CHAIRE SANTÉ

UNE CHAIRE DE
LA FONDATION DU RISQUE

Cahiers de la Chaire Santé

Asymmetric information and pooling contracts in hospital sector

Auteurs : Michel Mougeot, Florence Naegelen

N°2 - Octobre 2009

Asymmetric information and pooling contracts in hospital sector

Michel Mougeot*[G] and Florence Naegelen*
*University of Franche-Comte, CRESE, UFR SJEPG,
45 D Avenue de l'Observatoire, 25030 Besançon Cedex, France
[G]IEMS, University of Lausanne, Lausanne, Switzerland
E-mail address: michel.mougeot@univ-fcomte.fr

May 7, 2009

Abstract

Most of regulators in health care systems use pooling contracts such that payment do not depend on the level of severity. This policy is motivated by concerns about the moral hazard problem. In this paper, we show that it can be optimal when patient severity is private information because of the non-responsiveness phenomenon. We show in which cases the hospital may be non responsive to the regulator objective under adverse selection. We exhibit necessary conditions under which pooling contracts are optimal and we characterize these mechanisms when the hospital is self-interested and perfectly altruistic. In the first case, the fixed payment is equal to the cost of treating the patient with the highest severity whereas it is equal to the mean value of the treatment cost in the second one.

Keywords: Hospital regulation, patient severity, non-responsiveness, altruism, adverse selection.

JEL Classification: I1; L3; D8.

We gratefully acknowledge the financial support of the Risk Foundation (Health, Risk and Insurance Chair, AGF)

1. Introduction

Most of developing countries are currently implementing a prospective payment system under which hospitals are paid a fixed amount per admission for a given diagnosis. In a moral hazard setting, this fixed-price contract is a high-powered contract that gives the hospital socially incentives to reduce costs and to produce care efficiently. In practice, this mechanism is based on a relative performance evaluation, the payment received by a hospital for a given treatment falling within a specific Diagnosis Related Group being based on the average cost of the treatment in similar hospitals. This yardstick competition works because it does not let an inefficient choice by a hospital influence the price that it receives.[1] In most of countries, this prospective price policy does not adjust the payment for the severity of illness[2]: the providers receive the same payment for any admission in a given diagnosis whatever the degree of severity.

The adoption of this policy is mainly motivated by concerns about the moral hazard problem. However, the Diagnosis Related Groups (DRGs) classification is often such that there is a substantial variation in the cost of treating patients within some groups. This within DRG-variance arises because of differences in efficiency but also because of differences in the severity of illness of patients. Under adverse selection, new insights must be considered. When patient severity is not observable by the payer, the hospital may earn a rent when facing a low-cost patient if the payment received for a patient in a given DRG is calculated on the basis of the average cost incurred for that DRG nationally. As noted by Laffont and Tirole (1993), the main drawback of yardstick competition is that idiosyncrasies can prevail over common features. Though patients in a given DRG are often non comparable, hospital price regulation is mainly based on a fixed-price policy that solves the moral hazard problem without solving the adverse selection problem. In other words, most of regulators in health care systems use pooling contracts. According to Chalkley and Malcomson (2002), there are a number of reasons for this. For instance, providers could in fact receive little rent because they have not actually much better information about costs than regulators or because hospitals are altruistic and treat high cost patients even if

[1] See Shleifer (1985).

[2] An exception is the outlier payment policy that introduces retrospective factors for exceptionally costly patients (see Mougeot and Naegelen (2009)). For instance, Medicare PPS involves some cost sharing rules for these outlier patients. They are such that outlier payments are 5% of the total.

the payment does not cover the cost. Another reason is that prospective payments avoid costly cost monitoring procedures. The aim of this paper is to find some other explanations of this pooling contract practice and to show that efficiency can involve pooling. As under moral hazard, a ffxed-price contract is optimal, we focus on the adverse selection setting to exhibit cases in which a pooling contract is optimal.

In the theoretical literature, the main models of hospital regulation are moral hazard models. This is the case of the papers of Ellis and Mc Guire (1990), Ma (1994), Chalkley and Malcomson (1998) or Mougeot and Naegelen (2005). In this framework where adverse selection is not present, a fixed-price contract is always optimal when the regulator can use a lump-sum transfer to extract the provider's rent. As patient heterogeneity is not taken into account, this contract does not depend on the level of severity. Adverse selection is considered by De Fraja (2000) who assumes that each patient is characterized by a parameter which denotes her ability to benefit from the treatment (i. e. the opposite of the severity) and that efficiency varies across hospitals. Moreover, he assumes that the social benefit of treating a patient decreases with the severity. Under these assumptions, the optimal mechanism is characterized by a payment schedule depending on the efficiency parameter and a cut-off value under which a patient is dumped (which implies that patients with a high degree of severity are not treated). No pooling contract appears as optimal. In their empirical analysis, Chalkley and Malcomson (2002) consider simultaneously adverse selection and moral hazard in a Laffont-Tirole (1993) type model. The optimal contract solves a trade-off between productive efficiency and rent extraction and is such that the optimal transfer is a function of the severity. As in De Fraja, pooling does not occur at the optimum.

In all these models, either patient demand is random or it is a function of the quality of health care services. In both cases, it is independent of severity. However, these assumptions may be difficult to defend. The most severely ill patients are most costly but they receive more benefits from treatment than less severely ill patients. Then they have higher marginal values of quality. Hence, health care services demand may be increasing with severity and more quality elastic when severity increases. In the same way, the cost of treating a patient with severity /3 increases with the quality (or quantity) of health care services but also with /3. On the other hand, the social benefit can be considered as a function of the quality (or quantity) of health care services and as a function of severity. When social benefit of care and patient demand increase with severity, i.e., when the patient's type directly affects the regulator's objective function, new insights must be considered. First, if the marginal social benefit of health care services quantity perceived by a utilitarian regulator increases when severity increases, the first best quantity of services may increase with severity. In this case, a self-interested hospital can be not responsive to the regulator objective when severity is privately observed.[3] In an adverse selection setting, there may be a confiict between the optimal quantity and the feasible quantity. To ensure incentive compatibility, quantity must not increase with severity whereas it must increase with severity to achieve the first best. More generally, this phenomenon of non-responsiveness arises when there is a confiict between the regulator's preferences and the incentive constraints refiecting the hospital's preferences. Of course, this confiict depends on the degree of altruism of the provider.

