level.

---

[10] See Economides (1989), Gravelle(1999)

3.3 Stage three: the choice of the optimal quality level

In the third stage, HA sets q* to the level that makes it convenient for all the patients to receive hospital care. This condition requires that the marginal patient get his reservation utility when admitted. When both hospitals share the same objectives, they will share the market equally and the marginal patient is located at *d=1/2*. The market will be covered when:

$$\varphi q_i - \frac{s}{2} = 0 \qquad\qquad q_i = \frac{s}{2\varphi}$$

When hospital share their objectives with HA (DM and DC), the quality that allows to the marginal patient to hospital will be equal to :

$$q^* = \frac{s}{2\varphi}$$

If both hospitals are independent (S), their reaction function can be written as:

$$q_i = q^* + p\Delta^{IC} - \frac{s}{\varphi}$$

and the quality level that makes the marginal patient go to the hospital will be equal to:

$$q^* = \frac{3s}{2\varphi} - p\Delta^{IC}$$

It is interesting to note that q^* is not necessarily greater than the case where two hospitals on which HA has a degree of control compete. If $\frac{s}{\varphi} < p\Delta^{IC}$, i.e. if the position rent is lower than the information rent, the hospital will try to get patients by increasing its quality level above q^*.

When two hospitals with different objectives compete, they do not supply the same quality level hence they will not share the market equally.

In this case to have the market fully covered the marginal patient has to be indifferent between going to hospital or not receiving treatment, but its location is not ½. To find the quality level that clears the market and the shares of the two hospitals the problem can be written as:[11]

$$q^* = \frac{s}{\varphi} d_1 \qquad \text{(I)}$$

$$q^* + \frac{1}{2}\left[(1-\delta^R)p\Delta^K + \delta^R \Delta E^j) - \frac{s}{\varphi}\right] = \frac{s}{\varphi} d_2 \qquad \text{(II)}$$

$$d_1 + d_2 = 1 \qquad \text{(III)}$$

[11] The formal derivation is presented in appendix two

hence:

$$q^* = \frac{3s}{4\varphi} - \frac{1}{4}\left[(1-\delta^R)p\Delta^K + \delta^R \Delta E^j\right]$$

$$D_A = \left[\frac{3}{4} - \frac{\varphi}{4s}\left[(1-\delta^R)p\Delta^K + \delta^R \Delta E^j\right]\right]N \quad D_B = \left[\frac{1}{4} + \frac{\varphi}{4s}\left[(1-\delta^R)p\Delta^K + \delta^R \Delta E^j\right]\right]N$$

As in the previous case, the quality level necessary to have the market fully covered might be higher or lower than when HA share quality objectives with the hospitals. Both quality and market shares depend on the rents that the more independent organisation can enjoy; this element is in turn determined by *r*, i.e. the correlation between β's in the two hospital and by the chosen payment scheme.

3.4. The choice of the institutional setting

In the fourth stage, R compares the (minimum) costs arising from the different alternatives and chooses the price formula and the market structure that allows go minimise costs.

To start with, we will consider the problem of giving uniform, output based payment to both hospitals or, when this is possible, to use tailor made schemes.

This problem involves two market combinations where the cost for one of them can be observed with certainty, i.e. DM/S; DM/DC. In the first case, the cost for an output-related scheme will be equal to:

$$T = N\left[\frac{3s}{4\varphi} - \frac{1}{4}\Delta E^{DM} + E(C^{\min}) + \Delta E^{DM}\right]$$

while cost reimbursement would imply a total cost equal to:

$$T = D_A\left[\frac{3t}{4\varphi} - \frac{1}{4}p\Delta^{DM} + E(C^{\min}) + x_{DM}\right] + D_B\left[\frac{3t}{4\varphi} - \frac{1}{4}p\Delta^{DM} + E(C^{\min}) + \Delta AS^{DM}\right]$$

Price discrimination allows to reduce costs; the algebra for the demonstration is quite cumbersome; the intuitive reason for this result is that only half of the extra payment made to the more efficient hospital are recovered in the form of quality. The direct management of one of the hospital is necessary to reduce the costs deriving from asymmetry of information, but to achieve this result it is sufficient to pay the extra cost x_{DM} only to the patients treated by this hospital.

