Risk sharing and moral hazard under prospective payment to hospitals: how to reimburse services for outlier patients

François Marechal* and Michel Mougeot†

March 2004

Abstract

We analyze the regulation of a single health care provider (*e.g.* a hospital). According to several payment rules used in different countries, we consider a mixed linear payment in which the hospital is paid a fixed price per DRG (diagnosis related group) for most patients (inlier patients) and is reimbursed by a cost sharing payment for patients with exceptionally costly stays (outlier patients). Given this form of payment, we determine the optimal threshold above which to consider a patient as an outlier patient, as well as the optimal payment per DRG and the optimal cost sharing parameter. For the case where the regulator can use a two part tariff, we also determine the fixed charge the regulator has to impose in order to extract hospital rents.

JEL classification: I18.

keywords : Hospitals payment, cost, quality, incentives, DRG, risk sharing.

*DEEP-HEC, IEMS (University of Lausanne, Switzerland) and CRESE. francois.marechal@hec.unil.ch.
We are grateful to Marius Brulhart and to participants at seminars in Oslo and Lausanne for helpful comments. We acknowledge the financial support of the National Swiss Fundation ($n°$ 101412 − 101878). The usual disclaimer applies.

†CRESE, University of Besançon (France) and IEMS. michel.mougeot@univ-fcomte.fr.

1 Introduction

The trade-off between efficiency and selection is a major concern for the definition of contracts between health service providers and government agencies or insurance companies. As noted by Newhouse (1996), efficiency in production means least-cost treatment of a patient's medical problem holding quality constant, and selection means actions by suppliers to exploit risk heterogeneity with the result that some consumers may not obtain the health care services they desire. Hence, the optimal payment system should induce an efficient mix of quality enhancement and cost reduction effort without giving hospitals incentives to shun unprofitable patients. This is a complex problem because of the non-observability of the hospital's effort and of the patients' severity. Therefore, neither the effort nor the quality of services can be described in an enforceable contract.

Until the 1980s, the main form of payment in the U.S. and in most developing countries was cost reimbursement. Insurance companies or government agencies reimbursed realized hospital costs. This contract corresponded to Medicare payments prior to 1983: hospital services were reimbursed through a retrospective cost-based system. Since the early 1980s, a second form of payment has appeared in the form of a fixed price. In 1983, Medicare introduced the Prospective Payment System (PPS). Costs per case are used in the United Kingdom since the recent reform. Other systems of price per case are used in different European countries. Prospective payment sets prices prior to the period for which care is given. Under PPS, rates are determined by diagnosis related groups (DRGs). Each DRG is given a flat payment calculated in part on the basis of costs incurred for that DRG nationally.

It is well-known that the cost-reimbursement system encourages inefficient production and inflation of input prices. However, retrospective payment allows hospitals to recover their expenses. Thus, hospitals have no incentive for dumping or cream skimming. Furthermore, if quality enhancement is costly, the provider may offer the optimal level of health service quality. PPS has opposite properties. Whereas a cost-based contract offers no incentive for cost reduction, a fixed-price contract induces the first-best amount of cost reduction effort, because it makes the hospital residual claimant for cost savings. In other respects, cost reduction incentives and quality enhancement incentives work in opposite directions. Thus, the hospital may try to lower costs by skimping on quality and it has an incentive to refuse unprofitable patients.

In the last decade, a number of papers have discussed the incentive properties and drawbacks of prospective payments. Shleifer (1985) shows that DRG reimbursement by Medicare is very close to yardstick competition and induces efficient cost decisions by hospitals. However, Medicare does not adjust payments for the severity of illness because the moral hazard associated with reporting severity is great. Other papers relax Shleifer's assumptions and show that reimbursement should not necessarily be fully prospective. Dranove (1987) shows that rate setting by DRG encourages hospitals to specialize in those DRGs for which they have relatively low production costs. Pope (1989) shows that Shleifer's solution underpays hospitals with a high quality-related cost component. Ellis and Mc Guire (1986) introduce the physician as a utility maximizing agent for both the patient and the hospital. They demonstrate that a combination of fixed price and partial cost reimbursement is optimal when the doctor is not a perfect agent for the patient. Allen and Gertler (1991) show that some patients may receive inefficient quality under PPS when patients are heterogeneous within a given DRG.

More recently, some papers using the multitask agency approach (Holmstrom and Milgrom (1991)) have shown that prospective payment can achieve both efficient quality and efficient cost reduction effort. Ma (1994) considers that a hospital offering higher quality will attract more patients and shows that PPS can implement the efficient allocation of cost reduction and quality effort when the hospital cannot dump high-cost patients and when the DRG classification truly reflects cost variations. Chalkley and Malcomson (1998a) obtain a similar result under more general conditions on cost functions and patients' demand.

However, in practice the DRG classification fails to distinguish systematic cost variations. Therefore, hospitals have an incentive to shun unprofitable patients. Nevertheless, as noted by Ma (1994), based on the observation by Dranove (1984), "*it has been argued that through* *observing patient characteristics, a hospital can predict costs for 10-20% of its patients*". So, if a hospital decides to accept a patient, it has to bear some risks. The patient will only be treated if the price per discharge covers this risk, i.e. if the hospital receives a risk premium. Thus, we believe that the assumption of hospital managers exhibiting risk aversion is a useful assumption to reflect the decision to accept or refuse a patient.