In the following, we consider the regulation of a monopoly hospital treating patients characterized by the severity of their illness. Under adverse selection on this parameter, we show in which cases non-responsiveness forces the principal to use a pooling contract in which the same quantity of care and the same payment are implemented for any degree of severity. When the hospital is self-interested, we show that pooling contracts can be optimal when the social marginal benefit increases faster with severity than the virtual marginal cost. In contrast, when the hospital is perfectly altruistic, we show that pooling contracts can be optimal when the social marginal benefit of treatment increases with severity faster than the hospital marginal cost but more slowly than the marginal cost as perceived by the taxpayer. We exhibit necessary conditions under which pooling contracts are optimal for efficiency reasons and we characterize these mechanisms when the hospital is self-interested and perfectly altruistic. In the first case, the fixed payment is equal to the cost of treating the patient with the highest severity whereas it is equal to the mean value of the cost in the second one.

The paper is organized as follows. The model is presented in section 2. The optimal regulation of a self interested hospital is characterized in section 3 whereas the case of a perfectly altruistic hospital is analyzed in section 4. Some conclusions are drawn in section 5.

[3] See Guesnerie and Laffont (1984) for a general analysis of non-responsiveness.

2. The mode!

Let us consider the regulation of a monopoly hospital treating patients with a given diagnosis when patients are fully insured.

2.1. Assumptions

Assume that each patient is characterized by a parameter /3 which denotes the severity of her illness. The hospital observes the severity but the regulator only knows its distribution function. The uncertainty on /3 is represented by a cumulative distribution function $F(.)$ and an associated continuously differentiable density function $f(.) > 0$ on a support $[/3, /3]$, with $F(/3)/ f(/3)$ increasing in /3. The patients presenting for treatment at the hospital are a random sample from the distribution $F(/3)$. We assume that the marginal cost of treatment depends on the quantity of health care services as well as on the severity. More ill patients are assumed to be more costly. A hospital treating a patient /3 with a quantity of care services q has a cost function $C(q, /3)$ strictly increasing and convex in q, increasing in /3, with $C_{q\sim}(q, /3) > 0, V/3$.

Let $V(q, /3)$ denotes the benefit that the regulator (either a purchasing agency or a public insurer) attaches to having patient /3 treated with quantity q, with $V(q, /3)$ strictly increasing and concave in q. The influence of /3 on $V(q, /3)$ depends on the objective of the regulator. For instance, De Fraja (2000) assumes that the social benefit of treating a patient with severity /3 decreases with /3 (because the ability to benefit from the treatment decreases with /3)[4]. Chalkley and Malcomson (2002) do not assume that the social benefit decreases with /3 but they suppose that cost rises with /3 faster than benefit. On the contrary, Ma and Chone (2008) consider a managed care company maximizing the patient's utility less the payment to the physicians when the utility of a patient increases with the severity. In fact, this is an open question. There are probably kinds of diseases where a benevolent regulator would value treating a high cost patient more than a low cost patient. A caring provider could derive more utility from the act of providing medical services to patients with a high severity. In other respects, a benefit function increasing with /3 is in line with the principle of allocation ac-

[4] In De Fraja (2000), some patients benefit more than others because they are younger and not affected by other pathologies (and hence are likely to live longer). In our model, patients differ according to the severity of their illness which can be higher when they are older or affected by other pathologies.

cording to needs.[5] More generally, one can consider that $V(q, /3)$ is increasing in $/3$ when it represents the patient's benefit because the most severely ill patients receive more benefits from treatment than less severely ill patients. However, the benefit perceived by the regulator can incorporate other health policy issues and be different from the benefit perceived by the patient. Then $V(q, /3)$ may be either increasing or decreasing with $/3$. We will see in the following how the optimal mechanism depends on these assumptions on the social benefit function.

We assume that the hospital is partially benevolent and trades off its benefit and the benefit for the patients. If a is the degree to which the hospital takes the patient's benefit into account[6], the hospital utility can be written, when it receives a payment t

$$U(t, q, /3) = t - C(q, /3) + aV(q, /3) \tag{1}$$

The regulator maximizes a social welfare function equal to the sum of the net benefit of treatment and the hospital utility and takes a social cost of public funds $\grave{A}$ into account. After excluding altruistic preferences of the hospital to avoid undesirable double counting[7], social welfare can be written

$$W(t, q, /3) = V(q, /3) - (1 + \grave{A})t + t - C(q, /3) \tag{2}$$

2.2. Full information

Under complete information on the severity, the regulator would choose t^* and $q*$ maximizing $W(t, q, /3)$ in (2) subject to the participation constraint $U(t, q, /3) \sim 0$, if we normalize the minimum utility for which the hospital accepts a contract to 0, and the liability constraint $t - C(q, /3) \sim 0$[8]. The first best allocation is characterized by two functions $q^*(/3)$ and $t^*(/3) = t^*(q^*(/3))$ such that

$$V_q(q^*(/3) , /3) = (1 + \grave{A})C_q(q^*(/3), /3) \; V/3 \tag{3}$$

$$V_{qq}(q^*(/3)) - (1 + \grave{A})C_{qq}(q^*(/3), /3) <0 \; V/3 \tag{4}$$

$$t^*(q^*(/3))=C(q^*(/3),/3) \; V/3 \tag{5}$$

[5] See Culyer (1989). On the relationship between need and severity, see Culyer and Wagstaff (1993). Needs based allocation is often considered in rationing models which implies that only patients with severity greater than a threshold are treated (see for instance Cuff et al. (2007))

[6] We assume that c is common knowledge. See Jack (2005) for a model where c is private information and Ma and Choné (2008) for a bidimensional adverse selection model.

