Accepted Manuscript

Title: The effects of hospitals' governance on optimal contracts: bargaining vs. contracting

Authors: Matteo M Galizzi, Marisa Miraldo

Please cite this article as: Galizzi, M.M., Miraldo, M., The effects of hospitals' governance on optimal contracts: bargaining vs. contracting, *Journal of Health Economics* (2010), doi:10.1016/j.jhealeco.2010.12.001

The effects of hospitals' governance on optimal contracts: bargaining vs. contracting

Matteo M Galizzi[a*], Marisa Miraldo[b]
[a]Department of Economics University of Brescia
[b]Imperial College London Business School

galizzi@eco.unibs.it

Abstract

We propose a two-stage model to study the impact of different hospitals' governance frameworks on the optimal contracts designed by third-party payers when patients' disease severity is the private information of the hospital. In the second stage, doctors and managers interact within either a bargaining or a contracting scenario. In the contracting scenario, managers offer a contract that determines the payment to doctors, and doctors decide how many patients to treat. In the bargaining scenario, doctors and managers strategically negotiate on both the payment to doctors and the number of patients to treat. We derive the equilibrium doctors' payments and number of treated patients under both scenarios. We then derive the optimal contract offered by the government to the hospital in the first stage. Results show that when the cost of capital is sufficiently low, the informational rent is lower, and the social welfare is higher, in the contracting scenario.

***Keywords*:** Strategic Bargaining; Optimal Contracts; Hospitals; Asymmetric Information

JEL classification*:* I11, I18.

1

1. Introduction

Health economists and policy makers often debate as to which organisational framework leads to better health outcomes under a controlled budget. The relevance of such an analysis stems from two crucial features of the health care system. The first is the rise of health care spending as a share of GDP. Over the period from 1980 to 2005, OECD countries experienced an average increase of more than 2 percentage points, from around 6.7% to slightly more than 9% (OECD, 2007). Because public funding is the major funding source for most OECD countries, accounting for an average share of 73% of total expenditure in 2005 (OECD, 2007),[1] steady growth in health expenditure gives rise to concerns about the sustainability of public budgets.[2] The second feature concerns the upward pressure that medical advances, demographic changes and increased expectations exert on these costs. As a consequence, controlling health expenditure growth has assumed increasing relevance in the public policy agenda of most OECD countries, leading to several financial reforms.

Financial policies have focused on, but have not been restricted to, mechanisms to control the financial flows between payers, providers and patients. In particular, on the supply side, changes in EU financing mechanisms have focused on hospital cost control (Mossialos and Le Grand, 1999; Mossialos, 2002).

Several theoretical studies have analysed the design of payment systems to induce optimal provider behaviour in terms of quality level, cost containment effort and access to treatment (see for example Chalkley and Malcomson, 1998a and 1998b; Ma, 1994; Rickman and McGuire, 1999; Ellis and McGuire, 1986; Ellis, 1998).

A common assumption in this literature is that the agents composing a hospital are considered as a whole - a sort of *black box*. However, in practice, the hospital is a product of an array of different agents with different and sometimes conflicting objectives. For example, in England, Foundation Trusts' *board of governors* consists of diverse members elected by patients, the public and staff; these members are included so that their interests and views are reflected in the organisations' governance. The *board of governors* then interacts with the *board of directors*, comprising non-executive and executive directors who balance skills and experience to meet the organisation needs (Department of Health, 2006). More generally, a diversity of organisational and governance frameworks have been implemented across hospitals in the different health systems. The common feature of these is the existence of different agents within the hospital - with potentially different objectives – that interact and make decisions on the hospital's strategic variables. Thus, the specific governance and organisation framework is likely to significantly affect not only the bargaining power of the different agents, but also, and perhaps more importantly, the overall performance of the hospital. Therefore, the design of effective contractual incentives by third-party payers or regulatory bodies must acknowledge this organisational complexity.

In this paper, we analyse the impact of different organisational scenarios within the hospital on the optimal contracts designed by third-party payers when disease severity is the private information of the hospital. The different scenarios are: i) a *contracting* scenario in which managers offer a contract that determines the payment to doctors, and doctors decide how many patients to treat; and ii) a *bargaining* scenario in which doctors and managers negotiate on both the payment to doctors and the number of patients.

The results show that different governance solutions within the hospital affect the relationship between the

---

[1]Average of all OECD countries with the exception of Korea, Mexico and the United States.
[2]During the 1980s and 1990s, we also observe a shift of costs from the public to the private sector that might reflect the financial reforms introduced to control health expenditure

government and the hospitals. Compared with the contracting scenario, when the cost of capital is sufficiently low, the *informational rent is higher* and *the social welfare is lower* in the scenario where managers and doctors bargain within the hospital. Because guidelines for the use of technology, and hence its capital requirements and costs, typically vary by disease, our model suggests that the optimal organisational solution should differ from speciality to speciality and that the contracts between the government and the hospitals should accommodate such differences.

In the next section, we discuss how our analysis is situated within the literature. In section 3, we present the main model characteristics. Sections 4 and 5 develop the bargaining and contracting scenarios within the hospital, respectively. Optimal government contracts and results are discussed in section 6, and finally, section 7 concludes the paper.

2. Background and literature review

It is fair to say that the literature on models of hospital behaviour draws on previous developments in the *theory of the firm*. A natural starting point for an analysis of a firm or any economic organisation is the recognition that institutions are formed by different agents with different scopes, interests and objectives. This idea was implicit in Adam Smith's "*Wealth of Nations*" (Smith, 1776), which describes the complexity of the division of labour within a firm. This intricate net of relations of different parties within an institution was later discussed by Coase (1937) in his seminal paper on "*The Nature of the Firm*" that characterises "*the bounds of the firm as that range of exchanges over which the market system was suppressed and resource allocation was accomplished instead by authority and direction*" (Jensen and Meckling, 1976). This work was later challenged by Alchian and Demsetz (1972), who object to the notion that authority prevails as the ruling system of firms' activities, and who emphasise the role of contractual arrangements as a vehicle for voluntary exchange. The link between this literature and the health care sector is established by Jensen and Meckling (1976), who openly recognise that hospitals are indeed "*legal fictions which serve as a nexus for a set of contracting relationships among individuals*"**.**

The health care literature builds on these models to develop *theories of the hospital*. McGuire (1985) and Crilly and Le Grand (2004), in a survey of models of hospital, point out the existence of three main theories. The first considers hospitals as a unified institution. The second set of theories, known as managerial theories, considers organisations as the product of two separate units: ownership and management. Finally, the behavioural strand of theory considers the hospital as a product of multiple groups with converging or conflicting interests and which interact within the organisation. Regarding the behavioural strand of theory, most of the literature, while acknowledging the plurality of interests within the firm, assumes that one of the parties is dominant and therefore rules the organisation according to its objectives (Lee, 1971; Jacobs, 1991). In the health economics literature, the typology and nature of these relations, however, has only been weakly explored (Conrad et al., 1996; Jegers, 2004). In particular, Lindsay and Buchanan (1970) and Pauly and Redisch (1973) consider that clinicians dominate the decision-making process within the hospital. However, Harris (1977) as well as Rosko and Broyles (1988) highlight that the key issue in the analysis of hospital behaviour is the interaction between the conflicting interests of managers and doctors. We adopt this as the core assumption of our paper. We propose a model to analyse the effects of this interaction on the contractual relationship between the government and the hospital. Indeed, the heterogeneity in objectives within the hospital poses a challenge to the design of optimal financing mechanisms by the regulator. The crucial element is how these distinctive interests shape the performance of the hospital as a whole and to what extent the latter is aligned with the regulator's objectives.

The model has two *key elements*. The *first* is that it accounts for two different actors within the hospitals, namely, managers and doctors. Concerning doctors, it is assumed that various types of specialty physicians work in the hospital, each providing health treatments to patients suffering from one specific disease. Our model characterises one of these types of physicians as the representative doctor within the hospital. Doctors and managers have different objectives and act independently within the hospital, but it is assumed that they act jointly when contracting with the third-party payer. In particular, we assume that both doctors and managers elect their representatives to sit on the hospital board, to which they fully delegate the contractual relations with the government.

The *second key element* of our model is the explicit comparison between two alternative organisational scenarios within the hospitals. In the first scenario (*contracting*), managers offer the doctors a contract that determines their payment, and doctors decide how many patients to treat. In the second scenario (*bargaining*), both managers and doctors strategically bargain over resource allocation within the hospital. This second modelling element allows us to directly compare the most salient outcomes of two polar organisational solutions implemented in western health care systems: organisations in which the responsibilities for financial management and medical care are integrated and shared among the main actors within the hospitals (*bargaining*), and, in contrast, solutions in which the latter are completely split and separate (*contracting*).

The present model contributes to the existing *health economics literature*. Our setting is close to the work by Custer et al. (1990), Boadway et al. (2004) and Miraldo (2000), but departs from their analysis in many crucial ways. Custer et al. (1990) analyse the effects of a prospective payment system on hospital production, focusing on the relationship between the hospital and its medical staff. Other studies that account for different incentives within the hospital model hospitals' internal relations by means of *take-it-or-leave* offers made by one of the parties, which amounts to assuming that only the proposer has decision power. Boadway et al. (2004), in particular, propose a model with managers and doctors as decision makers and develop a two-stage agency problem in which contracts are designed to elicit information. Their paper is extended to consider cooperative bargaining within the hospital by Miraldo (2000). Our paper is closely related to these works, and it generalises Miraldo's (2000) model by introducing strategic negotiations between doctors and managers within the hospital in the spirit of non-cooperative bargaining and by explicitly comparing the effects of different interactions within the hospital on the contracts offered by the government.

Our model is also related to a broader strand of studies within the *economic literature*. For instance, the modelling issue of two subjects acting non-cooperatively within an organisation and jointly through their representatives when facing external agents, is directly inspired by the economic literature on intra-firm bargaining and coalitions (Jensen and Meckling, 1976; Hart and Moore, 1990; Stole and Zweibel, 1996a, 1996b) and by its applications to corporate finance (Rajan and Zingales, 1998; Scharfstein and Stein, 2000; Mudambi and Navarra, 2004). Similar assumptions are also common in the literature on bargaining and coalitions (Hart and Kurz, 1983; Zhang, 1995; Vidal-Puga, 2005; Bloch and Gomes, 2006) and on cartel formation (McAfee and McMillan, 1992; Bloch and Ghosal, 2000). More generally, our model also follows the general literature on the theory of the firm. In fact, Jensen and Meckling (1976) depart from Alchian-Demsetz's model by highlighting the importance of contractual relations between the firm and other organisations (such as regulators, customers, suppliers or others). In the healthcare sector the links between the market (e.g., market structure and composition) and the institutional environment (e.g., regulatory incentives; contractual mechanisms) condition the performance of health care organisations as a whole or by conditioning their elements (Putternam, 1986). Our model thus contributes to this broader strand of literature in that it is the first attempt to explicitly analyse the effects of the interaction within hospitals on the contractual relationship with the government in the presence of asymmetric information.

4

3. The model

The main actors in our model are the government (*G*) and the hospitals (*H*).

3.1 Hospitals

We open the hospital *black box* by considering that different agents exist within the hospital that have different objectives and that interact and make decisions on the hospital's strategic variables. In particular, we assume that hospitals are composed of doctors (*D*) and managers (*M*). It is assumed that managers aim at maximising their expected financial budget, whereas doctors care about their personal surplus, and they also care about improving the health care status of their patients.[3] Our assumptions about doctors' and managers' objectives are in line with the health economics literature (Pauly and Redisch, 1973). In particular, our assumption about the managers' objective is justified because, although a considerable part of the literature has assumed that the hospital's objective is to maximise either quantity, quality or both (Brown, 1970; Long, 1964; Rice 1966; Feldstein 1968; Newhouse 1970), profit maximisation has often been assumed to be the leading objective not only in for-profit hospitals (Davis, 1971; Jacobs, 1991) but also in not-for-profit organisations. Indeed, even though publicly-funded hospitals are usually constrained on the distribution of profits, it has become more common that public hospitals can add the financial surplus to their reserves. For example, Foundation Trusts in England, despite being public organisations, are the residual claimants of their gains or losses. With boards of directors who risk their position in the case that the hospital incurs significant losses, managers in the hospitals may therefore be considered profit maximisers (Department of Health, 2002a, 2002b).

Furthermore, it is also assumed that each hospital has a board that acts entirely on the behalf of the hospital as an institution when contracting funding and technological capacity with the government. The board is composed of representatives of both managers and doctors, and the government deals with hospitals only through their boards.

Note that because managers and doctors have different objectives, the specific governance and organisation within a hospital will determine the bargaining power of the different agents on the key decisions within the hospital and, therefore, on the overall hospital performance. As a consequence, the interaction within the hospital will shape the optimal contracts designed by the government to define the financing and technology capacity allowed to each hospital.