For the second type of market (DM/DC) we have:

$$T = N\left[\frac{s}{2\varphi} + E(C^{\min}) + (1-p)\left[Max\{x_{DM,}\Delta C^{DM} + x_{DC}\}\right] + pMax\{x_{DM,}\Delta^{DM} + x_{DC}\}\right] \qquad ()$$

$$T = \frac{N}{2}\left[\frac{s}{2\varphi} + E(C^{\min}) + x_{DM}\right] + \frac{N}{2}\left[\frac{s}{2\varphi} + E(C^{\min}) + x_{DC} + (1-p)\Delta C^{DM} + p\Delta^{DM}\right]$$

From () we can note that the output-related scheme is always more expensive. This form of market allows to reduce the extra costs caused by asymmetry of information, but increases the expenses arising from x-inefficiency. The best way to keep it to the minimum is to use a tailor-made payment to each hospital.

As concerns the institutional settings of the hospitals competing on the markets, the cost deriving from the different models are presented in table two.

The first model presented in table two is the optimal, first best solution where the two hospitals are directly managed, but do not have x-inefficiency cost. It represents the minimum cost to provide health care to the N patients and can be used as a benchmark to evaluate all the other contracts.

This first-best solution is followed by the six market structures presented in the previous section; the last column shows the cost in excess of the first-best.

The interpretation key for the choice of the market structure is the following: moving from DM/DM to S/S there are gains in efficiency, but there is a loss in information and common goals. Asymmetry of information inflates the cost to produce health care for any given quality level, if the hospital pursue his own objectives, the quality to be reimbursed is higher than the one delivered, the difference being represented by the monopoly rent each hospital enjoys.