[7] See Hammond (1987) for a justification of excluding altruistic preferences from social welfare.

[8] When the limited liability constraint is satisfied, the participation constraint is also satisfied and can be neglected in the full information setting.

which implies that the marginal social benefit of the treatment must be equal to its marginal social cost as perceived by the tax payers. As $À > 0$, the hospital receives no rent and the price is equal to the cost of providing the optimal quantity. Moreover, as increasing $/3$ increases the social marginal cost of the treatment, the influence of the severity on the marginal social welfare of the treatment depends on the influence of $/3$ on $V(q, /3)$.

Applying the implicit function theorem to (3), it can be shown that the sign of $\dot{q}^{*}(/3)$ depends on the sign of

$$W_q = V_q - (1 + À)C_q \tag{6}$$

Under our assumptions on the cost function, $W_q < 0$ and $\dot{q}^{*}(/3) < 0$ if the social marginal benefit V_q is decreasing with $/3$. When $V_q > 0$, $\dot{q}^{*}(/3)$ may be positive if the social marginal benefit of the treatment increases faster than its marginal cost when patient severity increases. In this framework, the phenomenon of non-responsiveness[9] can occur and makes the screening of types difficult under adverse selection.

2.3. Adverse selection

If the severity is privately observed, the hospital can increase its utility by announcing $/3' =6 /3$. Then, the regulator has to design a policy maximizing the expected social welfare subject to the constraints imposed by its lack of information. From the revelation principle (Myerson (1979)), we know that the optimal mechanism can be summarized by two functions $\{q(/3), t(/3)\}$, where $q(/3)$ and $t(/3)$ are respectively the requested quantity of health care services and the payment of the hospital when it announces $/3$. Thus the regulator's problem is to choose these functions maximizing[10]

$$EW(t, q, /3) = Z \quad (V(q(/3), /3) - Àt(/3) - C(q(/3), /3))f(/3)d(/3)$$

subject to three types of constraints:

[9] See Guesnerie and Laffont (1984), Caillaud et al. (1988), Laffont and Martimort (2002, p. 53-55). For an analysis of the links between implementability and responsiveness, see Morand and Thomas (2003).

[10] For simplicity, we assume that the social benefit of treating a patient is so high that it is worth treating any patient.

i) No dumping constraints that ensure that the hospital is willing to treat any patient of type /3

$$U(t(/3), q(/3), /3) = t(/3) - C(q(/3), /3) + aV(q(/3), /3) \sim 0 \; V/3 \qquad (7)$$

ii) Expected budget constraint: as patients presenting for treatment are a random sample from F(/3), the hospital must balance its expected budget.

$$EH = Z_{'3}^{'3} (U(t(/3), q(/3), /3) - aV\,(q(/3), /3))f(/3)d(/3) \sim 0 \qquad (8)$$

iii) Incentive compatibility constraints that ensure that the hospital reveals the true type of the patient

$$/3 \; E \; \arg\max_{'3'} U(t(/3'), q(/3'), /3) = t(/3') - C(q(/3'), /3) + aV(q(/3'), /3) \qquad V/3, V/3' \; (9)$$

Constraints i) are interim participation constraints whereas constraint ii) is an ex ante participation constraint. Usually, in incentives theory, when a principal and an agent contract before the agent discovers her type, the ex ante participation constraint replaces the interim participation constraints. Here those two constraints refer to two very different concerns. The expected budget constraint implies that the hospital accepts the contract for all the population of potential patients. The no dumping constraints imply that a hospital willing to participate is also willing to treat any peculiar patient. Note that these constraints are not redundant. On the one hand, (8) does not imply (7). On the other hand, if (7) is satisfied for any /3, the expected value of profit is greater than $-R_{'3}^{'3}$ aV(q(/3), /3))f(/3)d(/3) which does not ensure that (8) is verified.

Standard arguments imply that the necessary and sufficient conditions for incentive compatibility are given by the local optimality condition and the monotonicity constraint. When the hospital is partially altruistic, the local incentive compatibility constraint can be written

$$\dot{U}(/3) = -C_3(q(/3), /3) + aV_3(q(/3), /3) \qquad (10)$$

In the following, we assume that $\dot{U}(/3)$ is negative for all /3, which implies that $aV_{'3}(.) < 0$ or that cost increases with severity faster than $aV_{'3}(.)$[11]

[11] In this paper, we do not consider the countervailing issue that arises when $U_{(,6)}$ changes sign between ,6 and ,6. In some cases, it can also result in a pooling contract (see Lewis and Sappington (1989), Maggi and Rodriguez-Clare (1995)).

From (9), we obtain

$$\frac{dt(/3^0)}{d/3_0} — C_q(q(/3'), /3)\frac{dq(/3_0)}{d/3_0} + aV_q(q(/3'), /3)\frac{dq(/3_0)}{d/3_0} = 0 \sim\sim\sim\sim$$

and the identity

$$\frac{dt(/3)}{d/3} = (C_q(q(/3),/3) — \sim V_q(q(/3), /3))\frac{dq(/3)}{d/3} \qquad V/3 \qquad (11)$$

From the second order condition with respect to $/3'$ in $/3' = /3$ and the differentiation of (11) with respect to $/3,$ we obtain

$$(—C_{q\sim}(q(/3), /3) + aV_{q\sim}(q(/3), /3))\frac{dq(/3)}{d/3} \sim 0$$

and the monotonicity condition

$$q(/3) \quad 0 \text{ if } —C_{q\sim}(q(/3),/3)+aV_{q\sim}(q(/3),/3)<0V/3 \qquad (12)$$

$$q(/3) \sim 0 \text{ if } —C_{q\sim}(q(/3),/3)+aV_{qj}(q(/3),/3)>0V/3 \qquad (13)$$

Consequently, the monotonicity condition implies that the quantity requested from the hospital can be either decreasing or increasing with the severity according to the value of $aV_{q\sim}(.)$. In this setting, a confiict may arise between the feasible quantity (such that the hospital reveals its private information on patient severity) and the optimal quantity.