3.2 The first stage

Our model is developed in two stages. In the *first* stage, the hospital signs a financing contract with the government. The objective of the government is to optimally trade quality of care and health care expenditures. To achieve the optimal trade-off, the government proposes contracts to the hospitals, and these are signed by the board as outlined in the previous section. The format of the offered contracts does not differ according to the internal organisation of the hospitals. Indeed, under both scenarios, the government specifies a limited *capacity*, K_i, that constrains the use of high technology treatment; a *lump sum transfer*, T_i, paid *to the managers*; and a *fee*, g_i, paid directly *to the doctors* for each patient treated.[4]

Capacity K_i is a main element of our model and can be interpreted as hospitals' capital investments. The government invests in each type of hospital a certain level of capital *K*, paying a unitary cost of capital *r*. As already discussed, it is assumed that a number of specialty physicians work within the hospital. Our model

---

[3] Paraphrasing Le Grand (2003), we assume that managers are "knaves" while doctors are "knaves and knights".
[4] Note that under our setup, the fee can be equivalently interpreted as being a fee for service and as a capitation fee.

characterises one of these physician types at a time. Within a given specialty unit, the government's physical capital will be invested in a new medical technology, such as magnetic resonance imaging (MRI) machine, a cardiac cath lab or a new ultrasound sensor. The capital invested in the technology crucially constrains the level of health care services that can be run in the specialty unit and can therefore be seen as a cap on the health care activities within the hospital in terms of the treatments that can be provided to patients. Such capacity represents a target level for the hospitals, not only because the minimisation of fixed costs implies a full exploitation of investments up to the capacity level, but also because utilisation above capacity implies congestion. Both features illustrate recent trends in health care systems: on the one hand, national systems are pushing hospitals to rationalise their services to enjoy the most efficient operating scale; on the other hand, they also aim at avoiding the effects of congestion in technology utilisation, especially in terms of increased waiting times or waiting lists.[5]

The central role of technological capacity in our model is motivated by the substantial evidence that in recent decades, innovations in healthcare technology have been among the most important (if not the most important) drivers of the growth in health care expenditures on one hand, and improved outcomes on the other (Weisbrod, 1991; Newhouse, 1992; Bradley and Kominsky, 1992; Rettig, 1994; Blomqvist and Carter, 1996; Fuchs, 1996; Chernew et al., 1998; Zweifel et al., 1999; Cutler and Huckman, 2003; Meara et al., 2004; Goldman et al, 2005). Examining the mentioned literature, healthcare technology can be widely defined as the stock of knowledge about health treatments that is incorporated in drugs, medical devices and procedures.[6] This definition implies that healthcare technology is usually incorporated in physical and capital-intensive investments that shape and constrain its availability and utilisation (Scitovski, 1985; Cromwell and Butrica, 1995; Cutler et al., 1999; Hay, 2003; Baker et al., 2003; Pammolli et al., 2005)[7].

In our model, therefore, to promote allocative efficiency, the government decides the level of investment in technological capacity, thus constraining the use of healthcare technology in hospitals. This setting aims at capturing in a stylised way some trends in the contractual arrangements between governments and hospitals. In fact, several governments in the European Union, in accordance with their actual cost containment policies, attempt to design contracts with hospitals that improve health outcomes while keeping the use of some therapies under control and avoiding an over-utilisation of expensive treatments. In spite of such supply side pressures, many crucial variables can be exclusively controlled by managers and doctors, whose objectives may not include health care cost containment or, to a lesser extent, the improvement in health outcomes.

Under a situation of perfect information, this would be a straightforward issue and the purchaser of health care services could simply define a global budget for each hospital and list the type and volume of required

---

[5] This is particular relevant for the English health system because managers can be severely punished if, for example, waiting time targets are not achieved (Department of Health, 2002a).
[6] Clearly, the level of invested capital is more crucial in industries like energy or telecommunications, which produce single, standardised goods, rather than in the health care sector, which typically provides a wide range of complex and integrated labour-intensive services. Nevertheless, the magnitude of this type of investment is not negligible within the health care sector and has been identified as one of the major drivers of health care expenditure (Weisbrod, 1991).
[7] Many reasons can be advanced to explain the strong link between healthcare technology and capital-intensive investments. For example, the broadening concept of health care needs, which is associated with an aging population, increases the profitability of technological development. Indeed, patients increasingly not only live longer, but have higher expectations regarding the scope and breadth of health care provision and an increased willingness to pay for more sophisticated treatments. The latter, which is associated with changes in lifestyle and the increasing share of chronic diseases, has pushed health systems to find more sophisticated and effective, but also more costly, ways to treat these diseases: for example, *left ventricular assist devices*, which entail a mortality improvement of 15%, cost $120,000 per implant (Goldman et al, 2005). Finally, as also pointed out by a recent special report on innovation in health care (The Economist, 2009), the ongoing process of the digitalisation of medical records in several countries (e-health) will require high capital investments in scanning and data storing technologies in the hospitals. Outside the hospital setting, there is a similar trend. The increased need for long term care due to a steadily aging population has pushed governments to find more sophisticated devices to routinely manage long-term conditions outside the expensive inpatient-care setting, thus putting the emphasis on capital-, rather than labour-focused care (e.g. tele-care).

6

treatments. However, information on patients' case-mix is the private information of hospitals.[8] Therefore, hospitals with a low disease severity have an incentive to mimic high case-mix hospitals and to over-provide intensive treatments to attain financial benefits.

We follow a growing strand of the literature (De Fraja, 2000; Chalckley and Malcomson, 2002; Beitia, 2003; Jack, 2005; Siciliani, 2006) in deriving the optimal contract paid by the purchaser of health care services to the hospital when there is asymmetric information between the former and the latter. Conceptually, our model is an application of the Laffont–Tirole model (Laffont and Tirole, 1993; Baron and Myerson, 1982), which derives optimal contracts under asymmetric information. We therefore use a Principal-Agent model, where the principal is the government, as a purchaser of health services or as a regulator, and the agent is the hospital. The government maximises a utilitarian social welfare function composed by the difference between the patients' surplus plus the hospitals' profits and the total costs of health care. The total costs of health care, in turn, is composed by the sum of the lump-sum transfers to the managers, the fee for service paid to the doctors, and the cost of capital investment in technological capacity, all weighted by the shadow cost of public funding. As in any asymmetric information frame, the government designs such contracts to minimise informational rents while guaranteeing that each hospital operates without losses and has incentives to report its true case-mix.

3.3 The second stage

In the *second* stage, managers and doctors decide on resource allocation within the hospital and the number of patients treated with different technologies according to the decision-making process specific to the scenario.

In fact, the two organisational scenarios differ in the way decisions are made *within* the hospital. In particular, in the *contracting* scenario, the managers decide the hospital fees per treated patient (*h*) to be paid to doctors, while doctors decide the number (*n*) of patients to be treated in the hospital.[9]

In the *bargaining* scenario managers and doctors strategically bargain over both the hospital fees (*h*) and the number (*n*) of treated patients.

3.4 Patients and treatments

As already discussed, it is assumed that there are several specialty physicians within each hospital, each providing treatments to patients suffering from one specific disease. Our model focuses on the decision of one of these specialty physicians as a representative doctor within each hospital. We assume that hospitals differ only in the disease severity of their patients. Without loss of generality, we consider two types of hospitals differing in patients' case-mix, which we will denote as *type i* with $i = 1, 2$. All hospitals serve the same total

---

[8] Note that the assumption of perfect information is common to Miraldo (2000) but not to Boadway et al. (2004) who considered a two-stage asymmetric information between the hospital and the government as well as within the hospital. We believe that it is plausible to assume perfect information within the hospital because we are interested in analysing how different governance scenarios within the hospital affect hospital financing. The asymmetry of information would change the contractual arrangements within the hospital (both in the contract and in the bargaining scenarios) towards giving more power to the physicians. However, crucially, this effect would pull towards the same direction in both scenarios. Also, there are features of the healthcare sector that justify our choice and are related to the hypothesis that if one can *a priori* think of asymmetric info between managers and doctors, its level may be of less significance after examining hospital staff composition and duties because managers may: work closer to the clinical staff; have access to medical records and can therefore verify severity; or have worked as clinicians in the past. All these features would allow managers to ex-post verify patients' severity. Finally, in the bargaining scenario, the introduction of asymmetric information makes the overall analysis far more complicated because it requires a different solution concept in terms of Bayesian Nash Equilibria, which can result in extremely complex outcomes because of the simultaneous interaction between beliefs formation and dynamic strategies. Indeed, to the best of our knowledge, the literature on bargaining theory (see Osborne and Rubinstein, 1990; Muthoo, 1999) has not yet been able to work out a general analytical solution for multi-issue, infinite-horizon bargaining games under asymmetric information. To the best of our knowledge, the only model that analytically solves a multi-issues negotiation within asymmetric information is the one designed by Wang (1998) for the special case of a wage-quality linear relation.

[9] Note that both managers and doctors decide in this second stage and therefore play a simultaneous moves game. If the decisions under the contracting scenario were modelled as a sequential game but were modelled as a simultaneous moves game under the bargaining scenario, then it would not be possible to disentangle the effect of different organisational scenarios from that of the different timing in decision-making.

number of patients. Patients of hospital of type i differ in the severity of illness s_i, which is randomly distributed over a unitary length interval according to a uniform distribution with mean α_i,

$$s_i \in [\alpha_i - \frac{1}{2}, \alpha_i + \frac{1}{2}]$$

with $\alpha_i \geq 1/2$. Without loss of generality, the *type 2* hospital is assumed to have a higher case-mix, that is, $\alpha_2 > \alpha_1$. It is also assumed that the mass of patients with a particular realisation s_i is equal to one.

Information on patients' disease severity is shared perfectly within the hospital but is not known by the government: the government has no information concerning the hospital's case mix of patients in terms of illness severity. The government, however, knows the probability p_i of a hospital being of type i, with $p_1 = 1 - p_2$.

Patients seek care in their local hospital and are offered one of two treatments for the same disease: *high* technology treatment *H* or *low* technology treatment *L*. Alternatively but equivalently, we can think of *H* as *treatment* and *L* as *no treatment*.[10] The capacity allowed by the government can be seen as the capital invested in the technology for treatment *H*.

The benefit of using *high technology* therapy in the treatment of a patient increases with the patient's disease severity. In contrast, the health improvement when treatment occurs with therapy *L* is constant and does not depend on the patient's illness severity.[11] The marginal benefit from each treatment is given by:

$$q_H(s) = as \quad q_L(s) = b \quad with \quad 0 \leq b < a \leq 1$$

Note that while, from a patient's perspective, high technology therapy is a more effective treatment as long as $s > b/a$, from a cost containment perspective, its unlimited use may not always be efficient. In fact - as will be clear once the agents' payoff functions are formally introduced - when direct and indirect costs are accounted for, and from a social welfare perspective, only high illness-severity patients should be treated with high technology therapy *H*. That is, the use of high technology therapy is more beneficial when hospitals are of *type 2*.

For the sake of tractability we will conduct our analysis assuming that *b=0* and *a =1*. This should not be seen as a restrictive assumption because within the *treatment* versus *no treatment* taxonomy, it merely implies that patients with a positive disease severity must be treated while the others should be sent home, which is indeed a realistic simplification of hospitals' decisions under financial and capacity constraints. An alternative explanation could be that doctors treat both the high- and low-disease-severity patients, but while high-severity patients receive further treatment for the disease, the low severity-patients receive only a basic diagnostic exam that determines that they are not high-severity patients. [12] However, it is important to emphasise that our main results remain qualitatively unaffected by instead working with generic values for *a* and *b*: the only difference is

---

[10] Even though our treatment setting allows us to account for the multi-dimensionality of the hospital in a very stylised manner, it does ignore important issues such as cross-subsidisation between departments within the hospital or allocation of resources across medical teams that are out of the scope of this paper and do not affect our results qualitatively. However, our model does apply to settings within which the manager negotiates with each service/clinical team at a particular time and under an overall budget constraint. In this case, as mentioned, the doctors' problem would stand for a representative agent problem.

[11] Note that these only refer to the marginal benefit. Even though patients with higher severity will benefit relatively more from treatment, as will be clearer in Section 3.5, from the doctor's perspective, it would be more costly to treat a group of patients with a higher average severity.

[12] For instance, consider all patients who seek health care after suffering severe chest pain. On arrival at the hospital, all patients are physically examined, after which only the patients who have an acute myocardial infarction are actually treated within the hospital.

8

that computed explicit solutions and comparisons would be far more cumbersome otherwise.[13]

Our two-treatment setting can be illustrated with some examples. Consider, for instance, patients suffering from endometriosis. Endometriosis, a condition that affects women, occurs when endometrial cells are deposited in areas outside the womb, causing pain and possibly forming solid lumps that in some severe cases can affect organ function or cause infertility.

Different treatments are available to manage endometriosis in a manner that reduces pain and avoids the formation of benign tumours. One possibility is to resort to hormonal therapy and painkillers that control the growth of the endometrial tissue by suppressing or reducing the natural cycle. A second possibility is surgery to remove endometrial cells around the body. There are two possible types of surgery: laparoscopic and open surgery. Laparoscopic surgery is a less invasive treatment performed through small incisions and is usually performed in day-case settings; open surgery is a more invasive treatment performed through a single large incision. Both treatments are effective in reducing pain. However, hormonal treatment is more cost effective for low-severity patients: in fact, the cost of prescribing a contraceptive or another hormonal treatment is significantly lower than the cost of surgery when endometriosis is not too severe (Bulun, 2009).