When distributional concerns and hospitals' risk aversion are taken into account, the fully prospective price policy turns out not to be the optimal policy. If the regulator uses a fully prospective payment system, the hospital has a strong incentive to exert effort to lower costs. However, as the hospital bears all the risk, it will accept the contract only if the regulator offers it an appropriate level of profit. If a cost reimbursement system is used, the regulator bears all of the risks and the hospital none. Therefore, the hospital's behavior would be the same as in the risk neutral case (i.e. a zero cost reduction effort level and a zero risk premium). It is well-known that these two corner solutions are not optimal. The trade-off between risk aversion and moral hazard involves some cost sharing on the supply side as well as on the demand side. The need for cost sharing comes from the fact that costs of patients in the same DRG can differ widely. As a cost sharing contract reduces the risk borne by the hospital, it decreases the variance of the profit and the rent the regulator has to leave to the hospital.[1]

In this paper, we consider the hospital's payment problem as a regulation problem. There are three actors in the system: a hospital, patients and a regulator. The regulatory agency purchases the services from the hospital but is not the direct consumer. As in Laffont and Tirole (1993) we take the *"social cost of public funds"* approach and we assume that the regulator has a utilitarian objective function. We assume that the hospital incurs a disutility that increases with cost-reduction and quality-enhancement efforts. Thus, to accept a patient,

[1] Note that Chalkley and Malcomson (2002) investigate the need for cost sharing in a different context, i.e. with adverse selection and moral hazard when both the regulator and the hospital are risk neutral. Their optimal cost sharing rate solves the trade-off between rent extraction and incentives to reduce costs. In this paper, we do not consider the problem of adverse selection. Cost sharing is justified because the hospital is assumed to exhibit risk aversion whereas the regulator is risk neutral.

the hospital must be compensated by a monetary transfer in addition to the reimbursement of cost. Finally, we consider that patients' demand for a hospital's services is an increasing function of its quality of care. Thus, we assume that the hospital's choices are demand constrained. Our model is analogous to Ma (1994) with respect to demand functions and is a particular case of Ma with respect to the cost function. Departing from Ma, we assume that transfers from patients/tax payers to the regulator have to be multiplied by $(1+\lambda)$, where λ is the shadow cost of public funds. Thus, the deadweight loss from financing hospitals rents has to be considered. Moreover, we assume that the hospital is risk averse when patient severity is distributed randomly. As observed by Loeb and Magat (1979), in the absence of distributional concerns, moral hazard has no deleterious welfare effect: it suffices to give the hospital a reward equal to the net consumer surplus to obtain the first-best level of welfare. Taking the shadow cost of public funds into account implies that the regulator faces a three-way trade-off between (i) allocative efficiency (i.e. optimal choice of quality level), (ii) productive efficiency (i.e. optimal choice of cost reduction effort level), and (iii) minimizing the adverse distributional effect of the excess profits of the hospital[2].

When the hospital is risk averse, we have to add another trade-off between risk sharing and moral hazard and determine the optimal amount of cost sharing which solves this trade-off. In practice, as noted by Newhouse (2002), "*for patients with exceptionally costly stays, an additional payment is made equal to 80 percent of the accounting costs above some threshold or deductible. The threshold is set so as to be a given dollar amount above each DRG's mean payment rate. By law 5 percent of total payments are reserved for outlier payments; the outlier threshold is then set such that outlier payments will be 5 percent of the total.*" In many countries that have adopted prospective payment systems, hospitals are compensated for outlier patients through a two-part payment schedule. Below a given threshold, the hospital receives a flat payment. Above this threshold, it receives a given percentage of realized costs. Is this threshold optimally determined? What is the optimal cost-sharing rate? Does the value of the threshold depend on the value of the cost-sharing rate? This

[2] See Armstrong, Cowan and Vickers (1995).

paper proposes to investigate these questions. We consider a form of payment which is very close to what is used in practice (e.g. in the US or in Switzerland), i.e. a flat payment for inlier patients, and a cost sharing payment for outlier patients. Given this payment, we characterize the optimal contract, i.e. the optimal fixed price per patient, the optimal cost sharing parameter and the optimal threshold above which to consider a patient as an outlier patient. Since we consider that the regulator can use a two part tariff, we also determine the fixed charge the regulator has to impose in order to extract hospital rents.

The paper is organized as follows. The model is presented in section 2. Section 3 offers some numerical simulations of the optimal payment. Some conclusions are drawn in section 4.

2 The model

We consider the regulation of a single health care provider treating patients with a specific diagnosis. The regulator can be a public agency or a private insurance company contracting with a provider hospital. We assume that consumers are fully insured. Thus, demand is price inelastic and the patient's choice of a hospital depends solely on the quality of care it offers. According to the usual classification, we are in the case of search goods whose quality is observable by consumers but is not verifiable by the agency. Let us denote $x(q)$ the number of patients seeking treatment at the hospital when it provides care services of quality q. We assume that demand $x(q)$ is an increasing and concave function. We consider that quality has only one dimension and can be denoted by $q \geq 0$ such that $q = 0$ is the lowest quality level.