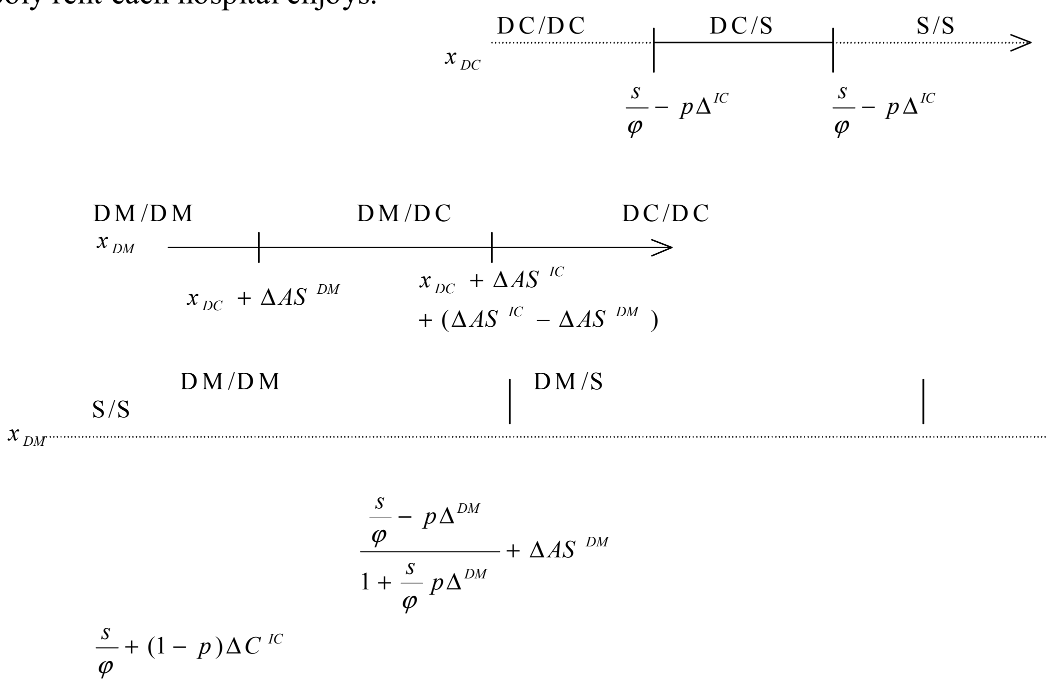

Figure two: Cost and market organisation

Table two: Comparison of costs

	Reimbursement scheme	Total cost	Difference
Bench	$t_l^* = \frac{s}{2\varphi} + C_l^{\min}$ $\quad t_h^* = \frac{s}{2\varphi} + C_h^{\min}$	$T = N\left[\frac{s}{2\varphi} + E(C^{\min *})\right]$	
DM/DM	$t_l^* = \frac{s}{2\varphi} + C_l^{\min} + x_{DM}$ $\quad t_h^* = \frac{s}{2\varphi} + C_h^{\min} + x_{DM}$	$T = N\left[\frac{s}{2\varphi} + E(C^{\min *}) + x_{DM}\right]$	Nx_{DM}
DC/DC	$t_l^{IC} = \frac{s}{2\varphi} + C_l^{\min} + x_{DC} + \Delta^{IC}$	$T = N\left[\frac{s}{2\varphi} + E(C^{\min}) + x_{PS} + p\Delta^{IC} + (1-p)\Delta^{IC})\right]$	$N(x_{DC} + p\Delta^{IC} + (1-p)\Delta C^{ic})$
	$t_h^{IC} = \frac{s}{2\varphi} + C_h^{\min} + x_{DC} + \Delta C^{IC}$		$N(x_{DC} + \Delta AS^{IC})$
DM/DC	$t_l^* = \frac{s}{2\varphi} + C_l^{\min} + x_{DM}$ $\quad t_h^* = \frac{s}{2\varphi} + C_h^{\min} + x_{DM}$	$T = \frac{N}{2}\left[\left[\frac{s}{2\varphi} + E(C^{\min *}) + x_{DM}\right]\right] +$	$\frac{N}{2}x_{DM} + \frac{N}{2}(x_{DC} + p\Delta^{DM} + (1-p)\Delta C^{DM})$
	$t_l^{DM} = \frac{s}{2\varphi} + C_l^{\min} + x_{DC} + \Delta^{DM}$	$+\frac{N}{2}\left[\frac{s}{2\varphi} + E(C^{\min *}) + x_{DC} + p\Delta^{DM} + (1-p)\Delta C^{DM}\right]$	$\frac{N}{2}x_{DM} + \frac{N}{2}(x_{DC} + \Delta AS^{DM})$
	$t_h^{DM} = \frac{s}{2\varphi} + C_h^{\min} + x_{DC} + \Delta C^{DM}$		
S/S	$t_l^{IC} = \frac{3s}{2\varphi} - p\Delta^{IC} + C_l^{\min} + \Delta^{IC}$	$T = N\left[\frac{3s}{2\varphi} - p\Delta^{IC} + E(C^{\min}) + p\Delta^{IC} + (1-p)\Delta C^{IC}\right]$	$N(\frac{s}{\varphi} + (1-p)\Delta C^{IC})$
	$t_h^{IC} = \frac{3s}{2\varphi} - p\Delta^{IC} + C_h^{\min} + \Delta C^{IC}$		
DC/S	$t_l^{IC} = \frac{3s}{4\varphi} - \frac{1}{4}\left[p\Delta^{IC} + x_{DC}\right] + C_l^{\min} + x_{DC} + \Delta^{IC}$	$T = N\left[\frac{3s}{4\varphi} + E(C^{\min}) + \frac{3}{4}(p\Delta^{IC} + x_{DC}) + (1-p)\Delta C^{I}\right.$	$N\left[\frac{1}{4}\frac{s}{\varphi} + (1-p)\Delta C^{IC} + \frac{3}{4}(p\Delta^{IC} + x_{DC})\right]$
	$t_h^{IC} = \frac{3s}{4\varphi} - \frac{1}{4}\left[p\Delta^{IC} + x_{DC}\right] + C_h^{\min} + x_{DC} + \Delta C^{IC}$		$N\left[\frac{1}{4}(\frac{s}{\varphi} - p\Delta^{IC}) + \Delta AS^{IC} + \frac{3}{4}x_{DC}\right]$