Another example concerns patients suffering from degenerative arthritis of the knee. This condition is caused by the breakdown and eventual loss of the cartilage of one or more joints. Depending on the severity of the disease and patient characteristics, there are two treatment paths for this condition. The first is non-invasive and includes rehabilitation and physical therapy, changes in lifestyle, or even the use of chondro-protective oral medication. The second treatment path involves surgery. There is a series of options within the surgery path that include arthroscopic debridement, tibial or femoral osteotomy, and uni-compartmental or total knee arthroplasty. The first treatment path is less costly and is therefore the option usually offered to low-severity patients (Willis-Owen et al, 2009).

Depending on the scenario of the model, treatments will be assigned to patients either by the negotiation between managers and doctors (in the *bargaining* scenario) or by doctors only (in the *contract* scenario). In general, patients with an illness severity above a certain level $\hat{s}_i$ will be treated with therapy H while the remaining patients will receive treatment L. Therefore, the number of patients treated with high-technology therapy n_i can be obtained by,

$$n_i \equiv \alpha_i + \frac{1}{2} - \hat{s}_i$$

which is clearly increasing in the average severity of disease of the population. The total health status improvement $Q_i(n_i, \alpha_i)$ is then given by:

$$Q_i(n_i; \alpha_i) \equiv \int_{\alpha_i - \frac{1}{2}}^{\alpha_i + \frac{1}{2} - n_i} q_L(s)ds + \int_{\alpha_i + \frac{1}{2} - n}^{\alpha_i + \frac{1}{2}} q_H(s)ds = \left(\alpha_i + \frac{1}{2}\right)n_i - \frac{1}{2}n_i^2$$

Differentiating, one can see how the total health status improvement is affected by n_i and α_i:

$$Q_{n_i} = \left(\alpha_i + \frac{1}{2}\right) - n_i, \quad Q_{\alpha_i} = n_i > 0, \quad Q_{n_i \alpha_i} = 1 > 0$$

---

[13] Of course, such a formulation ignores induced demand issues, which are beyond the scope of this paper. Note that the main purpose of defining two treatments is simply to represent medical cases for which there exists a threshold of severity beyond which patients have a higher benefit from receiving one treatment over the other. This is contemplated within our model.

Namely, the health status improvement increases with the average illness severity and the number of patients treated with therapy *H* as long as $\left(\alpha_i + \frac{1}{2}\right) > n_i$, which, by construction, is always verified as $n_i < \alpha_i + \frac{1}{2}$. Finally, the marginal improvement in the health status of treating an extra patient with high-technology therapy is higher for hospitals serving more severely ill populations.

3.5 Doctors' problem

In both organisational scenarios, doctors' utility function is defined by,

$$\Pi_{Di} = Q_i(n_i, \alpha_i) + \left(h_i + g_i\right)n_i - \frac{n_i^2}{2} \qquad (1)$$

In other words, within a hospital of type i, doctors directly benefit from the improvement in patients' health status $Q_i(n_i, \alpha_i)$ and from the fees paid by the government (g_i) and by the managers (h_i) for each patient treated with therapy *H*. This doctor reimbursement specification matches the way doctors are paid in Belgium (WHO, 2000)[14] and is similar to the fee for service system by which specialist physicians are paid in several OECD countries. Specifically, in Austria, France, Mexico, and New Zealand, specialists employed in private hospitals are mostly paid by fee for service. In addition, in countries where specialists are usually salaried, such as Australia and England, specialist physicians are paid a fee for service for treating private patients in public and private hospitals (Simoens and Giuffrida 2004).

Because high technology treatment requires more time and attention, it is also assumed that, when treating patients with therapy *H*, doctors should exert a higher effort than is required under the low technology treatment *L*. Therefore, doctors also experience a disutility term $n_i^2 / 2$, which is quadratic in the number of patients treated with therapy *H*.

If, as in our *contracting* scenario, doctors are allowed to decide the number of treated patients, they clearly set n_i at a level that maximises their utility, trading earned fees and a marginal increase in patients' health status for cost of effort. The optimal number of patients treated with therapy *H* in the *contracting* scenario is increasing in the average illness severity of the patients' population and in the fees and is given by:

$$n_i^* = \frac{\left(\alpha_i + \frac{1}{2}\right) + h_i + g_i}{2}$$

3.6 Managers' problem

Managers maximise surplus that consists of a budget from the government that is paid to managers via a lump-sum transfer T_i. From this budget, managers pay doctors a fee for service h_i for any patient treated with therapy *H*. Finally, we assume that managers derive disutility from a non-efficient use of the hospital's capacity, $\left(n_i - K_i\right)^2 / 2$. Therefore, managers' surplus is given by:

---

[14] Our setting is closely inspired by the Belgian health care system. In Belgium, non-medical hospital activity is funded via a fixed prospective budget system based on per diem and patient day quota rates, while medical services are funded by a fee for service system. Within the hospital, managers, who hold the overall hospital budget, pay medical services through fees paid directly to the doctors (WHO, 2000).

10

$$\Pi_{Mi} = T_i - \frac{(n_i - K_i)^2}{2} - n_i h_i \qquad (2)$$

With the disutility term $(n_i - K_i)^2 / 2$, we aim at capturing, in a simple way, management concerns related to *capacity* and *congestion* issues. In particular, we consider that there exists a technical capacity level K_i for the therapeutic treatment *H*. Capacity K_i measures the maximum number of patients that can be effectively treated without further costs. Additionally, it refers to the existence of some fixed costs necessary for the implementation of therapy *H*, such as the investment in a medical technology. Therefore, it is assumed that whenever the number of patients treated with therapy *H* is different from the technical capacity K_i, managers support an extra (quadratic) cost. In fact, the idea is that if the number of treated patients is lower than K_i, the costs from the medical technology are not optimally recovered. In contrast, if n_i is above the capacity, congestion arises in the supply of therapeutic treatment in the forms of both direct (e.g., extra hours worked by the staff) and indirect costs (e.g., an undesirable increase in queues and waiting times).

Clearly, if the managers were called to choose the number of treated patients, they would set it at a level to maximise their own surplus, $n_i^* = K_i - h_i$. However, as discussed above, the decision of how many patients to treat with therapy *H* is fully in the doctors' hands within the *contracting* scenario, while it is set through negotiations with the doctors in the *bargaining* scenario.

The disutility term $(n_i - K_i)^2 / 2$ exhibits a further intuitive interpretation in terms of the impact of capacity in the managers' surplus. In fact, $\partial \Pi_{Mi} / \partial K_i = n_i - K_i$ is positive for $n_i > K_i$ and negative otherwise: if the number of treated patients exceeds capacity, managers' surplus increases with K_i, and it decreases when $n_i < K_i$. Indeed, if the hospital is already operating at an activity level above K_i, expanding technological capacity will clearly be beneficial because it reduces congestion. However, if the hospital is operating below capacity, then extra capacity will simply increase fixed costs with no counterpart of benefits. An example of this is a hospital investing in a second MRI machine that requires additional equipment, staff and maintenance even though the first MRI is still used far below its potential.

3.7 Government' problem

The utilitarian government chooses $[T_i, g_i, K_i]$ to maximise the difference between the sum of consumer surplus [15]$CS = Q_i(n_i^*, \alpha_i)$ plus the hospital profits $\Pi_{Mi} + \Pi_{Di}$, and the total costs of health care, $-(1+\lambda)(rK_i + T_i + g_i n_i^*)$, subject to the hospitals' participation constraints, *IR_i*, by which $\Pi_{Mi} + \Pi_{Di} \geq 0$. In fact, when formulating its optimal contract, the government, within the total costs, accounts for the lump-sum transfers T_i, the total fees for the service provided $g_i n_i$, the cost of capital r borne for any unit of capital invested in technological capacity K_i and the direct and indirect distortions caused by the tax revenue raised to cover health expenditure: $\lambda \in [0,1]$ represents the shadow cost of public funds and captures the deadweight loss per unit of tax revenue. Hence, for each hospital $i = 1, 2$, the government problem is characterised by:

---

[15] Note that, in this case, because treatment is free at the point of use, the consumer's surplus coincides with the improvement in patients' health $Q_i(n_i^*, \alpha_i)$

$$\underset{K_i, g_i, T_i}{Max} \quad W = CS\!\left(n_i^*; \alpha_i\right) + \Pi_{Mi} + \Pi_{Di} - (1+\lambda)\!\left(rK_i + T_i + g_i n_i^*\right)$$

$$s.t. \quad \{IR_i\}: \quad T_i - \frac{\left(n_i^*\right)^2}{2} + Q\!\left(n_i^*; \alpha_i\right) - \frac{\left(n_i^* - K_i\right)^2}{2} + g_i n_i^* \geq 0$$

(3)

Under asymmetric information, the government's problem also incorporates some additional constraints that should be satisfied for the proposed contracts to be incentive-compatible (see Section 4.2.2 for a detailed discussion).

One final comment is appropriate before proceeding with the analysis. Note that when doctors and managers are considered as separate agents, the government shares with doctors the aim of providing an effective health status improvement to the most severe patients. However, the costs of provision taken into account by the government differ from those considered by the doctors, and their interests directly conflict in terms of h_i. In addition, the government shares with managers the objective to allocate patients to treatments in a cost-effective manner to control the fees and the overall costs of health care provision. However, their interests also conflict with respect not only to the lump sum transfer T_i, but also to the capacity level K_i for the high technology treatment.

We will now solve the model under both scenarios, starting from the decisions within the hospitals in the second stage and then moving back to the contract offered by the government to the hospitals in the first stage.

4. Bargaining scenario

Under the *bargaining* scenario, the second stage consists of a bargaining process between doctors and managers. The government allocation $\left(T_i, g_i, K_i\right)$ is taken as a given.

4.1 Strategic negotiations between managers and doctors

We solve the second stage of the game using a non-cooperative approach to negotiations among doctors and managers. In particular, managers and doctors bargain *simultaneously* on both the number of patients to be treated n_i and the fee for service h_i. Their dynamic strategic negotiations are modelled as a multi-issue bilateral bargaining with a random order of proposer following Osborne and Rubinstein (1990) and Muthoo (1999).

In fact, at any point in time t, either managers or doctors can be randomly selected respectively with probabilities $(1-\beta)$ and β to propose offers to the other party. An offer is a pair $\left(h_i, n_i\right)$ with $h_i \geq 0$ and $n_i \geq 0$. The other party can accept or reject that offer. If that party accepts, then each party *L=M,D* takes its payoff $\Pi_i[L]\!\left(h_i, n_i\right)$ corresponding to that offer, with $\Pi_i[L]\!\left(h_i, n_i\right)$ given by (1) and (2).

If, instead, the responding party rejects the offer, both managers and doctors enter a further stage of negotiations. Bargaining effort and trade delays are costly and managers and doctors share the same inter-temporal discount factor $\delta \in (0,1)$, with $\delta \to 1$ describing the limit case of the absence of impatience frictions.

At the beginning of each further bargaining stage, a new random selection of the party entitled to make proposal

12

occurs. The game is repeated until managers and doctors agree on a pair (h_i, n_i). If they reach an agreement on (h_i, n_i) at any time t, the discounted final payoff will be $\delta^{t-1}\Pi_i[L](h_i, n_i)$. If they perpetually disagree, their payoffs will be zero.

The described negotiation game is an infinite horizon dynamic game of complete information: in fact, players' payoff functions are common knowledge and at each move in the game, the players know the full history of the play thus far. Therefore, in the analysis that follows, we will solve the game for its *subgame perfect Nash equilibria*. More precisely, we will look for those players' strategies, describing a complete plan of proposals in the price-offer phases and of decisions of either acceptance or rejection in the response phases, that generate a Nash equilibrium in the immediately subsequent price-response phases and that constitute a Nash equilibrium in every subgame.

Because this is an infinite horizon game, one cannot use the backward induction method to solve for the *subgame-perfect Nash equilibria*, so any equilibrium solution of the present game would typically involve a high overall complexity. However, following a standard solution in strategic bargaining models (Osborne and Rubinstein, 1990; Fudenberg and Tirole, 1991; Muthoo, 1999), we will only focus on the *subgame-perfect Nash equilibria* in *pure* and *stationary* strategies *(PSSPNE)* satisfying *no delay* in the trade. Therefore, it is assumed that any trader always proposes the same price at every equivalent node where she has to make an offer, and that she always behaves in the same way whenever facing identical proposals if responding to an offer. Moreover, the property of no delay guarantees that whenever a player has to make an offer, her equilibrium offer is accepted by the other player.

The following proposition summarises the results.