We assume that hospital costs are endogenous. By expending more effort, hospitals can reduce cost and increase quality. We assume that the unit cost of treatment C is equal to $c + q - e + \xi$, where c is a common knowledge efficiency parameter (i.e. there is no adverse selection), q and e denote respectively the quality enhancement and the cost reduction efforts that the hospital's manager can exert, and ξ is a random variable which represents patient severity (i.e. unpredictable treatment cost fluctuation observed only by the hospital)[3]. ξ is assumed to be observed by the hospital only after it has decided to treat the patient and after it has chosen its effort levels e and q. Although the regulator cannot observe the realization of ξ, it knows that ξ is drawn from a distribution F with corresponding density f. Under these assumptions, a risk neutral provider has an expected cost function

$$E(C) = (c + q - e + E(\xi))\, x(q), \tag{1}$$

that is assumed to be convex.

Like Ma (1994), we use the same variable q to denote the quality of care and the quality enhancement effort[4]. If the provider exerts effort level e, it decreases the monetary unit cost by e. We suppose that $e = 0$ is the lowest level of cost reduction effort. If the provider exerts quality enhancement effort q, it increases the monetary unit cost by q. Like Ma (1994), we assume that these two efforts impose a disutility (in monetary units) of $\varphi(q+e)$ to the hospital's manager.[5] Without loss of generality, we consider that $\varphi(q+e)$ is a quadratic function

$$\varphi(q+e) = \frac{(e+q)^2}{2d},$$

where d is a positive constant[6].

The hospital's profit π, when receiving a payment T, is

$$\pi = T - (c + q - e + \xi)\, x(q) - \varphi(q+e). \tag{2}$$

[3]Our cost function is close to Ma (1994). Indeed, Ma first considers a general cost function $c(q,e)$, with $\frac{\partial c(q,e)}{\partial q} \geq 0$ and $\frac{\partial c(q,e)}{\partial e} < 0$. Then, he considers a random unit cost of treatment which is independent of the quality of care. In this paper, we assume simultaneously that the unit cost is random and that it depends on both quality and cost reduction efforts. To do so, we consider a special case of $c(q,e)$.

Note also that the specified cost function $C = c + q - e + \theta$ implies an assumption of strict first order stochastic dominance, which means that high levels of effort systematically reduce the likelihood of high cost realizations. Ma (1994) does not need this assumption, since he only assumes that the mean cost $\overline{c}(e)$ is decreasing in e, which is more general (on this discussion, see Sharma (1998) and Ma (1998)).

[4]If we had denoted $\widetilde{q}$ the base quality and $\widetilde{q} + q$ the total quality of care, we could redefine $\widetilde{q} + q$ as a new variable and eliminate $\widetilde{q}$.

[5]A more general formulation of hospital's disutility can be found in Chalkley and Malcomson (1998).

[6]φ is then increasing and convex in both e and q.

We assume that the regulator is risk neutral and seeks to maximize the sum of its perceived net benefit of treatment and the hospital's expected utility $U(\pi)$. Let λ be the shadow cost of public funds when the regulator is a public agency (like NHS or Medicare) using distortionary taxation to finance health care insurance. When the regulatory agency contracts with a hospital, it is concerned with the number of patients receiving the treatment and with the quality and cost of the treatment. If $V(x, q)$ is the benefit perceived by the regulator of treating x patients with care quality q, the net benefit can be written $V(x, q) - (1 + \lambda) T$. Since x is an increasing function of q, we simply consider $V(q) - (1 + \lambda) T$. We assume that $V(q)$ is increasing and concave. When the regulator is a private insurance company, $V(q)$ represents consumer surplus. When the regulator is a public agency, however, $V(q)$ can differ from the consumer surplus because of moral hazard or of a specific public policy. We consider that the ex post unit cost of treatment $C = c + q - e + \xi$ is ex post observable. However, the regulator can neither observe the hospital's choice of q and e nor the random variable ξ. Thus, the regulator can not directly make its monetary transfer contingent upon the value of q, e and ξ.

For a utilitarian regulator, expected social welfare is

$$E(W) = V(q) - (1 + \lambda) E(T) + U(\pi).$$

Moreover, we assume that the hospital's utility function is given by a mean-variance function

$$U(\pi) = E(\pi) - \frac{\gamma}{2} Var(\pi),$$

where π is given by (2), $Var(\pi)$ is the variance of π, and γ is a measure of risk aversion.

The assumption of risk averse hospitals deserves some comment. The literature on health economics generally assumes that the key trade-off on the supply side is between patient selection and efficiency. In Ma (1994) for instance, the hospital observes a patient's severity and may choose to refuse this patient if the prospective price is smaller than the cost. In this paper, we assume that the hospital cannot fully predict the cost of a patient when deciding to accept or refuse treatment. In this context, a patient will only be treated if the payment is high enough to cover both the expected cost of the patient and the risk borne by the hospital. The value of this risk premium has to be determined so as to prevent the hospital from dumping patients. We believe that the assumption of risk aversion captures this idea of patient selection when the hospital is not able to fully predict patients costs well. If the hospital's manager is risk averse and the regulator is risk neutral, the optimal contract might include some provision for risk sharing.

In this paper, we have chosen to adopt a positive approach. Rather than deriving the optimal payment system given our assumptions (which would be the normative approach), we have chosen to restrict attention to a mixed linear payment which is close to several payment rules used in practice. Namely, we consider that patients with exceptionally costly stays (i.e. outlier patients) are reimbursed by a cost sharing contract, whereas inlier patients are reimbursed by a flat payment. Given this form of payment, our aim is to determine the threshold above which a patient has to be considered as an outlier patient, as well as the amount of the flat payment and the value of the cost sharing parameter.