2.4. Optimal regulatory policy under adverse selection

Let us denote 'y the Kuhn and Tucker multiplier associated with the expected budget constraint. Then the expected Lagrangian can be written

$$EL = \int_{\sim}^{\sim} \{V(q(/3), /3)(1+\sim(\grave{A}—\sim))—(1+\grave{A})C(q(/3), /3)—(\grave{A}—\sim)U(t(/3), q(/3), /3)\}f(/3)d(/3)$$

As $U(/3) < 0\ V/3$, we have from (10)

$$U(/3) = U(/3) — \int_{\sim}^{\sim} (—C_{\sim}(q(s), s) + aV_{\sim}(q(s), s))ds \qquad (14)$$

EL can be rewritten after integration by part

$$\int_{\beta}^{s} \{V(q(e),e)(1+a(\grave{A}-y))-(^{1}+\grave{A})C(q(e),e)+$$

$$F(e)$$

$$(\grave{A}-y)\ f(e)(-C\,s(q(e),e)+aVs(q(e),e))\}\,f(e)d(e)-(\grave{A}--y)U(e) \quad (15)$$

and the optimal policy is characterized by a level of health services quantity $q(e)$ such that

$$\begin{aligned}&(V_q(q(e),e)(1+a(\grave{A}--y))-(1+\grave{A})C_q(q(e),e)\\&+(\grave{A}--y)F(e)/f(e)(-Cs_q(q(e),e)+aVs_q(q(e),e))f(e)=0\,v\,e\end{aligned} \quad (16)$$

and

$$\begin{aligned}&(V_{qq}(q(e),e)(1+a(\grave{A}--y))-(1+\grave{A})C_{qq}(q(e),e)\\&+(\grave{A}--y)F(e)/f(e)(-Cs_{qq}(q(e),e)+_{aVsqq(q(}e),e))f(e)<0\,v\,e\end{aligned} \quad (17)$$

Note that $(\grave{A}--y)$ cannot be negative at the optimum. If $\grave{A}>-y$, $U(e)=0.$ Then , the expected profit can be written

$$EH(q(e))=f$$

$$\overset{s}{\underset{3}{)}}$$

$$[(-aVs(q(e),e)+Cs\,(q(e),e))F(e)-aV\,(q(e),e)\,f(e)]d(e) \quad (18)$$

The expected budget constraint is satisfied if $a<a,$ with :

$$a=\overline{\qquad\qquad\qquad\qquad}\ fss\,Cs(q(e),e)F($$

$$e)d(e)$$

$$f_i\!:[Vs\,(q(e),e)F(e)+V(q(e),e)\,f(e)]d(e)$$

When $a>a$, $EH(q(e))<0$, which is impossible. Then $\grave{A}=-y$ and the first best is achieved. When $a<a$, $EH>0$ and $-y=0$. When $a=a$, the expected budget is balanced for a value $-y$ of the multiplier such that (18) is satisfied at equality and $EH=0$.

To characterize the cases in which pooling pricing policies are optimal, let us consider firstly the regulation of a self-interested hospital. Then we will consider the regulation of a perfectly altruistic provider ($a=1$).

3. Regulating a self-interested hospital

When the hospital is self-interested, a = 0, U = -(Y ~ < 0 and ˙q(/3) 0 V/3 because -$c_{q\sim}$(q(/3), /3) < 0 V/3. Moreover, as H(/3) = U(/3) V/3, EH > 0 at the optimum and 'y = 0. Then, the optimal policy is characterized by U(/3) = 0 and by a quantity of health care services q(/3) such that

$$V_q(q(/3),/3) -(1 + À)C_q(q(/3),/3) - ÀF(/3)/f(/3)C_{\sim q}(q(/3),/3) = 0 V/3 \qquad (19)$$

The optimal requested quantity q(/3) defined by (19) is such that the social marginal benefit is equal to the virtual marginal cost of treatment, that includes an information cost À$C_{\sim q}$ $\frac{F(j3)}{f(j3)}$. In this case, szgn ˙q(/3) = szgn (V_q - (1 + À + $\frac{À@F(\sim)=f()}{@\sim}$)C_q - $\frac{ÀF(\sim)}{f(\sim)}$$C_{q\sim}$). According to the signs of ˙q(/3) and ˙q(/3), three cases must be considered

3.1. The separating equilibrium

The separating equilibrium is obtained when ˙q(/3) < 0 and q˙(/3) < 0 for any /3 E [/3, /3]. The optimal requested quantity q(/3) is strictly decreasing in the degree of severity and lower than q^*(/3) for all /3 other than /3. To limit the informational rent, quantity is distorted downward and only the patient with the lowest severity receives the first best level of health care services quantity. q(/3) is the result of an optimal trade-off between rent extraction (reducing quantity reduces rent) and efficiency (reducing quantity reduces the social benefit of health care services). This case occurs when the social marginal benefit of the quantity of health care services increases more slowly than the marginal virtual cost of treatment when patient severity increases. In particular, it occurs when V_q (q(/3), /3) decreases with the severity[12] and when C_q(q(/3), /3) is convex in /3. For instance, if the regulator thinks that the ability to benefit from the treatment (measured for instance by the increase in QALYs) decreases with /3, V $_q$(.), representing the marginal benefit perceived by the regulator, and the patient's marginal benefit may vary in opposite directions with /3.