DM/S	$t_l^* = \frac{3t}{4\varphi} - \frac{1}{4} p\Delta^{DM} + C_l^{\min} + x_{DM}$ $t_h^* = \frac{3t}{4\varphi} - \frac{1}{4} p\Delta^{DM} + C_h^{\min} + x_{DM}$ $t_l^{DM} = \frac{3t}{4\varphi} - \frac{1}{4} p\Delta^{DM} + C_l^{\min} + \Delta^{DM}$ $t_h^{DM} = \frac{3t}{4\varphi} - \frac{1}{4} p\Delta^{DM} + C_h^{\min} + \Delta C^{DM}$	$T = D_A \left[\frac{3t}{4\varphi} - \frac{1}{4} p\Delta^{DM} + E(C^{\min}) + x_{DM} \right] +$ $+ D_B \left[\frac{3t}{4\varphi} - \frac{1}{4} p\Delta^{DM} + E(C^{\min}) + p\Delta^{DM} + (1-p)\Delta C \right.$	$D_A \left[\frac{1}{4} (\frac{s}{\varphi} - p\Delta^{DM}) + x_{DM} \right] +$ $D_B \left[\frac{1}{4} \frac{s}{\varphi} + \Delta AS^{DM} - \frac{1}{4} p\Delta^{DM} \right]$ $\frac{1}{4} (\frac{s}{\varphi} - p\Delta^{DM}) + D_A [x_{DM}] + D_B [\Delta AS^{DM}]$

The comparison of the costs arising from the different models is complicated given the nature of the extra costs that each structure implies; for this reason figure 2 offers just a partial ordering of the choices . In general, if the x-inefficiency cost is relatively low, direct management is the best alternative. As cost starts increasing, it might be convenient to make just one of the two hospitals more independent in order to reduce the x-efficiency loss from one side and the extra cost caused by asymmetry of information and local monopoly from the other. The choice on how independent the other structure should be depends on the relative importance of the costs involved in the different solutions.

Starting from a market where both hospitals are directly run, we can not that a complete switch between a directly controlled structure (DM/DM) and an independent market (S/S) depends on the relative importance of three key elements that characterises the model described in this paper, i.e. x inefficiency, information and position rent.

From figure two, we can note that the S/S structure should be preferred if $x_{DM} > \frac{s}{\varphi} + (1-p)\Delta C^{IC}$. In this case, in fact, the hospital increases its cost by his rent position and the cost inflation element deriving from asymmetry of information while the information rent ($p\Delta^{IC}$) is paid back through the Hotelling game.

The scope for hospitals that are directly managed appears to be limited; they are a viable solution only if:

$$x_{DM\leq} Min\left\{ \frac{\frac{s}{\varphi} - p\Delta^{DM}}{1+\frac{s}{\varphi}p\Delta^{DM}} ; \frac{s}{\varphi} - p\Delta^{IC} \right\}$$

In all the other cases a structure involving directly managed and/or independent hospital allows to cover the market a lower cost.

Finally, it is interesting to note that the combination DC/S is always dominated by the two pure form DC/DC or S/S.

This result is quite interesting since some of the reforms that have been proposed for the internal market for health care[12] have chosen a market where DC and S hospitals compete. The theoretical analysis presented here shows instead that only DM and S hospitals should be made compete on the internal market. Using DC hospitals, in fact, can reduce the position rent of the independent hospital, but at the cost of an increase in

[12] Italy for example

inefficiency. Since both hospitals have private information on the recovery parameter, HA has to pay the full information rent. This result is quite interesting from a policy point of view. It shows that the actual implementation of the internal market for health care might be flawed from the beginning and that in some cases the institutional arrangements for hospitals need a careful second thought.

5. Conclusions

The internal market for health care that several reforms in western countries have introduced have mixed performances as cost containment and health gains and it is difficult to define an institutional arrangement that allows to keep a good balance between cost minimisation and health outcomes.

In this article the cost minimising properties of alternative systems to reimburse hospital treatments have been discussed from the standpoint of a benevolent regulator (R) that wants to provide hospital care his population at the least possible cost through a purchaser (HA) and two providers with fix location. To do so, R can use several arrangements as concerns the organisation of the hospitals and their finance. The different institutional setting present a trade off between common goals and incentive to minimise cost: the more the control is stringent, the less incentive to an efficient use of the resources.