Assume that the regulator can commit to the following unit payment rule

$$\begin{cases} T = p \text{ if } C < C_H \\ T = (1-\alpha)\,p + \alpha C \text{ if } C > C_H \end{cases}, \tag{3}$$

where T is the payment per patient made by the regulator to the hospital for each patient, p is a fixed unit price, $0 \leq \alpha \leq 1$ is a cost sharing parameter, and C_H is a unit cost threshold[7].

[7] Note that (3) is close to the US payment rule which could be written as

$$\begin{cases} T = p \text{ if } C < \overline{C} \\ T = p + \alpha\left(C - \overline{C}\right) \text{ if } C > \overline{C} \end{cases},$$

where $\overline{C}$ is a cost threshold.

Note also that (3) is close to the payment rule used in Switzerland. Indeed, this rule can be written as

$$\begin{cases} T = p - \alpha\left(\underline{C} - C\right) \text{ if } C < \underline{C} \\ T = p \text{ if } \underline{C} < C < \overline{C} \\ T = p + \alpha\left(C - \overline{C}\right) \text{ if } C > \overline{C} \end{cases}.$$

A patient whose cost is below a threshold $\underline{C}$ is considered as a low outlier patient, whereas a patient whose cost is above $\overline{C}$ is considered as a high outlier patient (see *www.hospvd.ch/ise/apdrg*). In this paper, we concentrate on high outlier patients (whose proportion is higher than the proportion of low outlier patients), since they

The sequence of events is as follows. The regulator commits to payment rule (3). The hospital chooses q^* and e^*. Then it observes ξ. The regulator observes C, and the hospital is paid according to (3).

Note that given the regulator's commitment on p, α and C_H, the hospital will choose the optimal levels of q^* and e^* which maximize its utility. The ex post unit cost of the hospital will be $C = c + q^* - e^* + \xi$. Hence, given q^* and e^*, there exists a value of ξ, say ξ_H, such that the unit cost threshold can be written $C_H = c + q^* - e^* + \xi_H$. Then, from the hospital's point of view, $\text{Prob}(C < C_H)$ is equivalent to $\text{Prob}(\xi < \xi_H)$.

The same reasoning can be applied to the regulator. Indeed, at the equilibrium, the regulator knows the levels of q^* and e^* which will be chosen (as long as the hospital chooses q^* and e^* before the observation of ξ).[8] Insofar as c is common knowledge, determining a unit cost threshold $C_H = c + q - e + \xi_H$ is equivalent to determining a threshold value ξ_H.[9]

Following the previous argument, this unit payment is equivalent to

$$\begin{cases} T = p \text{ if } \xi < \xi_H \\ T = (1-\alpha)\,p + \alpha C \text{ if } \xi > \xi_H \end{cases},$$

when e and q are optimally determined.

In addition to the unit payment, we assume that a lump-sum payment t is included in the total payment. Given the distribution of ξ, the total expected payment can be written as

$$E\left(T.x\left(q\right)+t\right) = \left\{F\left(\xi_H\right)p + \left(1 - F\left(\xi_H\right)\right)\left[\left(1-\alpha\right)p + \alpha C\right]\right\}x\left(q\right) + t. \tag{4}$$

The three instruments of the regulator are now α, p and ξ_H.[10]

do introduce the need for cost sharing.

[8] This is so even though he cannot verify those levels of effort.

[9] This argument is developped by Dunne and Loewenstein (1995) in a slightly different context.

[10] Note that the form of payment (3) differs from conventional developments in contract theory which deal with the trade-off between the risk sharing effect and the moral hazard effect. Indeed, in Kawasaki and McMillan (1987) and McAfee and McMillan (1986) among others, the optimal sharing parameter is determined in expectation over θ. Note also that our contract would obviously be dominated by a contract which would also give a mixed payment even for inlier patients. This contract would propose two different cost sharing

To determine the optimal linear contract, given the payment (4), the regulator has to find the optimal values of α, p and ξ_H taking the hospital's optimal response to any particular contract offered into account. We can solve this problem in a two-stage principal agent model:

(i) First, we solve for the hospital's expected utility maximizing choice of quality enhancement and cost reduction efforts for any given contract.

(ii) Then, we determine the regulator's expected welfare maximizing choice of α, p and ξ_H, given the hospital's optimal response.

2.1 The hospital's choice of quality and cost reduction effort

We assume that the hospital chooses the level of q and e before observing the random variable ξ.[11] Therefore, when it faces the contract (3), its profit is

$$\pi = [p - c + e - q - \xi]\, x(q) - \varphi(q+e) \text{ if } \xi < \xi_H \tag{5}$$

and

$$\pi = (1-\alpha)[p - c + e - q - \xi]\, x(q) - \varphi(q+e) \text{ if } \xi > \xi_H \tag{6}$$

From (5) and (6), we have

$$Var(\pi/\xi < \xi_H) = x^2(q)\, Var(\xi/\xi < \xi_H),$$

and

$$Var(\pi/\xi > \xi_H) = (1-\alpha)^2 x^2(q)\, Var(\xi/\xi > \xi_H).$$

The hospital's utility is then

$$U(\pi) = [1 - F(\xi_H)]\{(1-\alpha)[p - c + e - q - E(\xi/\xi > \xi_H)]\, x(q)$$

parameters (α_L and α_H), two different fixed prices (p_L and p_H), and would for instance be written as

$$\begin{cases} T = p_L + \alpha_L (C - p_L) \text{ if } C < C_H \\ T = p_H + \alpha_H (C - p_H) \text{ if } C > C_H \end{cases}.$$

However, we could not find a clear characterization of this contract.