From the incentive compatibility condition, the optimal transfer is

$$t(q(/3)) = C(q(/3), /3) + \int_{\sim}^{3} C_{,}(q(s), s)ds \qquad (20)$$

[12] which corresponds to the hypothesis of De Fraja (2000)

The payment is equal to the cost plus an informational rent such that the hospital gets a positive rent on all treated patients except on the patient with the highest

:

severity. As q(/3) <0 , it is decreasing with /3.

3.2. The pooling equilibrium

:

When q(/3)and ˙q(/3) have opposite signs, non-responsiveness occurs. As ˙q(/3) cannot be positive, only one pooling equilibrium can occur when the hospital is self-interested. So we obtain the following proposition

Proposition 1. A necessary condition for the optimal contract to be pooling when the hospital is self interested is that

$$V_{q\sim}(q(/3),/3)> (1+ + \frac{\sim@(F(/3)=f(/3))}{@/3})C_{\sim q}(q(/3),/3) + \frac{F(/3)}{f(/3)} C_{\sim q\sim}(q(/3),/3)\ V/3$$

If the variation of the marginal social benefit with /3 is greater than the variation of the virtual marginal cost with /3, the optimal mechanism is such that q(/3) is strictly increasing with the degree of severity. Then there is a conflict between the regulator's desire to have the high severity patients receive more health services than the low severity patients and the monotonicity condition imposed by asymmetric information. Eliciting the true information on the severity would imply that the regulator chooses a quantity of health care decreasing with the severity but such a policy would reduce the expected social welfare. Then the regulator must not try to extract information on severity and must "bunch" all types of patients. To avoid a decrease of the expected social welfare, the regulator must leave all the informational rent to the provider and choose a contract that does not trade off efficiency and rent extraction.

According to Proposition 1, non-responsiveness occurs because the social marginal benefit of the treatment increases faster than its virtual marginal cost when patient severity increases. Regulator's preferences imply that the marginal (virtual) social welfare increases with the severity whereas the incentive constraint refiecting the self-interested hospital's preferences imply a decrease of the quantity of care with severity. In this case, pooling contracts are motivated by pure efficiency reasons. This situation cannot occur if the marginal benefit is decreasing with the severity when the marginal cost is convex in /3 but it can occur when V_q(q(/3), /3) increases with the severity and when C_q(q, /3) is concave in /3. For instance, it occurs when the marginal benefit as perceived by the regulator coincides with the patient's marginal benefit which increases with /3. Remark that Proposition 1 imply that a pooling contract is more likely to occur when the value of the shadow cost of public funds is low i.e., when the rent is socially less costly.

In this non-responsiveness context, any separable contract becomes very costly and the regulator must use a pooling allocation. The same quantity and payment will be implemented for all severity of illness. This allocation is the solution of the maximization of

$$\mathrm{Z}_{'3\,'3}\,(V(q, /3) - C(q, /3) - \grave{A}t)f(/3)d(/3) \qquad (21)$$

subject to

$$U(t, q, /3) = t - C(q, /3) \sim 0\ V/3 \qquad (22)$$

As U(.) is decreasing in /3, the harder participation constraint is that of /3 and the optimal transfer ^{e}tis such that the participation constraint associated with the highest cost patient /3 is saturated:

$$t = C(^{e}q, /3) \qquad (23)$$

Then the optimal pooling solution is ^{e}q, independent of /3 and such that

$$\mathrm{Z}_{33}(V_q(^{e}q,/3) - C_q(^{e}q,/3))f(/3)d(/3) = \grave{A}C_q(^{e}q,/3) \qquad (24)$$

q corresponds to an average level of health care services maximizing the net expected social welfare. Instead of achieving a goal of allocative efficiency, this contract results only in an average efficiency. In this case, the optimal contract consists of a fixed average amount of care independent of the severity and a fixed payment avoiding selection of patients. Insuring that no patient is dumped imply that this fixed price must be equal to the cost of treating the patient with the highest severity. Hence the provider earns a rent when treating any patient with a lower severity. This pooling contract corresponds to the usual practice of the Prospective Price Policy when no outlier payments are introduced for exceptionally costly patients.

3.3. The semi separating equilibria

:

When q(/3) changes sign between /3 and /3, the optimal mechanism is partially

:

pooling and partially separating. Assume that q(/3) changes sign one time on

[/3 , /3] [13]. As ˙q(/3) < 0, q(/3) can be increasing (resp. decreasing) then decreasing (resp. increasing) and the equilibrium can be separating (resp. pooling) for /3 lower than a cut-off value and pooling (resp. separating) for /3 greater than this cut-off. Let us assume for instance that q(/3) is increasing then decreasing, i. e., Vq~(q(/3), /3) - (1 + À + ÀôF(~)=f() @~)C~q(q(/3), /3) - ÀF(~) f() C~q~(q(/3), /3) positive for the low values of /3 and negative for the high values of/3. To determine the optimal cut-off e/3, the regulator has to maximize

E W (t, q, /3) = Z 3 ~ {V (q(/3), /3) - (1 + À)C(q(/3), /3)
β

-À(F(/3)/f(/3))C~(q(/3), /3)}f(/3)d(/3) - ÀU(/3)

under the constraint ˙q(/3) 0. If q(/3) is increasing then decreasing, the optimal solution eq(/3) is constant and equal to eq in [/3, e/3] and coincides with q(/3) on [e/3, /3], with

Z ~3 {Vq(eq, /3)(1 + aÀ) - (1 + À)Cq(eq, /3)-

À(F(/3)/f(/3))Cq~(eq, /3)}f(/3)d(/3) = 0

and eq = q(e/3) (See the proof in Appendix)

Then, for any level of severity lower than or equal to e/3, the requested quantity of care is constant and equal to q(/3). For any level of severity greater than /3,

[13] If j(3) changes sign several times, the generalization involves partitioning the interval [3, 3] so that ˙j(3) has the same sign in each subinterval.

q(/3) is decreasing with /3 and the contract is separating (see figure 1).