The paper shows that the choice of the organisation and financial setting depends on the relative importance of three main elements: the extra cost caused by lack of incentives to efficiency, the information rent deriving from the separation of purchasing and providing functions and finally a position rent each hospital enjoys and that can be extracted only in some cases.

Prospective payments should be avoided if the hospital can get full information on all the decision variables before defining his strategy.

The model shows that when the competitors on the market do not share the same objectives uniform payment schemes should be avoided as far as possible.

The introduction on the market of a direct controlled hospital, i.e. of an organisation that is financially independent, but share with HA quality goals might be optimal only when the x inefficiency cost is comparatively smaller than the position rent, in all the other cases hospitals that are directly managed or independent have better performances.

Directly controlled hospitals are used as a prevalent form to organise public sector health provision and the results of this paper cast some doubts on this decision. Furthermore, these structure should never be made compete with independent hospitals since the cost, mostly due to the use of uniform pricing is higher than for other institutional arrangements.

The paper presented here represents just a first step in the analysis of health care market; its structure is very simple and could be enriched by taking account of more complicated functional form for hospital costs and more diverging objectives for hospitals.

Appendix 1

$$Min \quad pt_l + (1-p)t_h$$
$$s.t.$$
$$C_i = q^* + \beta_i - e_i \qquad\qquad i = l,h$$
$$t_i - C_i - f(e_i) \geq 0$$
$$t_i - C_i - f(e_i) \geq t_j - C_j - f(e_{ij})$$

The problem has to be solved in terms of observable variables such as the cost and not in terms of effort *e* that only the hospital can observe. From the first constraint we can derive that $e_i = q^* + \beta_i - C_i$. The third constraint in the problem, the so-called Incentive Compatible Constraint can be written as:

$$t_l - C_l - f(q^* + \beta_l - C_l) \geq t_h - C_h - f(q^* + \beta_l - C_h)$$
$$t_h - C_k - f(q^* + \beta_h - C_h) \geq t_l - C_l - f(q^* + \beta_h - C_l)$$

The second order conditions on the disutility of the effort allows us to conclude that the second inequality is always satisfied. Let us now observe the first inequality. It states that the net payment to the hospital in the best states of the world has to be at least equal to the payment received in the worst state of the world plus a compensation for the disutility of the effort. Let us now observe the participation constraint. In the worst state of the world, the hospital receives a compensation that is equal to $t_h = C_h + f(q^* + \beta_h - C_h)$. Let us now observe the l.h.s. of the equation. We can observe that $f(q^* + \beta_l - C_h) < f(q^* + \beta_h - C_h)$, hence the utility received in the best state of the world is greater than zero, which in turns means that the first participation constraint is always satisfied. With all this in mind we can write the minimisation problem as:

$$Min \quad TR = p\{C_l + f(q^* + \beta_l - C_l) + [f(q^* + \beta_h - C_h) - f(q^* + \beta_l - C_h)]\}$$
$$+ (1-p)[C_h + f(q^* + \beta_h - C_h)]$$

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

$$\frac{\partial TR}{\partial C_l} = p[1 - f'(q^* + \beta_l - C_l)]$$
$$\frac{\partial TR}{\partial C_h} = p[f'(q^* + \beta_l - C_h) - f'(q^* + \beta_h - C_h)] + (1-p)[1 - f'(q^* + \beta_h - C_h)]$$

giving:

$$f'(e_l) = 1 \qquad\qquad e_l = e_l^*$$
$$f'(e_h) = 1 - p[1 - f'(q^* + x_j + \beta_l - C_h)] < 1 \qquad\qquad e_h = e_h^{IC} < e_h^*$$
$$t_l^{IC} = C_l^* + f(e_l) + [f(e_h^{IC}) - f(e_{lh}^{IC})] \qquad\qquad t_l^{IC} = q^* + C_l^{\min} + x_j + \Delta^{IC}$$
$$t_h^{IC} = C_h^{IC} + f(e_h^{IC}) = \qquad\qquad t_h^{IC} = q^* + C_h^{\min} + x_j + \Delta C^{IC}$$
$$U_l = f(e_h^{IC}) - f(e_{lh}^{IC})$$
$$U_h = 0$$