Proposition 1: The bargaining game between doctors and managers shows a unique equilibrium in which agreement is reached at the first round of negotiations. The unique equilibrium entails the pair of strategies $(n_i^*, h_i^*[M])$ if managers are selected to make offers at the first stage, and the alternative pair of strategies $(n_i^*, h_i^*[D])$ if doctors are selected, where

$$n_i^* = \frac{(\alpha_i + \frac{1}{2}) + g_i + K_i}{3} \qquad (4)$$

and

$$h_i^*[M](\delta, \beta, \alpha_i; K_i, T_i, g_i) \neq h_i^*[D](\delta, \beta, \alpha_i; K_i, T_i, g_i)$$

with the exact expressions being reported in Appendix A.
Proof: In Appendix A.

It is interesting to notice some salient properties of this equilibrium. First, in the unique *PSSPN* equilibrium of the multi-issue bargaining game, the equilibrium agreed *number of patients* n_i^* maximises the joint surplus of managers and doctors $T_i - \frac{(n_i^*)^2}{2} - \frac{(n_i^* - K_i)^2}{2} + Q(n_i^*) + g_i n_i^*$. This implies that the bilateral negotiation between doctors and managers is the Pareto efficient in the equilibrium number of treated patients.

Second, the equilibrium *fee for service* is $h_i^*[D]$ or $h_i^*[M]$, respectively, if the doctors or the managers are selected to make offer at the first round. Equilibrium offers for the fee are therefore different than for the number of treated patients. However, notice that, as $\delta \to 1$, that is, if the game approaches the limit case of perfectly

patient players, both proposals for the fee for service converge to the same equilibrium value:[16]

$$h_i^* [M]_{\delta \to 1} \equiv h_i^* [D]_{\delta \to 1} \to \overline{h_i^*} (\delta, \beta, \alpha_i; K_i, T_i, g_i)$$

This is clearly due to the fact that when frictions due to impatience are negligible, the advantage experienced by the first player selected to make offers disappears, and the corresponding asymmetries in the proposed equilibrium fee for service also vanish.

Notice, moreover, that the equilibrium level of the fee for service $\overline{h_i^*}(.)$ corresponds to the one found by Miraldo (2000). This is due to the property by which the equilibria of a non-cooperative bargaining game approximate the cooperative Nash bargaining solution as inter-temporal frictions disappear and $\delta \to 1$ (Muthoo, 1999). Because such an expression for $\overline{h_i^*}$ fully corresponds to the Nash solution and can be directly compared with the equilibrium fee for service under a *contract* scenario (in which impatience does not play any role), hereafter we refer to $\overline{h_i^*}$ as the unique *equilibrium fee* for a service emerging under the *bargaining* scenario.

The equilibrium of our non-cooperative negotiations game also has an intuitive interpretation: when bargaining within the hospital, doctors and managers agree to set the number of treated patients at a level that maximises hospital profits and then use the fee for service system as an instrument to divide the generated profits.

4.2 Government problem

4.2.1 Full information
We now characterise the menu of contracts proposed to hospitals by the government. However, before doing so, it is helpful to provide, as a benchmark, the so-called *full information* contracts. These contracts would be chosen by the government if it was aware of each hospital's average illness severity α_i.

In such a case, for a given α_i, for each hospital $i = 1, 2$, the government's problem is characterised by (3).

Because public funds are costly, $\partial W / \partial T_i < 0$ and, consequently, T_i will be chosen such that the participation constraints (*IR$_i$*) are just satisfied. Substituting T_i into *W* and maximising with respect to $\{K_i, g_i\}$, the full information (*first-best*) contracts must satisfy the following first-order conditions:

$$\{K_i\} : \quad \frac{\partial W}{\partial K_i} \equiv \left[\frac{\partial Q_i}{\partial K_i} + \frac{\partial J\Pi_i}{\partial K_i} \right] - (1+\lambda) \left[r + \frac{\partial T_i}{\partial K_i} + g_i \frac{\partial n_i^*}{\partial K_i} \right] = 0 \qquad (5)$$

$$\{g_i\} : \quad \frac{\partial W}{\partial g_i} \equiv \left[\frac{\partial Q_i}{\partial g_i} + \frac{\partial J\Pi_i}{\partial g_i} \right] - (1+\lambda) \left[\frac{\partial T_i}{\partial g_i} + g_i \frac{\partial n_i^*}{\partial g_i} + n_i^* \right] = 0$$

where $J\Pi_i = \Pi_{Mi} + \Pi_{Di}$ is the total hospital surplus and n_i^* is the equilibrium number of treated patients, which is an outcome of the negotiations between doctors and managers (as in Proposition 1). Note that in the optimal contract, K_i and g_i will be chosen such that the social marginal benefits (respectively $\partial Q_i / \partial K_i + \partial J\Pi_i / \partial K_i$ and $\partial Q_i / \partial g_i + \partial J\Pi_i / \partial g_i$) equal the social marginal costs (respectively

---

[16] The exact expression for $\overline{h_i^*} (\delta, \beta, \alpha_i; K_i, T_i, g_i)$ is in Appendix A.

14

$r + \partial T_i / \partial K_i + g_i \partial n_i^* / \partial K_i$ and $\partial T_i / \partial g_i + g_i \partial n_i^* / \partial g_i + n_i^*$) weighted by the shadow cost of public funds. Hence, in a full-information contract, the government chooses g_i and K_i to optimally trade health outcomes and health care costs. Namely, the third-party payer will weigh the improvement in patients' health status against the costs of treatment, ensuring that every population is served by one provider. When weighing the financial burden in particular, the government takes into account that g_i and K_i directly affect the equilibrium number of treated patients n_i^* as negotiated within the hospital at the last stage of the game. Moreover, the government also explicitly considers the cost of capital r and the indirect effects that both g_i and K_i exert on the levels of the lump-sum transfers T_i necessary to induce hospitals' participation.

It is worthwhile to notice, moreover, that g_i and K_i affect, in a positive and symmetric way, the equilibrium number of treated patients and are, therefore, perfect substitute instruments for the government in this respect. In fact, $\partial n_i^* / \partial g_i = \partial n_i^* / \partial K_i = 1/3$. Such a symmetric and positive impact also characterises the effects of g_i and K_i on the improvement in patients' health status $Q_i(\alpha_i, n_i^*)$. In fact, $\partial Q_i / \partial g_i = \partial Q_i / \partial K_i > 0$.

Computing explicit derivatives with respect to K_i and g_i and substituting in the above first-order conditions, the policy instruments in the optimal contracts are as follows:[17]

$$g_i^{b,pi} = \frac{\alpha_i + 1/2 + r}{2\lambda + 3} \qquad K_i^{b,pi} = \frac{(\alpha_i + 1/2)(2 + \lambda)}{2\lambda + 3} - \frac{r(4 + 3\lambda)}{2\lambda + 3}$$

The full information fee per patient is increasing with the case-mix and decreasing with the shadow cost of the public funds. The capacity offered by a full information contract is increasing with the case-mix and decreasing with the cost of capital. As a consequence, the equilibrium number of treated patients agreed upon in the subsequent stage is:

$$n_i^{*b,pi} = \frac{(\alpha_i + 1/2)(2 + \lambda)}{2\lambda + 3} - \frac{r(1 + \lambda)}{2\lambda + 3} \qquad (6)$$

It can been noted that, in the full information benchmark, the number of patients treated by the hospital only differs from the contractual capacity by a positive (but small) term related to the cost of capital.[18] Intuitively, this is due to the fact that the hospital's interests with respect to the number of treated patients are not perfectly aligned with the government's goals, which also encompass the cost of capacity *r*.
Finally, by comparing the number of treated patients across hospitals of different types,

$$n_2^{*b,pi} - n_1^{*b,pi} = (\alpha_2 - \alpha_1)\frac{2 + \lambda}{2\lambda + 3}$$

it is shown that hospitals with a higher case-mix treat more patients and that this difference is directly increasing with the distance in the average illness severity between the two populations.

4.2.2 Asymmetric information
Of course, the full information policy cannot be implemented when the government is not aware of the hospital's type. Indeed, as in a standard *asymmetric information* problem, the hospital with the less-severe casemix has an incentive to claim to be a hospital of *type* 2 to benefit from the more generous bundle $[T_2, g_2, K_2]$. In other

---

[17] Note that the subscript b, pi has been used to identify the perfect information benchmark case within the bargaining scenario.
[18] In fact, $(1 + \lambda)/(4 + 3\lambda) < 1$ is verified as $\forall \ \lambda > 0$.

words, the full information contracts are not incentive-compatible. Therefore, the government should design its menu of contracts to induce the hospitals to truthfully reveal their true type. This is equivalent to ensuring that each hospital is better off by choosing a specific contract $[T_i, g_i, K_i]$ designed for its type rather than mimicking the other type and selecting its contract. This is achieved when in the government's problem (3), the additional incentive-compatible constraints (IC_i) are satisfied

$$\{IC_i\}: \quad T_i - \frac{\left(n_i^*\right)^2}{2} + Q_i\left(n_i^*; \alpha_i\right) - \frac{\left(n_i^* - K_i\right)^2}{2} + g_i n_i \geq T_{-i} - \frac{\left(\hat{n}_{-i}^*\right)^2}{2} + \hat{Q}_i\left(\hat{n}_{-i}^*; \alpha_i\right) - \frac{\left(\hat{n}_{-i}^* - K_{-i}\right)^2}{2} + g_{-i}\hat{n}_{-i}^*$$

where T_{-i}, K_{-i} and g_{-i} stand for the elements of the contract intended for hospitals of the type opposite to $i = 1, 2$, while

$$\hat{n}_{-i}^* = \frac{\left(\alpha_i + \frac{1}{2}\right) + g_{-i} + K_{-i}}{3}, \quad \widehat{Q}_{-i} = \left(\alpha_i + \frac{1}{2}\right)\hat{n}_{-i}^* - \frac{1}{2}\left(\hat{n}_{-i}^*\right)^2$$

denote the number of patients treated and the corresponding health improvement by a *type* $i = 1, 2$ hospital mimicking one of the opposite *type* $-i$.

Before proceeding with the optimal contract, it is useful to investigate the properties of the hospital iso-profit curves. The hospitals' indifference curves are monotonic in average severity and therefore the *Single Crossing* conditions are satisfied (see calculations in Appendix B).

From the *Single Crossing* conditions, it follows that the IC_2 constraint is always automatically satisfied as long as constraints IR_2 and IC_1 are verified. This standard result is justified because only hospitals of *type 1* have incentives to mimic hospitals of the other type to be allowed higher budgets.

Moreover, given that public funds are costly, constraint IC_1 will bind at the optimum: the informational rent to be paid to hospitals of *type 1* will be minimised, leaving such hospitals just indifferent as to whether to reveal their true type or to mimic hospitals of *type 2*. That is, the government will choose T_1 such that the following equality would be satisfied:

$$\{IC_1\}: \quad T_1 + Q_1\left(n_1^*; \alpha_1\right) + g_1 n_1^* - \frac{\left(n_1^*\right)^2}{2} - \frac{\left(n_1^* - K_1\right)^2}{2} = R\left(K_2, g_2; \alpha_1\right)$$

where,

$$R\left(K_2, g_2; \alpha_1\right) = \left\{T_2 - \frac{\left(\hat{n}_2^*\right)^2}{2} + \widehat{Q}_2\left(\hat{n}_2^*; \alpha_1\right) - \frac{\left(\hat{n}_2^* - K_2\right)^2}{2} + g_2\hat{n}_2^*\right\} \qquad (7)$$

stands for the *informational rent* accruing to hospitals of *type 1*. Clearly, the lump sum remuneration of managers T_1 will be higher than in a full information case to induce hospitals to choose the contract intended for their type. Because this rent is positive, the participation constraint for a hospital of *type 1* (IR_1) is also automatically verified.

In contrast, in the contract intended for hospitals of *type 2* there will be no informational rent: T_2 will be chosen such that the hospital participation constraint (IR_2) just binds.

16

To investigate the role of K_i, g_i and T_i in the government contract, it is useful to take a closer look at the *informational rent*. Substituting T_2 in $R^b\left(K_2, g_2; \alpha_1\right)$, comparative statics analysis shows that the impact of both K_2 and g_2 on the above informational rent is negative and equal in magnitude. In fact, it can be seen that $\partial R^b / \partial K_2 = \partial R^b / \partial g_2 = -(\alpha_2 - \alpha_1)/3 < 0$. Therefore, the capacity and the fee for service system are perfect substitutes as policy instruments targeted at the minimisation of informational rent.

However, when looking at the design of the menu of contracts, we need to keep in mind that the government objective of the provision of cost-effective health care involves not only the minimisation of informational rent, but also the maximisation of patients' total health improvement and the control of health treatment costs. Therefore, the policy instruments in the optimal contracts will reflect the trade-off between all of these elements.