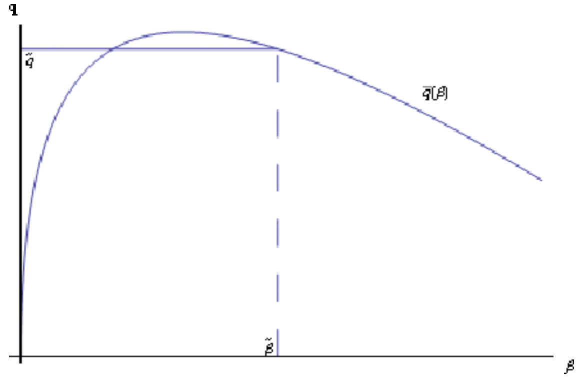

Figure 1

As for a patient ^{e}/3, the hospital rent is equal to U(^{e}/3) = R k k Ck(q(s), s)ds, the payment received by the hospital for a patient with gravity lower than or equal to ^{e}/3 is

Z k
t=C(eq,-/3)+ Ck(q(s), s)ds when/3 < ^{e}/3
k

and

Z k
t(q(/3)) = C(q(/3), /3) + Ck(q(s), s)ds when/3 ~ $^-$/3
k

The optimal payment is equal to a constant ^{et} when /3 < /3 ^{e} and is decreasing with /3 when /3 ~ ^{e}/3. It is a consequence of the decrease of W_q in /3 for the patients with severity higher than ·/3.

4. Regulating a perfectly altruistic hospital

Let us now consider the case of a perfectly altruistic hospital (a = 1). When U = -Ck + V $_k$ <0 V/3, the optimal policy is ^{b}q(/3, 'y) such that

V_q(^{b}q(/3, 'y), /3)(1 + À - 'y) - (1 + À)C_q(^{b}q(/3, 'y), /3)+
(À - 'y) F (/3)
f(/3) (-C_qk(^{b}q(/3,'y),/3) + V_qk(^{b}q(/3,'y),/3)) = 0 V/3 (25)

(À-'y)U(/) =0

'yEH=0

At the optimum, 'y cannot be greater than À (otherwise U(/) would tend to infinity). Note that the exp ected profit can be written

EH = Z 3 3 (U(ᵇq(/, 'y), /) - V (ᵇq(/, 'y), /))f(/)d/

= U(/) + Z 3 3 [(C3(ᵇq(/, 'y), /) - V3(ᵇq(/, 'y), /))F (/) - V (ᵇq(/, 'y), /)f(/)]d/

= U(/) + EH(ᵇq(/, 'y)) (26)

with ᵇq(/, 'y) solution of (25). Using the implicit function theorem, we have from (25)

sign ᵇdq / d'y = sign[-V_q(ᵇq(/,'y),/) + F(/) / f(/) (C_q3(ᵇq(/, 'y),/) - V_q3(ᵇq(/, 'y), /))]

Then

dEH / d'y = Z 3 3 ᵇdq / d'y (-V_q(ᵇq(/, 'y), /)+F (/) / f(/) (C_q3(ᵇq(/, 'y), /)-V_q3(ᵇq(/, 'y), /))f(/))d/ > 0

If EH > 0, 'y = 0 and U(/) = 0 and EH(ᵇq(/, 0)) > 0. As the expected profit is increasing with 'y, EH(ᵇq(/, 'y)) > 0 V'y > 0 which is impossible at the optimum because 'y > 0 implies EH = 0. Then 'y = 0 cannot be obtained at the optimum and EH(ᵇq(/, 0)) <0. Moreover, as EH(ᵇq(/, 'y)) is increasing with 'y and not strictly positive, the highest value of the expected profit is obtained when 'y = À. Then, at the optimum, 'y = À and either U(/) > 0 (if EH(ᵇq(/, À)) <0) or U(/) = 0 (if EH(ᵇq(/, À)) = 0). The first best is achieved and the optimal quantity q^*(/ , À) = q^*(/) is such that V_q(q^*(/) , /) = (1 + À)C_q(q^*(/), /)

To characterize the different equilibria, remind that ˙q(/) = $\overset{>}{<}$ 0 if Vq3 $\overset{>}{\underset{<}{=}}$ Cq3. As the first best is achieved, ˙q *(/) $\overset{>}{\underset{<}{=}}$ 0 if Vq3 $\overset{>}{\underset{<}{=}}$ (1 + À)C_q3. As in the self-interested case, three kinds of equilibria can be obtained.

4.1. Separating equilibria

Two separating equilibria can be obtained when $\dot{q}^*$ and $\dot{q}$ have the same sign for any $/3$ 2 $[/3, /3]$. Both functions are positive when $V_{q\sim} > (1 + À)C_q > C_{q\sim}$. Both are negative when $V_q < C_q < (1 + À)C_{q\sim}$. The requested quantity is either increasing or decreasing with the severity. As the hospital is highly altruistic and the expected buget constraint binding, rent extraction is not a concern for the regulator. Then the first best is achieved and no distortion of the quantity of care is needed.

The first separating equilibrium occurs when the marginal benefit of health care services increases faster than the marginal cost of the treatment taking the shadow cost of public finds into account, whereas the second occurs when the marginal cost increases faster than the marginal benefit. In particular, it occurs when V_q $(q\ (/3), /3)$ decreases with the severity.