Notation

A = Hospital A

B = Hospital B

s = transport cost

d= distance

C_i = cost to provide health care to patient of type i

β_i = factor affecting the ability of patients to take advantage of health care

e_i = state contingent effort of the management

q = quality of health care offered

U = utility of the management of the hospital

t_i= state contingent reimbursement scheme

p = probability that the patient is a high recovery (hence low cost)

r = correlation coefficient between β_i^A and β_i^b

$\Delta C^{IC} = \left[(e_h^* - e_h^{IC}) - ((f(e_h^*) - f(e_h^{IC}))\right]$ cost by asymmetry of information

$\Delta^{IC} = p[f(e_h^{IC}) - f(e_{lh}^{IC})]$ information rent

$\Delta C^{DM} = \left[(e_h^* - e_h^{DM}) - ((f(e_h^*) - f(e_h^{DM}))\right]$ cost by asymmetry of information

$\Delta^{DM} = p[f(e_h^{DM}) - f(e_{lh}^{DM})]$ information rent

ΔE^j = rent S can enjoy when competing with DC or DM

$\Delta A^{IC} = p\Delta^{IC} + (1-p)\Delta C^{IC}$ expected cost from asymmetry of information

$\Delta A^{DM} = p\Delta^{DM} + (1-p)\Delta C^{DM}$ expected cost from asymmetry of information

C_{min} = minimum cost to provide care of quality equal to zero

$E(C_{min})$ =minimum expected cost

z= conditional probability

π= joint probability

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

Culyer, A.J., Maynard A.K., Posnett J.W. (eds) (1990) *Competing in health care*, Mc Millan Press, Basingstoke

Dawson, D. (1994) Costs and prices in the internal market: markets vs the Management Executive guidelines, *Working Paper n. 115*, Centre for Health Economics, York

Dawson, D. e D. Howart (1994) Competition and the law in the internal market, *mimeo*

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

Enthoven A. (2002) Introducing market forces into health care: a tale of two countries, Paper presented at the 4th European Conference on Health Economics, Paris, 7-10 july. http://perso.wanadoo.fr/ces/Pages/english/PLS3.pdf

Gravelle, H. (1999) Capitation contracts: access and quality, *Journal of Health Economics*, 18, 315-340

Laffont, J.J. and J. Tirole (1993) *A Theory of Incentives in Procurement and Regulation, The MIT Press*, Cambridge, MA

Levaggi, R. (1999) Optimal procurement contracts under a binding budget constraint, *Public Choice*, 101, n.1-2, 23-37

Levaggi, R. (1996) NHS contracts: an agency approach, *Health Economics*, 5, 341-352

Levaggi, R. - L. Rochaix-Ranson (2001) Alternative systems for organising primary care and remunerating GP's: the role of patients' behaviour, *mimeo*

Levaggi, R. (2003) Cost and quality control in the market for hospital care, *mimeo*

Ma, C. A. (1994) Health care payment systems: costs and quality incentives, Journal of Economics and Management Strategy, vol.3 (1) pp.93-112

Newhouse, J. (1992) Medical care costs: how much welfare loss?, Journal of Economic Perspectives,6(3)3-21

Newhouse, J. (1996) Reimbursing health plans and health providers: efficiency in production vs. selection, *Journal of Economic Literature*, 34, 1236-63

Rees, R. (1985) The Theory of Principal and Agent: Part 1, Bulletin of Economic Research v. 37(1), iss. 1, pp. 3-26

Rochaix, L. (1997) Asymétries d'information et incertitude en santé: les apports de la théorie des contracts, *Economie et Prevision*, n.129-130, pp.11-24

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

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

Sappington, D. (1983) "Limited liability contracts between Principal and Agent", *Journal of Economic Theory*, 29:1-21.

Weisbrod, B.A. (1991) The health care quadrilemma: an essay on technological change, insurance, quality of care and cost containment, *Journal of Economic literature* 29(2)523-552