Maximising *W*, with respect to K_i and g_i, it can be seen that the optimal contract for *type 1* hospitals is described by the following first-order conditions:

$$\{K_1\}: \quad \frac{\partial W}{\partial K_1} \equiv p_1\left[\frac{\partial Q_1}{\partial n_1^*}\frac{\partial n_1^*}{\partial K_1} + \frac{\partial J\Pi_1}{\partial K_1}\right] - (1+\lambda)p_1\left[r + n_1^*\frac{\partial n_1^*}{\partial K_1} + \left(K_1 - n_1^*\right)\left[1 - \frac{\partial n_1^*}{\partial K_1}\right] - \frac{\partial Q_1}{\partial n_1^*}\frac{\partial n_1^*}{\partial K_1}\right] = 0$$

$$\{g_2\}: \quad \frac{\partial W}{\partial g_1} \equiv p_1\left[\frac{\partial Q_1}{\partial n_1^*}\frac{\partial n_1^*}{\partial g_1} + \frac{\partial J\Pi_1}{\partial g_1}\right] - (1+\lambda)p_1\left[n_1^*\frac{\partial n_1^*}{\partial g_1} + \left(K_1 - n_1^*\right)\frac{\partial n_1^*}{\partial g_1} - \frac{\partial Q_1}{\partial n_1^*}\frac{\partial n_1^*}{\partial g_1}\right] = 0$$

where $\partial J\Pi_1 / \partial K_1$ and $\partial J\Pi_1 / \partial g_1$ capture the impact of K_1 and g_1 on the overall hospital surplus while $\left(\partial Q_1 / \partial n_1^*\right)\left(\partial n_1^* / \partial K_1\right)$ and $\left(\partial Q_1 / \partial n_1^*\right)\left(\partial n_1^* / \partial g_1\right)$ stand for the effects of the policy instruments on the improvement in patients' health status, an objective that the government is able to control only indirectly through its influence on the equilibrium number of patients n_1^* negotiated within the hospital. The negative term in the second brackets multiplied by $(1+\lambda)$ represents the overall costs associated with the effect of the policy instruments on the equilibrium number of treated patients: in particular, concerning capacity, the costs are net of the beneficial effect on the increase in patients' health status $\left(\partial Q_1 / \partial n_1^*\right)\left(\partial n_1^* / \partial K_1\right)$ and include the cost of capacity *r*, the induced cost $n_1^*\left(\partial n_1^* / \partial g_1\right)$ of treated patients in terms of a higher fee for services paid by the government, and the indirect cost $\left(K_1 - n_1^*\right)\left[1 - \left(\partial n_1^* / \partial K_1\right)\right]$ of the sub-optimal utilisation of the medical technology, arising any time the number of treated patients diverges from the allowed capacity and as long as their relative rates of adjustment differ. Concerning the fee for service, because there is no related cost of capacity, costs, net of the beneficial health status improvement $\left(\partial Q_1 / \partial n_1^*\right)\left(\partial n_1^* / \partial K_1\right)$, only include the induced cost $n_1^*\partial n_1^* / \partial g_1$ of treated patients in terms of higher fees paid by the government and the indirect cost $\left(n_1^* - K_1\right)\left[\partial n_1^* / \partial g_1\right]$ of congestion.

The first-order conditions for the contract intended for *type 2* hospitals are analogous to those for *type 1* hospitals, with the addition of some supplementary terms representing the impact of g_2 and K_2 on the informational rent $R^b\left(K_2, g_2; \alpha_1\right)$:

$$\{K_2\}: \quad \frac{\partial W}{\partial K_2} \equiv p_2\left[\frac{\partial Q_2}{\partial n_2^*}\frac{\partial n_2^*}{\partial K_2} + \frac{\partial J\Pi_2}{\partial K_2}\right] - (1+\lambda)p_2\left[r + n_2^*\frac{\partial n_2^*}{\partial K_2} + \left(K_2 - n_2^*\right)\left[1 - \frac{\partial n_2^*}{\partial K_2}\right] - \frac{\partial Q_2}{\partial n_2^*}\frac{\partial n_2^*}{\partial K_2}\right] -$$

$$- (1+\lambda)p_2\frac{\partial R^b\left(K_2, g_2; \alpha_1\right)}{\partial K_2} = 0$$

$$\{g_2\}: \quad \frac{\partial W}{\partial g_2} \equiv p_2\left[\frac{\partial Q_2}{\partial n_2^*}\frac{\partial n_2^*}{\partial g_2} + \frac{\partial J\Pi_2}{\partial g_2}\right] - (1+\lambda)p_2\left[n_2^*\frac{\partial n_2^*}{\partial g_2} + \left(K_2 - n_2^*\right)\frac{\partial n_2^*}{\partial g_2} - \frac{\partial Q_2}{\partial n_2^*}\frac{\partial n_2^*}{\partial g_2}\right] -$$

$$- (1+\lambda)p_2\frac{\partial R^b\left(K_2, g_2; \alpha_1\right)}{\partial g_2} = 0$$

Hence, it can be immediately seen that the allocation of g_i and K_i will differ across the various types. In particular, solving explicitly the first-order conditions above, it can be seen that the allocation of both g_1 and K_1 - targeted at hospitals with lower case-mix – will not be distorted with respect to the full information benchmark:

$$g_1^b = g_1^{b,pi} = \frac{\alpha_1 + 1/2 + r}{2\lambda + 3} \qquad K_1^b = K_1^{b,pi} = \frac{(\alpha_i + 1/2)(2+\lambda)}{2\lambda + 3} - \frac{r(4+3\lambda)}{2\lambda + 3} \tag{8}$$

As a consequence, the number of patients of *type 1* hospitals treated in equilibrium corresponds to the number treated in first best:

$$n_1^b = n_1^{b,pi} = \frac{(\alpha_1 + 1/2)(2+\lambda)}{2\lambda + 3} - \frac{r(1+\lambda)}{2\lambda + 3}$$

This set of results is standard in asymmetric information problems and is known as the property of *non-distortion at the top.*

However, because both g_2 and K_2 are allocated taking into account the effects conveyed by the extra terms, the contract intended for *type 2* hospitals clearly differs from the full information benchmark. In fact, the government will choose the optimal allocations g_2 and K_2 in such a manner as to trade off between the allocative efficiency goal and the minimisation of the informational rent.

More precisely, the exact choice of $\{K_2, g_2\}$ in the optimal contract will depend on the interaction between the impact of these policy instruments on the informational rent reaped by *type 1* hospitals and their direct and indirect effects on the other objectives envisaged by the government. Because $\partial R/\partial g_2 = \partial R/\partial K_2 < 0$, the minimisation of the informational rent implies that at least one of these two instruments should be over-provided by the government. However, the socially-optimal care supply must also consider patients' health status, hospitals' surplus, and the number of patients treated within the hospital. From the government's perspective, the latter target should optimally trade off the benefit from the enhancement of patients' health status with the reduction of the overall costs in health care provision, including the cost of capacity, the fees paid to doctors and the costs due to the sub-optimal utilisation of medical technology or to congestion, as captured by the terms *r*, $n_2^*\partial n_2^*/\partial K_2$, $n_i^*\partial n_2^*/\partial g_2$, $\left(K_2 - n_2^*\right)\left(1 - \partial n_2^*/\partial K_2\right)$ and $\left(n_2^* - K_2\right)\left(\partial n_2^*/\partial g_2\right)$, respectively. Therefore, the optimal choice of g_2 and K_2 will ultimately depend on the complex interaction between all of the above effects.

18

Solving explicitly the first-order conditions, the optimal menu of contracts can be characterised as,[19]

$$g_2^b = \frac{2\lambda(\alpha_2 - \alpha_1) + \alpha_2 + r + 1/2}{2\lambda + 3} + \lambda r \quad n_2^b = \frac{2(2\alpha_2 - r)(1+\lambda) + (2+\lambda) - 2\lambda\alpha_1}{2(2\lambda + 3)}$$

$$K_2^b = \frac{2\lambda(\alpha_2 - \alpha_1) + (\lambda + 2)(2\alpha_2 + 1) - 4(\lambda + 2)(1+\lambda)r}{2(2\lambda + 3)}$$

(9)

Comparing with the optimal contract under perfect information, it can be seen that the optimal contract's fee for patients is always higher than in a first best allocation, i.e., $g_2^b - g_2^{b,pi} > 0$. Because the effects on the (reduction of) informational rent and on (the maximisation of) patients' health status prevail, hospitals with a more severe case-mix receive higher fees than under full information.

Moreover, it can be seen that the difference between the capacity levels under asymmetric and full information depends on the cost of capacity r:

$$K_2^b - K_2^{b,pi} = \frac{\lambda(\alpha_2 - \alpha_1)}{3 + 2\lambda} - r\lambda$$

(10)

In particular, for sufficiently *low (high) cost of capacity*, namely if $r < (>)(\alpha_2 - \alpha_1)(3 + 2\lambda)$, the capacity allowed to *type 2* hospitals is greater (lower) than under perfect information, $K_2^b > K_2^{b,pi}$ ($K_2^b > K_2^{b,pi}$). As capacity becomes cheaper, the government is willing to over-provide capacity to hospitals with more severe case-mixes.

That the allocations of both K_2 and g_2 are distorted is standard under asymmetric information contracts. However, it should be noticed that even though both instruments are perfect substitutes (both in controlling the informational rent and in influencing the equilibrium number of patients and, thus, health status) the extent to which one will be preferred over the other will ultimately depend on the relative (social) cost of their usage. In that respect, when capacity becomes too expensive, then the government will substitute it away for g_2.

Finally, the above trade-offs also imply that the number of patients treated in high case-mix hospitals is larger than under perfect information: $n_2^b - n_2^{b,pi} = \lambda(\alpha_2 - \alpha_1)/(2\lambda + 3) > 0$. Note that this distortion, with respect to the first best, not only does not depend on the capacity cost, but it is also greater than the corresponding positive distortion in the capacity. Hence, hospitals with a more severe case-mix treat more patients, and such a difference is greater than in the first best, i.e.,

$$n_2^b - n_1^b \equiv 2(\alpha_2 - \alpha_1)(1+\lambda)/(3+2\lambda) > (\alpha_2 - \alpha_1)(2+\lambda)/(3+2\lambda) \equiv n_2^{b,pi} - n_1^{b,pi}.$$

5. Contracting Scenario

5.1 Managers' and doctors' decisions

In the contracting scenario, doctors and managers make simultaneous decisions separately on n_i and h_i, respectively. After the government sets the optimal contract, doctors decide on the number of patients to be treated with therapy *H* by maximising their utility function (1), treating the managers' decisions as given. The

---

[19] The second-order conditions are always satisfied. In fact, it is possible to see that the four leading minors are of alternating signs with the sign of the first leading minor being negative, i.e., $|H1| < 0$

optimal number of treated patients is the level $n_i(h_i) = \arg\max\left\{\Pi_{D_i} \equiv Q_i(n_i; \alpha_i) + (h_i + g_i)n_i - \frac{n_i^2}{2}; 0\right\}$.

Because doctors treat patients with technology *H* until the treatment's marginal benefit equals its marginal cost, it follows that the doctor's best response function is: $n_i(h_i) = \left(\left(\alpha_i + \frac{1}{2}\right) + h_i + g_i\right)/2$.

Treating the doctors' decision as given, managers offer the fee h_i that maximises their own utility (2) subject to the doctors' participation constraint. The optimal doctors' fee is the level $h_i(n_i)$ that solves:

$$\underset{h_i}{Max}\ \Pi_{Mi} = T_i - \frac{(n_i - K_i)^2}{2} - n_i^* h_i \quad s.t. \quad \Pi_{Mi} \geq 0 .$$

That is, in the contracting scenario, the managers offer the lowest possible level of the fee that satisfies the doctors' non-negative utility constraint. It can be checked that because $\partial\Pi_{Mi}/\partial h_i < 0$ and $U\big|_{h_i=0} > 0,$ the manager will optimally set $h_i^* = 0$.[20]

Substituting on the doctors' reaction function, the equilibrium number of patients is given by $n_i^* = \left(\left(\alpha_i + \frac{1}{2}\right) + g_i\right)/2$. Notice that the equilibrium number of treated patients n_i^* increases with the case-mix α_i. Also notice that, under a contract scenario, n_i^* is independent of the capacity K_i contracted by the government, and it clearly depends on the fee g_i set by the government. Moreover, contrary to the bargaining scenario, the impact of the capacity and the fee for services offered by the government is no longer symmetric. Indeed, while the capacity has no impact on doctors' choices ($\partial n_i^*/\partial K_i = 0$), fee for service positively affects the number of patients treated, and this effect is greater than the analogous effect under bargaining scenario. The same holds with respect to health benefits: $\partial Q_i/\partial g_i = \left(\alpha_i + 1/2 - n_i^*\right) > 0, \partial Q_i/\partial K_i = 0.$

5.2 Government's problem

5.2.1 Full information
Analogously to the bargaining scenario, in a full information scenario, the menu of contracts offered by the government is given by: [21]

$$g_i^{c,pi} = \frac{\alpha_i + 1/2 - 2r(1+\lambda)}{2\lambda + 3}, \quad K_i^{c,pi} = \frac{(2+\lambda)(\alpha_i + 1/2) - r(4 + 3\lambda)}{2\lambda + 3}$$

$$n_i^{c,pi} = \frac{(2\alpha_i + 1)(\lambda + 2) - 2r(1+\lambda)}{2(2\lambda + 3)}, \quad n_2^{c,pi} - n_1^{c,pi} = (\alpha_2 - \alpha_1)\frac{2+\lambda}{2\lambda + 3}$$

(11)

5.2.2 Asymmetric information

Because the policy that would be optimal in case of full information cannot be implemented when the government is not aware of the hospital's type, the menu of contracts is designed by the government to

---

[20] Notice that such a "corner" solution is also common to Boadway et al. (2004) and serves mainly as a benchmark scenario.