[11]Thus, we depart from the analysis of Laffont and Rochet (1998) who derive the optimal contract in a context of adverse selection, moral hazard and risk aversion while considering an ex ante risk, i.e. a risk that materializes before the agent chooses the effort level.

$$
\begin{aligned}
&-\varphi(q+e)-\frac{\gamma}{2}(1-\alpha)^2 x^2(q)\,Var(\xi/\xi>\xi_H)\}\\
&+F(\xi_H)\{[p-c+e-q-E(\xi/\xi<\xi_H)]\,x(q)\\
&-\varphi(q+e)-\frac{\gamma}{2}x^2(q)\,Var(\xi/\xi<\xi_H)\}+t.
\end{aligned}
$$

Let us write

$$E_O=E(\xi/\xi>\xi_H),\ E_I=E(\xi/\xi<\xi_H),\ Var_O=Var(\xi/\xi>\xi_H)\text{ and }Var_I=Var(\xi/\xi<\xi_H).$$

The first order conditions with respect to e and q when φ is quadratic can be written

$$e=d[1-\alpha(1-F(\xi_H))]\,x(q)-q, \tag{7}$$

and

$$
\begin{aligned}
0 \;=\;& [F(\xi_H)-1]\{\frac{2(e+q+d(1-\alpha)x(q))}{d}\\
&-2(1-\alpha)x'(q)[p-c+e-q-E_O-\gamma(1-\alpha)x(q)\,Var_O]\}\\
&-\frac{2F(\xi_H)}{d}\{e+q+dx'(q)(-p+c-e+q+E_I)\\
&+x(q)\left(d+dVar_I\gamma x'(q)\right)\}.
\end{aligned}
\tag{8}
$$

(8) defines a function $p(\alpha,q,\xi_H)$. When this function is strictly increasing with respect to q and ξ_H, it can be inverted to yield a function $q(\alpha,p,\xi_H)$. This function and equation (7) define the hospital's optimal choice of quality and cost reduction effort in response to any contract which specifies α,p and ξ_H.

2.2 The regulator's choice of contract

Consider now the regulator's optimal choice of p, α and ξ_H given $e(p,\alpha,\xi_H)$ and $q(p,\alpha,\xi_H)$. Expected welfare can be written as

$$EW(p,\alpha,\xi_H)=V(q(p,\alpha,\xi_H))-(1+\lambda)\,ET(p,\alpha,\xi_H)+U(\pi).$$

The regulator has to maximize $EW(p,\alpha,\xi_H)$ subject to the participation constraint

$$U(\pi)\geq 0,$$

where 0 is the reservation level of utility. Given (7), $ET(.)$, $U(.)$ and given the fact that the constraint is binding, the regulator's optimization problem becomes

$$\begin{aligned} EW(.) &= V(q(.)) - (1+\lambda)\, x(q(.))\,[c + E_O + 2q(.) + F(\xi_H)(E_I - E_O)] \\ &\quad -\frac{1}{2}(1+\lambda)\, x^2(q(.))\, A, \end{aligned} \tag{9}$$

with

$$A = d\alpha^2 [1 - F(\xi_H)]^2 - d + \gamma F(\xi_H)\, Var_I + \gamma Var_O (1-\alpha)^2 [1 - F(\xi_H)].$$

The first order condition with respect to p is

$$\begin{aligned} 0 &= \frac{dq}{dp}\{V_q(.) - (1+\lambda)\,[x'(q(.))\,[c + E_O + 2q(.) \\ &\quad + F(\xi_H)(E_I - E_O)] + 2x(q(.)) + x'(q(.))\, x(q(.))\, A]\}. \end{aligned} \tag{10}$$

The first order condition with respect to α is

$$\begin{aligned} &\frac{dq}{d\alpha}\{V_q(.) - (1+\lambda)\,[x'(q(.))\,[c + E_O + 2q(.) \\ &+ F(\xi_H)(E_I - E_O)] + 2x(q(.)) + x'(q(.))\, x(q(.))\, A]\} \\ &= \frac{1}{2}(1+\lambda)\, x^2(q(.))\, \frac{dA}{d\alpha}. \end{aligned} \tag{11}$$

The first order condition with respect to ξ_H is

$$\begin{aligned} &\frac{dq}{d\xi_H}\{V_q(.) - (1+\lambda)\,[x'(q(.))\,[c + E_O + 2q(.) + F(\xi_H)(E_I - E_O)] \\ &+ 2x(q(.)) + x'(q(.))\, x(q(.))\, A]]\} \\ &= \frac{1}{2}(1+\lambda)\, x^2(q(.))\, \frac{dA}{d\xi_H} \\ &+ (1+\lambda)\, x(q(.)) \left[\frac{\partial E_O}{\partial \xi_H} + f(\xi_H)(E_I - E_O) + F(\xi_H)\frac{\partial (E_I - E_O)}{\partial \xi_H}\right]. \end{aligned}$$