From the incentive compatibility condition, the optimal transfer is

$$t(q^*(/3)) = C(q^*(/3), /3) - V\,(q^*(/3), /3) + \overset{3}{\underset{\sim}{Z}}\ (C_\sim(q^*(s), s) - V_s(q^*(/3), /3))ds + U(/3) \tag{27}$$

The payment is equal to the cost minus the benefit of care plus an informational non monetary rent such that the hospital gets a positive rent on all treated patients except on the patient with the highest severity plus a subsidy $U(/3)$ allowing to balance the expected budget. It is decreasing with $/3$ when $\dot{q} > 0$ and increasing with $/3$ when $\dot{q} < 0$

4.2. Pooling equilibrium

Pooling could occur in two cases: when $\dot{q}(/3) < 0$ and $\dot{q}^*(/3) > 0$ and when $q(/3) > 0$ and $\dot{q}^*(/3) < 0$. The first case arises when $C_{q\sim}(.) > V_{q\sim}(.) > (1{+}À)C_{q\sim}(.)$,
:

which is impossible. Then, only the second case is possible and arises when $C_{q\sim}(.) < V_{q\sim}(.) < (1 +))C_{q\sim}(.)$. Consequently, we obtain the following proposition

Proposition 2. A necessary condition for the optimal contract to be pooling when the hospital is perfectly altruistic is that

$$C_{qs}(q^*(/3), /3) < V_{q\sim}(q^*(/3)\ , /3) < (1 + À)C_{qs}(q^*(/3), /3)$$

When the hospital is perfectly altruistic, non-responsiveness occurs because the social marginal benefit of health care services increases faster than the marginal cost of treatment but more slowly than the social marginal cost of treatment as perceived by the tax payers (taking the shadow cost of public funds into account) when patient severity increases. Conditions stated by Proposition 2 are the opposite of conditions of Proposition 1. Regulator's preferences imply that the marginal social welfare decreases with the severity whereas the perfectly altruistic hospital's preferences imply an increase of the quantity of care with severity. Proposition 2 is rather restrictive. Firstly, V_q(.) must be increasing in /3. Secondly, V_q must belong to the interval]$C_{q\sim}$(.), (1 + À)$C_{q\sim}$[that can be narrow for the usual values of the shadow cost of public funds. In contrast with the self-interested case, pooling contract are more likely to be optimal when the value of À is high (because no socially costly expected monetary rents are left to the hospital when the expected budget balance constraint is binding).

As in the self-interested case, optimal screening of types works against efficiency. Any separable contract becomes very costly and the regulator must use a pooling contract for pure efficiency reasons. The same quantity and payment must be implemented for all severity of illness. This allocation is the solution of the maximization of

Z~'3 (V (q, /3) - C(q, /3) - Àt)f(/3)d(/3) (28)

As a price such that U(/3) = 0 does not ensure that the expected budget is balanced, the transfer must be such that Et = EC(q, /3). Then social welfare can be written

Z 3

(V (q, /3) - (1 + À)C(q, /3))f(/3)d(/3)

~

and the optimal pooling solution is ^{e}q, independent of /3 and such that

p Z(V_q(^{e}q; /3) - (1 + À)C_q(^{e}q, /3))f(/3)d(/3) = 0 (29) ~ In this case of bunching, the quantity of health care services maximizes the social net expected benefit whereas the price is equal to the mean value of the cost. As the provider is perfectly altruistic, deterring dumping is not costly. The hospital agrees to losses on high severity patients being offset by gains on low severity patients provided expected budget is balanced.

4.3. Semi separating equilibria

As in the self-interested case, when $q^{\cdot}\ (e)$ changes sign between e and $e,$ the optimal mechanism is partially pooling and partially separating. Assume for instance that $e)$ changes sign one time on $[e\ ,\ e]$. Several semi-separating equilibria can occur. When $^{\cdot}q(e) < 0$, $q(e)$ can be increasing (resp. decreasing) then decreasing (resp. increasing) and the equilibrium can be pooling (resp. separating) for e lower than a cut-off value and pooling (resp. separating) for e greater than this cutoff. In the same way, when $^{\cdot}q(e) > 0,$ the equilibrium can be either separating or pooling for e lower than a cut-off value and either pooling or separating for e greater than this cut-off. Using the same method than in 3.3, the semi-separating equilibria can be characterized in each cases.

5. Conclusion

Usual explanations of the prevalence of pooling contracts policies in health care systems are based either on the provider's altruism or on the cost of monitoring procedures. In this paper, we have looked for other justifications. We have considered the influence of unobservability of patient's severity under different assumptions on the objective functions of the regulator and the hospital. We have shown in which cases the hospital may be non responsive to the regulator objective. When there is a conflict between the regulator's preferences and the incentives constraints, the regulator must design an optimal pooling contract for pure efficiency reasons. We have shown under which conditions this non-responsiveness phenomenon occurs and results in a fixed-price contract whatever the degree of severity. When the hospital is self-interested, non-responsiveness occurs when the marginal social virtual welfare increases when severity increases. When the hospital is perfectly altruistic, it occurs when the marginal social welfare decreases when severity increases. In the first case, the fixed payment is equal to the cost incurred by the patient with the highest severity whereas it is equal to the mean value of the cost in the second one.

While a fixed price policy is usually motivated by concerns about the moral hazard problem, we have shown that it can also be motivated by concerns about the regulator's and the provider's preferences under asymmetric information on patient severity. In particular, when the regulator's benefit function does not differ from the patient benefit function, it may be increasing with severity. Due to the non-responsiveness phenomenon, the regulator must use a pooling allocation implementing the same quantity and payment for all severity of illness. Whereas an optimal separating contract solves a trade-off between rent extraction and efficiency, non-responsiveness forces the regulator to give up extracting the hospital's rent.

Taking simultaneously moral hazard and adverse selection would not change the main insights of this paper. A question remains open. Non-responsiveness is the result of a confiict between the regulator and the agent preferences. What do we know about these preferences? What is the regulator objective? Is it increasing or decreasing with the severity of treated patients? In the theoretical literature, some models are based on increasing benefit functions whereas some others are based on decreasing benefit functions. Empirical analysis should be useful to reveal the regulator preferences when choosing specific payment rules.