[21] The subscript c, pi has been used to identify the perfect information contract scenario.

maximise the total welfare function *W* subject to the usual *Individual Rationality* (*IR$_i$*) and *Incentive Compatibility* (*IC$_i$*) constraints for both types $i = 1, 2$, as in the bargaining scenario.

It can be checked that the *Single Crossing Conditions* are also verified in this scenario (see Appendix B).

In the optimum, the contract for *type 1* is not distorted with respect to the perfect information benchmark (11) [22] ($g_1^c = g_1^{c,pi}, K_1^c = K_1^{c,pi}$ and $n_1^c = n_1^{c,pi}$) , and the contract for *type 2* hospitals is such that:

$$g_2^c = \frac{2\lambda(\alpha_2 - \alpha_1) + \alpha_2 + 1/2 - 2r(1+\lambda)}{2\lambda + 3} \qquad n_2^c = \frac{2(2\alpha_2 - r)(1+\lambda) + (2+\lambda) - 2\lambda\alpha_1}{2(2\lambda + 3)}$$

$$K_2^c = \frac{\lambda(\alpha_2 - \alpha_1)(5 + 2\lambda) + (2\alpha_2 + 1 - 4r(1+\lambda))(\lambda + 2)}{2(2\lambda + 3)}$$

(12)

Finally, hospitals serving populations with a higher average illness severity treat more patients than hospitals with a lower average illness severity,

$$n_2^c - n_1^c = \frac{2(\alpha_2 - \alpha_1)(1+\lambda)}{2\lambda + 3} > 0$$

Interestingly, this difference is the same under the bargaining and the contracting scenario.

Comparing the fee for *type 2* with the full information benchmark: $g_2^c - g_2^{c,pi} = 2\lambda(\alpha_2 - \alpha_1)/(2\lambda + 3) > 0$, as in the bargaining scenario, *type 2* hospitals are given higher fees than under perfect information. However, given $g_2^b - g_2^{b,pi} > 0$, this distortion is smaller in the contracting than in the bargaining scenario.[23]

Proceeding in a manner that is analogous to the bargaining scenario, it can be shown that although both K_2 and g_2 have a negative impact on the informational rent,

$$\frac{\partial R^c(K_2, g_2; \alpha_1)}{\partial K_2} = \frac{\alpha_1 - \alpha_2}{2} < 0, \quad \frac{\partial R^c(K_2, g_2; \alpha_1)}{\partial g_2} = \frac{\alpha_1 - \alpha_2}{4} < 0$$

capacity K_2 is twice as effective in reducing the informational rent with respect to the fee for service g_2: $\left| \partial R^c / \partial K_2 \right| > \left| \partial R^c / \partial g_2 \right|$.

Comparing capacity levels: $K_2^c - K_2^{c,pi} = \frac{\lambda}{2}\left[\frac{2\lambda + 5}{2\lambda + 3}(\alpha_2 - \alpha_1) - 2r \right]$, it can be seen that the capacity allowed to *type 2* hospitals is the same as under perfect information only in the special case in which the cost of capacity assumes the punctual value $r = (\alpha_2 - \alpha_1)(2\lambda + 5)/(2(2\lambda + 3))$. For sufficiently low (high) capacity prices, the *distortion* in the allowed capacity is *positive* (negative): $K_2^c > K_2^{c,pi}$ $(K_2^c < K_2^{c,pi})$. That is, the outcome and the distortion with respect to the perfect information case are qualitatively the same under the contracting and bargaining scenarios.

---

[22] The expressions for the lump-sum transfers are cumbersome. Given that their role is not very informative, we do not explicitly state them in the paper.

[23] In particular, the distortion is smaller by a factor $r\lambda$, which is directly related to the cost of capacity. Indeed, though under the bargaining scenario g_2 and K_2 were found to be perfect substitutes and the government was able to directly substitute away K_2 with g_2 when the cost of capacity was high, this is no longer the case under the contracting scenario.

6. Discussion

6.1 The results

In the last sections, we presented the optimal financing contracts offered by a third-party payer under bargaining and contracting scenarios within the hospital. In this section, we discuss those findings further, starting from the benchmark case of full information and moving to the asymmetric information case. The following propositions summarise the main results.

Proposition 2. Under full information, the government is able to achieve the same levels of technological capacity, patients treated and social welfare in both the bargaining and the contracting scenario.

Proposition 3. Under asymmetric information, for sufficiently low levels of the cost of capital r, the government allows lower informational rents to hospitals and attains higher levels of social welfare under the contracting scenario.

6.1.1 Full information

As a benchmark, we will first proceed with a direct comparison between the government's contracts that would be offered in the two scenarios under *perfect information*. In particular, we first compare the *fee* for service that would be paid to doctors under the two scenarios, and then briefly mention the levels of *capacity*, the *number of patients* treated and the *social welfare* across the two scenarios.

Regarding the *fees for service* g_i, it can be seen that under perfect information, these are higher in the bargaining than in the contracting scenario: $g_i^{b,pi} - g_i^{c,pi} = r > 0$. Notice that the difference equals the cost of capacity. Intuitively, this is related to the different role of g_i under the two scenarios. In fact, as discussed above, the fee paid by the government to the doctors is a policy instrument that can be used in both scenarios to impact the different elements that determine social welfare: the *informational rent*; the social and private *costs of treatment* and their related monetary incentives; and the *number of patients* treated and their health benefit.

The *informational rent* is clearly not an issue in the perfect information case. With respect to the monetary incentives related to the *costs of treatment*, the two scenarios differ. First, while in the bargaining scenario, the equilibrium fee for service h_i^* is set such that the hospital's surplus is shared equally between doctors and managers, in the contracting scenario, h_i^* is set so that doctors are provided the minimum monetary payment to induce their participation. Therefore, the government's scope to integrate monetary payments and to induce optimal doctors' behaviour is clearly higher in the contracting than in the bargaining scenario. Consequently, because h_i and g_i are perfect substitutes for doctors, one should expect higher fees paid by the government to doctors when managers choose h_i unilaterally than when it is decided through bilateral negotiations. However, the different manner in which the number of patients is set in the two scenarios may counterbalance and even outweigh this effect. In fact, in the *bargaining* scenario, the number of patients is also negotiated by both managers and doctors, and therefore, in equilibrium, it is set at a level that encompasses the interests of both parties, taking into account the direct and indirect costs of the hospital, including the costs associated with the inefficient use of capacity. In contrast, under the *contracting* scenario, the number of treated patients is decided exclusively by the doctors in an opportunistic way, disregarding the impact of their decision on the overall costs

22

to the hospital. In the latter scenario, doctors are more likely to strategically manipulate the number of treated patients to compensate the loss in their revenues due to the zero fees paid by managers with artificially high revenues earned by inflating the number of treated patients above the optimal level. The complete split in decision-making between the doctors and managers within the hospital described in the contracting scenario, would thus generate not only a failure of coordination on the optimal level of patients to be treated (from the hospital's point of view), but also a negative externality on the government in the form of higher fees to be paid to doctors. To counterbalance this effect, the government reduces its fee g_i to compensate or even to dominate the first effect.

To implement a (socially) optimal *number of patients treated*, the government must weigh increases in overall costs of health care provision against improved health outcomes. As discussed in the previous sections, the fee for service paid to doctors is one of the key instruments that can be used to achieve this result: in particular, an increase in the fee for service g_i induces a higher number of treated patients in both organisational scenarios. However, because the absolute size of such a positive effect is higher in the contracting scenario $(\partial n_i^* / \partial g_i^c > \partial n_i^* / \partial g_i^b)$, a smaller fee level in the contracting scenario is sufficient to induce the same number of treated patients as in the bargaining scenario: this may further explain why $g_i^{b,pi} > g_i^{c,pi}$. Furthermore, in the contracting scenario, doctors decide n_i^* without internalising the overall costs to the hospital. Therefore, because the number of treated patients set in equilibrium by the doctors does not depend on the capacity offered by the government, g_i is the only instrument that can be used by the government to control such number. In contrast, in the bargaining scenario, both g_i and K_i can be used by the government to control the equilibrium number of treatments. Moreover, because g_i and K_i have exactly the same impact on n_i^* in the bargaining scenario, g_i and K_i are perfect-substitute policy instruments concerning this objective. Hence, under bargaining, the government can use g_i instead of K_i as an instrument to affect n_i^* whenever capacity is a relatively more expensive policy tool. Due to the cost of capital *r*, the relative price of fees with respect to capacity is *1/(1+r)*. Hence, in the bargaining scenario, whenever the government substitutes away one unit of capacity K_i with a unitary increase in the fee g_i, it saves an amount r, which also explains why the difference between the full information fees under the two scenarios equals $g_i^{b,pi} - g_i^{c,pi} = r$.

Furthermore, under perfect information, the government would allow exactly the same *capacity* under both scenarios $K_i^{b,pi} = K_i^{c,pi}$. With perfect information, capacity is chosen to maximise overall efficiency and to minimise the total costs of health care provision, including the costs of capacity, which are identical in both scenarios. Therefore, the *number of treated patients* is also identical under both the bargaining and the contracting scenarios, although, as discussed above, the policy instruments used by the government to achieve such a result differ between the two scenarios. This result is of potential interest to policy makers. In fact, it suggests that, whenever the third-party payer is in the position of obtaining full information from the hospital regarding the disease severity of the patients, it is possible to completely neutralise the impact of different governance within the hospitals by designing contracts that implement the same envisaged outcomes independently of the organisation and governance framework adopted by the hospitals. In particular, the third-party payer is able to achieve not only the same desired level of technological capacity - and therefore of health care costs - but also the same health outcomes. These are the reasons why, under full information, government contracts achieve the same level of social welfare under both the bargaining and the contracting scenario.

6.1.2 Asymmetric information

Under *asymmetric information*, given the *non-distortion-at-the-top* result, the difference in the contracts offered to *type 1* hospitals between bargaining and contracting scenarios exactly corresponds to the analogous difference under *perfect information*. However, perfect information contracts targeted at higher case-mix hospitals are distorted under asymmetric information to minimise the informational rent. In what follows, we will compare, for *type 2* hospitals: first, the levels of *fees* and *capacity* under the two scenarios; then, the *informational rents*; and, finally, the overall *social welfare* associated with the contracts offered by the government in the two scenarios.

In the bargaining scenario, the *fee for service* g_2 and the *capacity* K_2 are perfectly substitutable in reducing the informational rent, while in the contracting scenario K_2 is twice as effective. In fact,

$$\left|\frac{\partial R^c}{\partial K_2}\right| = \frac{\alpha_1 - \alpha_2}{2} > \left|\frac{\partial R^b}{\partial K_2}\right| = \left|\frac{\partial R^b}{\partial g_2}\right| = \frac{\alpha_1 - \alpha_2}{3} > \left|\frac{\partial R^c}{\partial g_2}\right| = \frac{\alpha_1 - \alpha_2}{4}.$$

Therefore, *ceteris paribus*, the distortion on the fee for service for the *type 2* hospitals should indeed be larger in the bargaining than in the contracting scenario: $g_2^b - g_2^c = (\lambda + 1)r > 0$. Indeed, because the use of capacity is costly, under the bargaining scenario, when the government substitutes away one unit of K_2 with a unitary increase in g_2, it still experiences the same envisaged reduction in the informational rent while it also benefits from a savings equal to *r*.

Therefore, when the cost of capital *r* is low, the capacity allowed to *type 2* hospitals under asymmetric information is greater than would be optimal under perfect information in both contractual scenarios: $K_2^b > K_2^{b,pi}$, $K_2^c > K_2^{c,pi}$. However, the effect is not the same in magnitude across the two organisational scenarios: in particular, for any given level of the costs of capital, the capacity allowed to hospitals of *type 2* is always smaller in the bargaining than in the contracting scenario: $K_2^b - K_2^c = \frac{\lambda}{2}(\alpha_1 - \alpha_2) < 0$. Indeed, for a given cost of capacity *r*, K_2 is a less cost-effective policy in the bargaining than in contracting scenario. Therefore, because the optimal level of capacity in the perfect information case is the same under both scenarios - even for low capacity costs - the over-provision of capacity, with respect to the optimum, is less pronounced under the bargaining than in the contracting scenario.

The next result we are going to compare across different scenarios is the *informational rent*. When the costs of capacity are sufficiently low (high), the rent allowed to *type 2* hospitals is lower (higher) in the contracting than in the bargaining scenario (comparisons in Appendix C1).