When $x(q) \neq 0$, we obtain the following condition from (10) and (11)

$$\frac{dA}{d\alpha} = 0,$$

which yields the optimal value of the cost sharing parameter

$$\alpha = \frac{\gamma Var_O}{d[1 - F(\xi_H)] + \gamma Var_O}. \tag{12}$$

Note that the optimal level of α is decreasing in the moral hazard parameter d, increasing in the conditional variance Var_O and increasing in the risk aversion parameter γ. However the effect of ξ_H on the value of α is more ambiguous. Indeed, we have

$$\frac{\partial \alpha}{\partial \xi_H} \lesseqgtr 0 \text{ if } \frac{\partial Var_O}{\partial \xi_H}(1 - F(\xi_H)) + Var_O f(\xi_H) \lesseqgtr 0 \Leftrightarrow \frac{\frac{\partial Var_O}{\partial \xi_H}(1 - F(\xi_H))}{-Var_O f(\xi_H)} \gtreqless 1, \tag{13}$$

where obviously $\frac{\partial Var_O}{\partial \xi_H} < 0$.

Condition (13) can be interpreted in terms of the elasticity of the conditional variance Var_O with respect to $(1 - F(\xi_H))$, i.e. the probability that $\xi > \xi_H$. This elasticity can be written as

$$e_{[Var_O,\ (1-F(\xi_H))]} = \frac{\frac{\frac{\partial Var_O}{\partial \xi_H}}{Var_O}}{\frac{-f(\xi_H)}{(1-F(\xi_H))}} = \frac{\frac{\partial Var_O}{\partial \xi_H}(1 - F(\xi_H))}{-Var_O f(\xi_H)}. \tag{14}$$

From (13) and (14), we have[12]

$$\frac{\partial \alpha}{\partial \xi_H} \lesseqgtr 0 \text{ if } e_{[Var_O,\ (1-F(\xi_H))]} \gtreqless 1.$$

We now compute the optimal value of ξ_H. From (10) and (15), this value is determined by the following equation

$$0 = \frac{\partial E_O}{\partial \xi_H} + f(\xi_H)(E_I - E_O) + \frac{1}{2}x(q(.))\frac{dA}{d\xi_H} + F(\xi_H)\frac{\partial(E_I - E_O)}{\partial \xi_H}, \tag{15}$$

with

$$\begin{aligned} \frac{dA}{d\xi_H} &= 2d\alpha^2 f(\xi_H)[F(\xi_H) - 1] + \gamma\left[\frac{\partial Var_I}{\partial \xi_H}F(\xi_H) + Var_I f(\xi_H)\right] \\ &\quad + \gamma(1-\alpha)^2\left[\frac{\partial Var_O}{\partial \xi_H}(1 - F(\xi_H)) - Var_O f(\xi_H)\right]. \end{aligned}$$

Once α is determined, p can be obtained from (10) and (8). Indeed, with the linear contract the unit price can be shown to be equal to

$$\begin{aligned} p &= \frac{V'(q)}{(1+\lambda)x'(q)} - \frac{1}{2x'(q)[1 - \alpha(1 - F(\xi_H))]} \\ &\quad \left\{-2\alpha x(q)(1 - F(\xi_H))\left[(1-\alpha)^2 x(q)(d + \gamma Var_O)\right] + F(\xi_H)(E_I - E_O + Xx(q))\right\}, \end{aligned} \tag{16}$$

[12] An example of the value of α and $e_{[Var_O,\ (1-F(\xi_H))]}$ is given in section 3 for the special case of a gamma distribution.

with

$$X = 2d\alpha\left(1-\alpha\right) + \gamma\left(Var_I - Var_O\left(1-\alpha\right)^2\right) + d\alpha^2 F\left(\xi_H\right).$$

Since the regulator dislikes leaving rents to the hospital, total expected welfare can be enhanced if the agency is able to extract those rents through a fixed charge

$$t = \frac{x^2\left(q\right)}{2x'\left(q\right)}\left\{4\left[\alpha\left(1-F\left(\xi_H\right)\right)-1\right] + x'\left(q\right)\left[\left(1-\alpha\right)^2\left(d-\gamma Var_O\right) + F\left(\xi_H\right)X\right]\right\}. \tag{17}$$

Our results are summarized in the following proposition.

Proposition 1 *When the hospital is risk averse and the regulator is risk neutral, the optimal contract, given the form of payment (3), is*

$$\begin{cases} T = p + t \text{ if } C < C_H \\ T = \left(1-\alpha\right)p + \alpha C + t \text{ if } C > C_H \end{cases}, \tag{18}$$

with

- $\alpha = \frac{\gamma Var_O}{d[1-F(\xi_H)]+\gamma Var_O}$,
- p *and* t *respectively given by (16) and (17),*
- $C_H = c - e + q + \xi_H$*, when* q *is the optimal provided quality given by (8),* e *is the optimal effort level given by (7), and* ξ_H *solves condition (15).*

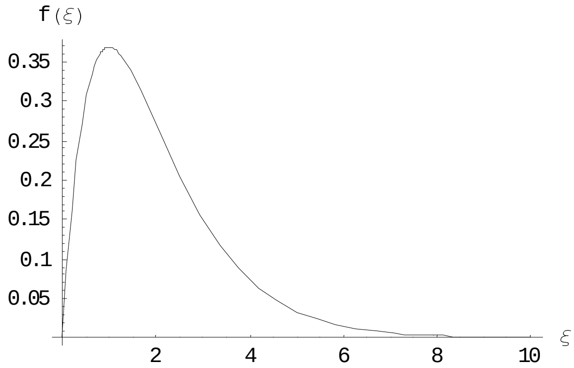
The optimal value of α trades off the risk sharing effect and the moral hazard effect, whereas the value of p induces the optimal choice of quality. If the regulator can extract the rent by means of a fine, our initial four-way trade-off is solved. As the hospital's participation constraint is binding, the provider can accept this contract.