:

Appendix 1

The problem of the regulator is

$$\underset{qC@),yC@)\sim}{\text{Max}} \overset{Z}{} \{V\,(q(/3),/3) - (1 + \grave{A})C(q(/3),/3) - F\,(/3)=f(/3)(C_{\sim}(q(/3),/3))\}f(/3)d(/3)$$

under

$$\begin{array}{rll} q_{/3}(/3) & \dot{}q(/3)=y(/3) & (A.1) \\ y(/3) & 0 & (A.2) \end{array}$$

where $q(/3)$ is the state variable and $y(/3) = \dot{}q(/3)$ the control variable. Let us denote by $i(/3)$ the multiplier associated with (A.1). The Hamiltonian is

$$H(q, y, , /3) = \{V(q(/3), /3)-(1+\grave{A})C(q(/3), /3)-\grave{A}F(/3)/f(/3)(C_{\sim}(q(/3), /3))\}f(/3)+uy$$

From the Pontryagin principle, we have

$$\text{:}\;\; (/3) = -\frac{0\;\;H}{Oq}$$

$$= -\{V_q(q(/3), /3) - (1 + \grave{A})C_q(q(/3), /3) - \grave{A}F(/3)/f(/3)(C_{\sim_q}(q(/3), /3))\}f(/3) \quad (A.3)$$

Maximizing with respect to y(/3) with y(/3) 0 yields p(/3) ~ 0 with y(/3) = 0 ifp(/3) >0.

On an interval where y(/3) < 0, p(/3) = 0 and ˙p(/3) = 0 and we obtain q(/3) solution of

$$V_q(q(/3), /3) - (1 + \grave{A})C_q(q(/3), /3) - \grave{A}F(/3)/f(/3)(C_{\sim q}(q(/3), /3)) = 0 \qquad (A.4)$$

If q(/3) solution of (A.3) is increasing then decreasing, the monotonicity constraint is binding on [/3, e/3] and q(/3) is constant in the interval and equal to eqi As the multiplier is continuous, after integrating (A.3) between /3 and e/3, we obtain

$$Z_{\sim 3e}(V_q(eq,/3) - (1 + \grave{A})C_q(eq, /3) - \grave{A}F(/3)/f(/3)C_{\sim q}(eq, /3))f(/3)d(/3) = 0 \text{ and } eq =$$

q(e/3)

References

[1] Caillaud B., R. Guesnerie, P. Rey and J. Tirole (1988), Government Intervention in Production and Incentives Theory: A Review of Recent Contributions, Rand Journal of Economics, 19, 1-26.

[2] Chalkley M. and J.M. Malcomson (1998), Contracting for Health Services when Patient Demand does not Refiect Quality, Journal of Health Economics, 17(1), 1-19.

[3] Chalkley M. and J.M. Malcomson (2002), Cost Sharing in Health Service Provision: an Empirical Assessment of Cost Savings, Journal of Public Economics, 84, 219-249.

[4] Choné P. and C-t. A. Ma (2008), Optimal Health Care Contract under Physician Agency, Boston University Working Paper Series

[5] Cuif K., J. Hurtley, S. Mestelman, A. Muller and R. Nuscheler (2007), Public and Private Health Care Financing with Alternative Public Rationing Rules, CHEPA working paper, Mc Master University.

[6] Culyer A.J. (1989), The Normative Economics of Health Care Finance and Provision, Oxford Review of Economic Policy, 5(1), 34-58.

[7] Culyer A.J and A. Wagstaff (1993), Equity and Equality in Health and Healh Care, Journal of Health Economics, 12(4), 431-457.

[8] De Fraja G. (2000), Contracts for Health Care and Asymmetric Information, Journal of Health Economics, 19, 663-677.

[9] Ellis R.P. and Th. Mc Guire (1986), Provider Behavior under Prospective Reimbursement: Cost Sharing and Supply, Journal of Health Economics, 5, 129-152.

[10] Ellis R.P. and Th. Mc Guire (1990), Optimal Payment Systems for Health Services, Journal of Health Economics, 9, 375-396

[11] Guesnerie R. and J.J. Laffont (1984), A Complete Solution to a Class of Principal-Agent Problems with an Application to the Control of a self-Managed Firm, Journal of Public Economics, 25, 329-369

[12] Hammond P.J (1987), Altruism, in Eatwell J., M. Milgate and P. Newman (eds;), New Palgrave Dictionary of Economics, London, Mc Millan Press Ltd, 85-87.

[13] Jack W. (2005), Purchasing Health Care Services from Providers with Unknown Altruism, Journal of Health Economics, 24, 73-93.

[14] Laffont J.J. and D. Martimort (2002), The Theory of Incentives, The Principal-Agent Model, Princeton, Princeton University Press (2002).

[15] Laffont J.J. and J. Tirole (1993), A Theory of Incentives in Procurement and Regulation, Cambridge, MIT Pres

[16] Lewis T. and D. Sappington (1989), Countervailing Incentives in Agency Problems, Journal of Economic Theory, 49, 294-313.

[17] Ma C-t. A. (1994), Health Care Payment Systems: Cost and Quality Incentives, Journal of Economics and Management Strategy, 3(1), 93-112.

[18] Maggi G. and A. Rodriguez-Clare (1995), On Countervailing Incentives, Journal of Economic Theory, 66, 238-263.

[19] Morand P.H. and L.Thomas (2003), On Non-Responsiveness in Adverse Selection models with Common Value, Topics in Theoretical Economics, 3(1), article 3

[20] Mougeot M. and F. Naegelen (2005), Expenditure Cap Policy and Hospital Regulation, Journal of Health Economics,24, 55-77

[21] Mougeot M. and F. Naegelen (2009), Adverse Selection, Moral Hazard and Outlier Payment Policy, Journal of Risk and Insurance, 76(1), 177-195.

[22] Newhouse J.P. (2002), Pricing the Priceless: a Health Care Conundrum, Cambridge, MIT Press.

[23] Shleifer A. (1985), A Theory of Yardstick Competition, Rand Journal of Economics, 16, 319-327.