The ultimate reason for such a result is attributed to the interactions and trade-offs among all of the described effects in both scenarios. In the *contracting* scenario, the government faces more than one trade-off when designing the menu of contracts to offer hospitals with a more severe case-mix. On the one hand, an increase in the capacity K_2 is twice as effective in reducing the informational rent as an equivalent increase in fees paid to doctors g_2. On the other hand, capacity K_2 is also a more costly policy instrument than the fees paid to doctors. Hence, in the control of informational rents, the relative use of these two instruments will be decided on cost-effectiveness grounds. In particular, when *r* is low enough, the allowed capacity K_2 is much more cost-

24

effective in reducing the informational rent than the fee g_2; therefore, the government uses K_2 more intensively than g_2 to reduce the informational rent.

Note that while K_2 and g_2 are (imperfect) substitute instruments *vis-à-vis* the informational rent, in the contracting scenario they cannot be used interchangeably to control the number of treated patients or to balance total costs of health care against patients' improved health. Indeed, because doctors decide n_i^* independently (without internalising managers' objectives), their decision is not affected by capacity levels $\left(\partial n_2^* / \partial K_2\right)^c = \left(\partial Q_2^* / \partial K_2\right)^c = 0$. Moreover, the fee paid by managers does not affect doctor's decisions because in equilibrium it is set at $h_i^* = 0$. Therefore, under the contracting scenario, the fee for service g_2 is the only available instrument to control the number of treated patients and will be set to optimally balance the direct and indirect costs of health care provision and the enhancement of patients' health status. It is as if the separation of decisions within the hospital that is typical of the contracting scenario induces the government to use its main instruments K_i^c and g_i^c in a similar "dedicated" way, by clearly separating their role and their impact: the capacity K_i^c is used to minimise the informational rent, provided that the cost of capacity *r* is low enough; the fee for service g_i^c is set to affect the doctors' decision on the number of treated patients and to optimally balance the improvement in health status with the total costs of health care.

Also note that, in the *bargaining* scenario, there is a trade-off in the use of K_i^b against g_i^b. The government knows that the decision-making within the hospital is not separate but integrated, such that doctors and managers negotiate simultaneously both on the number of treated patients and the fees for doctors. Hence, the government can no longer easily identify which agent is ultimately responsible for an objective and which instrument is more appropriate to induce desired changes in the behaviour and the relevant variables. In fact, because K_2 and g_2 are *perfect* substitutes in their impact on the rent and on the number of treated patients, the government uses the instrument associated with the lowest cost of utilisation, namely, g. Therefore, it is as if the integrated decision-making within the hospital induces the government to also treat its different policy targets in an integrated manner and to use only doctors' fees for services as a policy instrument.

Therefore, in the bargaining scenario, g_i^b is the unique policy instrument tied to the achievement of government's two main objectives: the minimisation of the informational rent, and the implementation of the socially optimal number of treated patients. Thus, the fee paid to doctors is set to balance these two different objectives.

Hence, in the bargaining scenario, g_i^b is necessarily an imperfectly accurate instrument for targeting the specific objective of minimising the informational rent. This may explain the occurrence described above of higher informational rents under the bargaining than the contracting scenario. In the contracting scenario, as long as the cost of capacity *r* is low, K_i^c is used as a dedicated instrument to minimise the informational rent, while g_i^c is exclusively targeted at the control of the optimal number of treated patients. Consequently, when the cost of capacity is sufficiently low, there are two more precise, tailored and specialised instruments in the contracting scenario, instead of only one as in the bargaining scenario. This explains why a lower informational rent is achieved under the contracting scenario.

Finally, we explicitly compare the *social welfare* attained under the bargaining and the contracting scenarios under asymmetric information. It is found that under asymmetric information, for low levels of the cost of capital *r*, the contracting scenario delivers a higher social welfare, while for sufficiently high levels of *r*, the bargaining scenario is associated with higher values of social welfare (see Appendix C2). The fact that, under full information, welfare is the same in both scenarios (see Proposition 2) is perfectly in line with the comparison of the informational rents. In fact, because for a given level of health care treatment, informational rents increase the government's costs, social welfare is clearly higher when informational rents are minimised. This occurs under the contracting scenario when the cost of capital is low and in the bargaining scenario otherwise.

6.2 Modelling issues and extensions

Even though we have found some results of potential interest within our model, it should be noted that the analysis could be extended in various ways. In particular, throughout the paper, we assume that the outside option for both doctors and managers confers them a profit and utility level of zero in both the bargaining and the contracting scenario. This is a very stylised version of what happens in reality because physicians can often choose among different employers, either public or private, who demand clinical services or even non-clinical services (e.g., research). The importance of these outside options is reinforced by the practice of maintaining two jobs, which is often designated in the literature as "moonlighting" (Biglaiser and Ma, 2006). Indeed, in many countries, it is common for physicians to work for two different public employers or even to work in both the public and private sector (for a review of physician dual practice see García-Prado and Gonzalez, 2006). The existence of a private sector alternative, which very often offers better working conditions (Biglaiser and Ma, 2006), empowers physicians in their contractual arrangements *vis-a-vis* managers in the public sector and should, therefore, be accounted for in the analysis of negotiations within the hospital. In particular, there are *two* main analytical *options* that could be used to introduce this issue in our scenario.

The *first option* would be to model such an outside option as deriving *endogenously* from the private sector, where doctors can strategically decide to supply their labour when formulating their decisions within the public hospital. This model would account for an explicit interaction between the private and the public hospitals and for a direct effect of the endogenous outside option on doctors' bargaining power. This general equilibrium approach would be extremely interesting, but because it is beyond the scope of the present paper, it is left for future work.

The *second option* would be to model doctors' and managers' outside options as *exogenous*. In turn, at least *four* solutions are possible for modelling *exogenous outside options*.

1. The first solution is the one we propose in the model. To keep the bargaining protocol as neutral as possible, we assume that doctors and managers enjoy symmetric, exogenous outside options, standardised to zero for the sake of analytical tractability.

2. A second solution for modelling exogenous outside options would be to consider that the alternative wage paid within the private sector, ω, represents a strictly positive outside option to doctors. It can be quickly checked that this modelling variation would leave the qualitative results of the current version of our model completely unaffected. In fact, both scenarios would simply reduce to a re-scaling of the doctor's utility function. In particular, in the *bargaining* scenario, one can modify the existing negotiation protocol by imagining that, in the case that doctors reject an offer made by the managers, they can opt out to work into the private sector, receiving a salary ω for one period, while still continuing negotiations with the managers. If so, by an argument similar to the one above, it must be that in the unique sub-game perfect Nash equilibrium, doctors would receive from the managers an offer such that

26

$\Pi_i[D](h_i[M], n_i[M]) = \omega + \delta V_i^*[D]$. Therefore, the two continuation payoffs, the expected payoffs and the equilibrium fee for service offered by doctors and managers would all also be functions of ω. In addition, an analogous introduction of a positive outside option in the *contracting* scenario would imply that the doctors' participation constraint in the decision by the managers would be binding at the ω level. It can also be seen that the effect of this modelling variation in both scenarios would consist of re-scaling the final equilibrium levels of the fees by a factor depending on ω. However, the introduction of this new variable would also make all the analytical expressions in the model much more cumbersome. Because the main objective of our work is to compare the salient equilibrium outcomes across scenarios, and because by introducing the outside option ω, the difference between the fee for service in the bargaining scenario and the *contract* scenario was still positive and substantially unaltered with respect to the current version of the model, we chose to standardise ω to zero for the sake of clarity and analytical tractability.

3. A third solution for modelling exogenous outside options would be to introduce a non-zero, possibly asymmetric, outside option for the managers in terms of the unitary income, *v*, earned in an alternative private sector whenever they disagree with the doctors in the hospital's negotiation process. The bargaining protocol would be modified accordingly in a manner similar to that previously discussed for the doctors. The final result would correspondingly be a simple re-scaling of the equilibrium outcomes at the expense of a further complication of the analytical expressions.

4. Finally, a fourth modelling solution for exogenous outside options would consist in interpreting the bargaining power of doctors and managers in the negotiation protocol as an implicit function of their relative outside options. Our model can fully account for this solution. In fact, in the current bargaining protocol, doctors and managers make offers at each negotiation stage with a probability *β* and *1-β* respectively. Clearly, the higher the probability of doctors making an offer, for example, the more likely that they enjoy a first mover advantage in the negotiations, and the better are their contractual conditions at the end of the game. As can be clearly seen in the formal analysis (in Appendix A), the proposed solution of the bargaining game is such that the probabilities *β* and *1-β* of doctors and managers being selected to make an offer do not affect the number of treated patients, which is set to such a level as to maximise hospitals' joint surplus. Rather, they only affect the internal distribution of resources in terms of higher or lower levels of fee for service paid to doctors: intuitively, both of the fees proposed in equilibrium, $h^*(M)$ and $h^*(D)$, are increasing in *β*, the probability that doctors make offers. Thus, a modelling variation would be to let *β* be a function of the outside options of doctors and managers, for instance of the type *β (ω,v)* with *0 ≤ β(ω,v) ≤ 1*, *∂β(ω, v)/∂ω>0*, and *∂β(ω, v)/∂v<0*. This would be a direct extension of the model, but would imply a great complication of the analysis in terms of explicit tractability.

7. Conclusion

We have studied the impact of different hospitals' organisation and governance solutions on the optimal contracts designed by third-party payers when disease severity is the hospital's private information. In particular, we have compared a *contracting* scenario, in which managers offer a contract setting the payment to doctors and doctors decide how many patients to treat, with a *bargaining* scenario, in which doctors and managers strategically negotiate on both the payment to doctors and the number of patients.

We have shown that the nature of governance and decision-making within the hospital affects hospital performance and consequently shapes the contracts that third-party payers offer hospitals.

Our model shows not only that the government's contracts must be tailored according to the hospitals' organisational structure, but also that their effectiveness in eliciting information and enhancing social welfare

differs across hospitals' governance scenarios.

Our results support the idea that radical organisational reforms within hospitals must be based on a thorough and articulated analysis of the policy implications. A scenario under which all strategic variables are negotiated within the hospital between managers and doctors can lead to less opportunistic and less biased unilateral behaviour regarding key financial and medical decisions, namely, the monetary incentives to doctors and the number of treated patients. However, a bargaining scenario can also lead to further distortions in the government's ultimate objectives, and, at least when the cost of capacity is sufficiently low, to less accurate and less effective policy tools aimed at reducing informational rents accruing to secondary care providers. In particular, the most desirable option from the government's perspective ultimately depends on the cost of capacity. Namely, for a low (high) cost of capacity, the contracting (bargaining) scenario delivers higher social welfare.

Although the main focus of the paper is on the theoretical analysis, our model has a number of potential policy implications. To the extent that technological investments and, therefore, the cost of invested capital, is specific to the type of disease and health treatment, one of the main policy implications of our analysis is that the optimal form of hospital governance may vary across different clinical specialities. For example, while expensive robotic surgery is currently widely-used in urology, in orthopaedics, surgeons still rely on conventional surgery methods. Our model suggests that the differences in the capital investments in medical technologies across these specialities certainly establish the case for the design of speciality-specific financing contracts.

Furthermore, if we interpret the cost of capacity as the cost of capital, because public and private providers are likely to face different capital market conditions, the policy implications of our results will depend on the type of providers serving the market. Because public providers normally benefit from a rate of return below the commercial rate faced by their private counterparts,[24] we would expect a contracting scenario to be more effective as an organisational structure if the secondary care market is served by public providers.

Finally, our model emphasises the key role of the governance frameworks adopted within the hospitals and their crucial effect on the financial flows from the third-party payer. Indeed, even if the model does not capture all of the complexity of real hospitals,[25] it at least partially mirrors the institutional landscape created by recent reforms in the healthcare sector. In the UK, for instance, one of the main institutional innovations in recent years has been the transformation of some hospitals into *Foundation Trusts*. Foundation Trusts are residual claimants of their gains and losses, have large managerial and financial independence, and are run by a *board of governors*, which is largely constituted by members of the staff and representatives of the public and the patients, and by a *board of directors*, constituted by executive and non-executive directors.

Our model can thus be applied to the analysis of the different types of interaction between these two bodies, loosely corresponding to what we denote *doctors* and *managers*, respectively. For instance, a scenario in which the different actors within the hospitals strategically bargain over all of the relevant economic variables may be a preferable governance solution to Foundation Trusts if the cost of the capital investments is relatively high. This is more likely to be the case for Foundation Trusts that offer a service mix characterised by large investments in capital-intensive medical technologies.

---

[24] For example, in England, the commercial rate of return faced by private providers is 6.1% while the analogous rate for public providers is 3.5% (Mason *et al* 2007).
[25] For instance, an interesting line for future research concerns how different organisational scenarios can endogenously emerge through strategic interaction among the actors within the hospital.