3 Numerical simulations of the payment mechanism

To provide some numerical simulations of the payment mechanism, let us assume that the random variable affecting hospital's cost follows a Gamma distribution with values $\alpha = 2$ and $\theta = 1$ on the interval $[0, \infty[$.[13] The gamma density function is then

$$\begin{aligned} f(\xi) &= \xi e^{-\xi} \text{ if } \xi > 0 \\ &= 0 \text{ elsewhere.} \end{aligned}$$

It is illustrated by the following figure



Probability density of ξ

The cumulative density function is $F(\xi) = 1 - (1+\xi)e^{-\xi}$. Then, we can compute

$$E_O = \frac{\int_{\xi_H}^{\infty} x f(x)\, dx}{1 - F(\xi_H)} = 1 + \xi_H + \frac{1}{1+\xi_H},$$

$$E_I = \frac{\int_0^{\xi_H} x f(x)\, dx}{F(\xi_H)} = 2 + \frac{\xi_H^2}{1 + \xi_H - e^{\xi_H}},$$

$$Var_O = \frac{\int_{\xi_H}^{\infty} x^2 f(x)\, dx}{1 - F(\xi_H)} - [E(\xi/\xi > \xi_H)]^2 = \frac{2 + \xi_H(4+\xi_H)}{(1+\xi_H)^2},$$

$$\begin{aligned} Var_I &= \frac{\int_0^{\xi_H} x^2 f(x)\, dx}{F(\xi_H)} - [E(\xi/\xi < \xi_H)]^2 \\ &= \frac{2 + 2e^{2\xi_H} + \xi_H(4+\xi_H) - e^{\xi_H}[4 + \xi_H(4 + \xi_H(\xi_H - 1))]}{(1+\xi_H - e^{\xi_H})^2} \end{aligned}$$

[13]Several empirical studies show that most of the cost distributions observed within a DRG can be fitted with a gamma distribution (see e.g. Marazzi et al. (1998)).

and

$$\alpha = \frac{1}{1 + \frac{d(1+\xi_H)^3 e^{-\xi_H}}{\gamma[2+\xi_H(4+\xi_H)]}}.$$

In this example, note that we have

$$\frac{d\alpha}{d\xi_H} > 0, \text{ since } e_{[Var_O,\,(1-F(\xi_H))]} = \frac{2}{2+\xi_H\,(4+\xi_H)} < 1.$$

The following table gives the optimal values of α and ξ_H for different values of parameters γ and d.[14] Note that d reflects the magnitude of moral hazard.[15] Besides, estimating the risk aversion coefficient is a difficult task. Using data from Japanese subcontracting, Kawasaki and McMillan (1987) show that the estimated absolute risk-aversion coefficient of subcontractors ranges from 0.3 to 0.003. Therefore, we have chosen γ to be close to these values in our example.

$\gamma \backslash d$	0.25	0.5	1
0.3	$\alpha = 0.8$ $\xi_H = 1.74$	$\alpha = 0.7$ $\xi_H = 2$	$\alpha = 0.57$ $\xi_H = 2.21$
0.15	$\alpha = 0.7$ $\xi_H = 2$	$\alpha = 0.57$ $\xi_H = 2.21$	$\alpha = 0.42$ $\xi_H = 2.38$
0.05	$\alpha = 0.48$ $\xi_H = 2.31$	$\alpha = 0.33$ $\xi_H = 2.45$	$\alpha = 0.21$ $\xi_H = 2.52$

As depicted by the previous table, the trade-off between risk sharing and incentives to reduce costs not only relies on α but also on the cost threshold (ξ_H in our model) above which a patient is considered as an outlier patient. According to Newhouse (2002), US hospitals are given an additional payment (for outlier patients) equal to 80 percent of the costs above a threshold (i.e. $\alpha = 0.8$) . Furthermore, the outlier threshold is set such that outlier payments will be 5 percent of total payments. If we assume that ξ still follows the above gamma distribution, then we have $\text{Prob}(\xi > \xi_H) = 0.05$ for $\xi_H = 4.75$. For instance, for a value

[14] We check that the second order conditions are satisfied.

[15] More exactly, if we do not consider the problem of quality incentives, obviously d represents the difference between the total cost under a fixed price contract and the total cost under a cost reimbursement contract.

$\alpha = 0.8$, a cost threshold $\xi_H = 4.75$ is much higher than the optimal values we find in the table. This suggests that hospitals' contracts seem designed rather to share risks than to give hospitals incentives to reduce costs.

4 Conclusions

We have considered hospitals' payment problem as a regulation problem. Specifically, we analysed a demand constrained case where the quality of care is observable by patients but unverifiable by the regulator and where public funds are socially costly.

Our main contribution is to examine an informational structure which is not considered in other models. Indeed, the cost sharing rate is not the same for different values of patient severity. This allows us to consider a contract which is very close to the contracts proposed in practice. Our mechanism solves the trade-off between moral hazard and risk sharing. It induces the optimal decentralized choice of quality and cost levels. It also takes into account distributional concerns resulting from the shadow cost of public funds. The optimal contract is determined by three endogenous variables: the cost sharing parameters, the fixed part of the payment and the total cost threshold above which a patient has to be considered as an outlier patient. Moreover, we have shown that, to extract hospital rent, the regulator has to require the hospitals to pay a fixed lump-sum charge.

However, as in related models, the main drawbacks of our optimal policy arise from informational requirements. The optimal contract requires information on the risk aversion parameter and on the measure of moral hazard. Another limitation is that the tax that the hospitals may have to pay could be difficult to enforce. Finally, we have not considered the possibility of competition between hospitals.

References

[1] ALLEN R. and GERTLER P. [1991], Regulation and the provision of quality to heterogeneous consumers : the case of prospective pricing of medical services, *Journal of Regulatory Economics*, vol 3, pp 361-375.

[2] ARMSTRONG M., COWAN S. and VICKERS J. [1995], Regulatory Reform, Cambridge, MIT Press.

[3] CHALKLEY M. and MALCOMSON J. M. [1998], Contracting for health services when patient demand does not reflect quality, *Journal of Health Economics*, vol 17 pp. 1-19.

[4] CHALKLEY M. and MALCOMSON J. M. [1998a], Contracting for health services with unmonitored quality, *The Economic Journal*, 108, pp 1093-1110.

[5] CHALKLEY M. and MALCOMSON J. M. [2002], Cost sharing in health service provision: an empirical assessment of cost savings, *Journal of Public Economics*, 84, pp 219-249.

[6] DRANOVE D. [1984], An empirical study of a hospital-based home care program, *Inquiry*, 22, pp 59-66.

[7] DRANOVE D. [1987], Rate-setting by DRG and hospital specialization, *Rand Journal of Economics*, 18, pp 417-427.

[8] DUNNE, S.A. and LOEWENSTEIN, M.A. [1995], Costly Verification of Cost Performance and the Competition for Incentive Contracts, *Rand Journal of Economics,* 26 (4), 690-703.

[9] ELLIS R. P. and MCGUIRE T. [1986], Provider behavior under prospective reimbursement : cost sharing and supply, *Journal of Health Economics*, 5 (2), pp 129-51.

[10] ELLIS R. P. and MCGUIRE T. [1990], Optimal payment systems for health services, *Journal of Health Economics*, 9 (4), pp 375-96.

[11] HOLMSTROM B. and MILGROM P. [1987], Aggregation and linearity in the provision of intertemporal incentives, *Econometrica*, 55, pp 303-328.

[12] HOLMSTROM B. and MILGROM P. [1991], Multitask principal agent analyses : incentives contracts, asset ownership and job design, *Journal of Law, Economics and Organization*, vol 7, pp 24-52.

[13] KAWASAKI S. and MCMILLAN J. [1987], The design of contracts : evidence from Japanese subcontrating, *Journal of the Japanese and International Economies*, 1, pp 327-349.

[14] LAFFONT J.-J. and ROCHET J-C. [1998], Regulation of a risk averse firm, *Games and Economic Behavior*, 25, 149-173.

[15] LAFFONT J.-J. and TIROLE J. [1993], A theory of Incentives in Procurement and Regulation, MIT Press, Cambridge.

[16] LOEB M. and MAGAT W.A. [1979], A decentralized method of utility regulation, *Journal of Law and Economics*, 22, pp. 399-404.

[17] MA C. t. A. [1994], Health care payment systems : cost and quality incentives, *Journal of Economics and Management Strategy*, 3 (1), pp 93-112.

[18] MA C. t. A. [1998], Health care payment system: cost and quality incentives - reply, *Journal of Economics and Management Strategy*, 3, pp 139-142.

[19] MARAZZI A., PACCAUD F., RUFFIEUX C. and BEGUIN C. [1998], Fitting the dstributions of length of stay by parametric models, *Medical Care*, 36, pp. 915-927.

[20] MCAFEE R. P. and MCMILLAN J. [1986], Bidding for contract : a principal agent analysis, *Rand Journal of Economics*, 17, pp 326-338.

[21] MCCLELLAN M. [1997], Hospital reimbursement incentives : an empirial analysis, *Journal of Economics and Management Strategy*, 6 (1), pp. 91-128.

[22] NEWHOUSE J.-P. [1996], Reimbursing health plans and health providers : efficiency in production versus selection, *Journal of Economic Literature*, vol XXXIV, pp 1236-1263.

[23] NEWHOUSE J.-P. [2002], Pricing the priceless: a health care conondrum, MIT Press.

[24] POPE G. C. [1989], Hospital non price comptetition and medicare costs policy, *Journal of Health Economics*, 9 (3), pp 237-51.

[25] SHARMA R. [1998], Health care payment system: cost and quality incentives - comment, *Journal of Economics and Management Strategy*, 7, pp 127-137.

[26] SHLEIFER A. [1985], A theory of yardstick competition, *Rand Journal of Economics*, 16, pp 319-327.