28

References

1. Alchian AA, Demsetz H. Production, information costs, and economic organization. American Economic Review 1972; 62; 777-795.
2. Baker L, Birnbaum H, Geppert J, Mishol D, Moyneur E. The relationship between technology availability and health care spending. Health Affairs 2003; 3; 537-551.
3. Baron D, Myerson R. Regulating a monopolist with unknown costs. Econometrica 1982; 50; 911–930.
4. Beitia A. Hospital quality choice and market structure in a regulated duopoly. Journal of Health Economics 2003; 22; 1011–1036.
5. Biglaiser G, Ma CT. Moonlighting: public service and private practice. Boston University Working paper 2006.
6. Bloch F, Ghosal S. Buyers' and sellers' cartels on markets with indivisible goods. Review of Economic Design 2000; 5; 129-147.
7. Bloch F, Gomes A. Contracting with externalities and outside options. Journal of Economic Theory 2006; 127; 172-201.
8. Blomqvist AG, Carter RA. Is health care really a luxury? Journal of Health Economics 1997; 16; 207-229.
9. Boadway R, Marchand M, Sato M. An optimal contract approach to hospital financing. Journal of Health Economics 2004; 23; 85-110.
10. Bradley TB, Kominski GF. Contributions of case mix and intensity change to hospital cost increases. Health Care Financing Review 1992; 14; 151-163.Brown M. Economic analysis of hospital operation. Hospital Administration 1970; 15; 67-74.
11. Bullun SE. Endometriosis. New England Journal of Medicine 2009; 360; 1911-1912.
12. Chalkley M, Malcomson JM. Contracting for health services with unmonitored quality. The Economic Journal 1998a; 108; 1093-1110.
13. Chalkley M, Malcomson JM. Contracting for health services when patient demand does not reflect quality, Journal of Health Economics 1998b; 17; 1-19.
14. Chalkley M, Malcomson JM. Cost sharing in health service provision: an empirical assessment of cost savings. Journal of Public Economics 2002; 84; 219–249.
15. Chernew M, Fendrick AM, Hirth RA. Managed care and medical technology: implications for cost growth. Health Affairs 1997; 16; 196-206.
16. Coase R. The nature of the firm. Economica 1937; 4; 386-405.
17. Conrad D, Wickizer T, Maynard C, Klastorin T, Lessler D, Ross A, Soderstrom N, Sullivan S, Alexander J, Travisk K. Managing care, incentives, and information: an exploratory look inside the "Black box" of hospital efficiency. Health Services Research 1996; 31; 235-259.
18. Crilly T, Le Grand J. The motivation and behaviour of hospital trusts. Social Science and Medicine 2004; 58; 1809-1823.
19. Cromwell J, Butrica B. Hospital department cost and employment increases: 1980-92. Health Care Financing Review 1995; 147-166.
20. Custer WS, Moser JS, Musacchio RA, Willke RJ. The production of health care services and changing hospital reimbursement: the role of hospital-medical staff relationships. Journal of Health Economics 1990; 9; 167—192.
21. Cutler DM, Huckman RS. Technological development and medical productivity: the diffusion of angioplasty in New York State. Journal of Health Economics 2003; 22; 187-217.
22. Cutler D, McClellan M, Newhouse JP. The costs and benefits of intensive treatments for cardiovascular disease. In Triplett J (Ed): Measuring the prices of medical services. The Brookings Institution: Washington DC; 1999.
23. Davis K. Economic theories of behaviour in non-profit, private hospitals. The Brookings Institution: Washington DC; 1971.
24. De Fraja G. Contracts for health care and asymmetric information. Journal of Health Economics 2000; 19;

663-677.

25. Department of Health. A guide to NHS Foundation Trusts. Department of Health: London; 2002a. Also available at http://www.dh.gov.uk/assetRoot/04/06/04/80/04060480.pdf

26. Department of Health. Reforming NHS financial flow: introducing Payment by Result. Department of Health: London; 2002b.

27. Department of Health. NHS Foundation Trusts: A sourcebook for developing governance arrangements. Department of Health: London; 2006. Also available at http://www.dh.gov.uk/en/Publicationsandstatistics/Publications/PublicationsPolicyAndGuidance/DH_4130231

28. Ellis RP, McGuire TG. Provider behaviour under prospective reimbursement: cost sharing and supply. Journal of Health Economics 1986; 5; 129-51.

29. Ellis RP. Creaming, skimping and dumping: provider competition on intensive and extensive margins. Journal of Health Economics 1998; 17; 537-55.

30. Feldstein P. Applying economic concepts to hospital care. Hospital Administration 1968; 4.

31. Fuchs VR. Economic values and health care reform. American Economic Review 1996; 8; 1-24.

32. Fudenberg D, Tirole J. Game theory. MIT Press: Cambridge Massachusetts; 1991.

33. Gaffney D, Pollock A, Allyson M, Price D, Shaoul J. NHS capital expenditure and the Private Finance Initiative. British Medical Journal 1999; 319.

34. Garcia-Prado A, Gonzalez P. Who do physicians work for? Working paper Department of Economics Universidad Pablo Olavide 2006.

35. Goldman D, Shang B, Bhattacharya J, Garber A, Hurd M, Joyce C, Lakdawalla D, Panis C, and P Shekelle. Consequences of health trends and medical innovation for the future elderly. Health Affairs 2005.

36. Harris JE. The internal organization of hospitals: Some economic implications. Bell Journal of Economics 1977; 8; 467-488.

37. Hay JW. Hospitals cost drivers: An evaluation of 1998-2001 state-level data. American Journal of Managed Care 2003; 13-24.

38. Hart O, Moore J. Property rights and the nature of the firm. Journal of Political Economy 1990; 98; 1119-1158.

39. Hart S, Kurz M. Endogenous formation of coalitions. Econometrica 1983; 51; 1047-1064.

40. Jack W. Purchasing health care services from providers with unknown altruism. Journal of Health Economics 2005; 24; 73–93.

41. Jacobs P. The economics of health and medical care. Aspen: Gaithersburg, MD; 1991.

42. Jegers M, The economics of non profit accounting and auditing: suggestions for a research agenda. International Journal of Not-For-Profit Law 2002; 5.

43. Jensen M, Meckling W. Theory of the firm: managerial behaviour, agency costs and ownership structure. Journal of Finance 1976; 3; 305–360.

44. Laffont JJ, Tirole J. Theory of incentives in procurement and regulation. MIT Press: Cambridge Massachusetts; 1993.

45. Lee ML. A conspicuous production theory of hospital behaviour. Southern Economic Journal 1971; 28; 48-58.

46. Le Grand J. Motivation, agency, and public policy: of knights & knaves, pawns & queens. Oxford University Press: Oxford; 2003.

47. Lindsay CM, Buchanan JM. The organisation and financing of medical care in the United States. Health Service Financing. British Medical Association: London; 1970.

48. Long MF. Efficiency use of hospitals. In: Muskin S (Ed) The Economics of health and medical care. University of Michigan Press: Ann Arbor Michigan; 1964.

49. Ma AC. Health care payment systems: cost and quality incentives. Journal of Economics and Management Strategy 1994; 3; 93-112.

30

50. Mason A, Miraldo M, Siciliani L, Sivey P, Street A. Establishing a fair playing field for Payment by Results.Report to the Department of Health: London; 2007.

51. McAfee RP, McMillan J. Bidding rings. American Economic Review 1992; 82; 579-599.

52. McGuire A. The theory of the hospital: A review of the models. Social Science & Medicine 1985; 20; 1177–1185.

53. Meara E, White C, Cutler DM. Trends in medical spending by age: 1963-2000. Health Affairs 2004; 23; 176-183.

54. Miraldo M. Hospital financing. Master Thesis, DEA-UCL: Louvain-la-Neuve; 2000.

55. Mossialos E, Le Grand J. Health care and cost containment in the European Union. Aldeshot Brookfield: Ashgate; 1999.

56. Mossialos E. Funding health care: options for Europe. Open University Press, European Observatory on Health Care Systems series: Buckingham; 2002.

57. Mudambi R, Navarra P. Divisional power, intra-firm bargaining and rent-seeking behaviour in multidivisional corporations. Economics Bulletin 2004; 4; 1-10.

58. Muthoo A. Bargaining theory with applications. Cambridge University Press: Cambridge; 1999.

59. Newhouse JP. Toward a theory of nonprofit institutions: an economic model of a hospital. American Economic Review 1970; 60; 64–74.

60. Newhouse JP. Medical care costs: how much welfare loss?. Journal of Economic Perspectives 1992; 6; 3-21.

61. OECD Health Data 2007: Statistics and Indicators for 30 Countries. Organisation for Economic Co-operation and Development: Paris; 2007.

62. Osborne M, Rubinstein A. Bargaining and markets. Academic Press Inc: San Diego; 1990.

63. Pammolli F, Riccaboni M, Oglialoro C, Magazzini L, Baio G, Salerno N, Medical devices: competitiveness and impact on public health expenditure. CERM study for Directorate Enterprise of the European Commission: Rome; 2005.

64. Pauly MV, Redisch M. The non-for-profit hospital as a physicians cooperative. American Economic Review, 1973; 63; 87–99.

65. Putterman L. The economic nature of the firm: a reader. Cambridge University Press: New York; 1986.

66. Rajan RG, Zingales L. Power in a theory of the firm. Quarterly Journal of Economics 1998; 62; 387-432.

67. Rettig RA. Medical innovation duels cost containment. Health Affairs 1994; 13; 7-27.

68. Rice RG. Analysis of the hospital as an economic organism. Modern Hospital 1966; 106; 88-92.

69. Rickman N, McGuire A. Providers' reimbursement in the health care market, Scottish Journal of Political Economy 1999; 46; 53-71.

70. Rosko M, Broyles R. The economics of health care: a reference handbook. Greenwood Press: Westport, CT; 1988.

71. Scharfstein D, Stein JC. The dark side of internal capital markets: divisional rent-seeking and inefficient investment. Journal of Finance 2000; 55; 2537-2564.

72. Scitovski AA. Changes in the cost of treatment of selected illnesses, 1971-1981. Medical Care 1985; 23; 1345-1357.

73. Siciliani L. Selection of treatment under prospective payment systems in the hospital sector. Journal of Health Economics 2006; 25; 479-499.

74. Simoens S, Giuffrida A. The impact of physician payment methods on raising the efficiency of the healthcare system. Applied Health Economics and Health Policy 2004; 3; 39-46.

75. Smith A. An enquiry into the nature and causes of the wealth of nations. Black: Edinburgh; 1776.

76. Stole LA, Zweibel J. Organizational design and technology choice under intra-firm bargaining. American Economic Review 1996a; 86; 195-222.

77. Stole LA, J Zweibel. Intra-firm bargaining under non-binding contracts. Review of Economic Studies 1996b; 63; 375-410.

78. The Economist. Medicine goes digital. A special report on health care and technology. The Economist, 18th April 2009. London

79. Vidal-Puga JJ. A bargaining approach to the Owen value and the Nash solution with coalition structure, Economic Theory 2005; 25; 679-701.

80. Wang G. Bargaining over a menu of wage contracts. Review of Economic Studies 1998; 65; 295-305.

81. Weisbrod B. The health care quadrilemma: an essay on technological change, insurance, quality of care and cost containment. Journal of Economic Literature 1991; 523-552.

82. Willis-Owen CA, Brust K, Alsop H, Miraldo M, Cobb JP. Unicondylar knee arthroplasty in the UK National Health Service: an analysis of candidacy, outcome and cost efficacy. The Knee, 2009 Dec;16(6):473-8

83. WHO. Belgium – Health system profiles. European Observatory of Health Systems, World Health Organization: London; 2000.

84. Zhang X. The pure bargaining problem among coalitions. Asia-Pacific Journal of Operational Research 1995; 12; 1-15.

85. Zweifel P, Felder S, Meiers M. Ageing of the population and health care expenditure: a red herring? Health Economics 1999; 8; 485-496.

Appendix

A.1 Strategic bargaining

Define $\left(h_i^*[L], n_i^*[L] \right)$ as the *PSSPN*-equilibrium offer that L proposes whenever she is selected to make an offer, and $\delta V_i^*[L]$ as the (discounted) expected equilibrium continuation payoff to player L from accessing a further round of negotiations. Suppose *M* have been selected to make offers. The equilibrium payoff from rejecting any offer for D is $\delta V_i^*[D]$. Therefore, D will accept from M any proposal $\left(h_i[M], n_i[M] \right)$ s.t..

$\Pi_i[D]\left(h_i[M], n_i[M] \right) \geq \delta V_i^*[D]$, while rejecting any offer $\left(h_i[M], n_i[M] \right)$ s.t.

$\Pi_i[D]\left(h_i[M], n_i[M] \right) < \delta V_i^*[D]$. Note that we have taken advantage of the standard tie-breaking rule that if a player is indifferent between rejecting and accepting an offer, she accepts it. Furthermore, the property of no delay implies that in equilibrium, $\Pi_i[D]\left(h_i^*[M], n_i^*[M] \right) \geq \delta V_i^*[D]$. However, it is immediately reckoned that in equilibrium, it can never be that $\Pi_i[D]\left(h_i^*[M], n_i^*[M] \right) > \delta V_i^*[M]$ because otherwise, M could always increase their profit by offering a pair $\left(h_i'[M], n_i^*[M] \right)$, which would still be accepted, with $h_i'[M] < h_i^*[M]$, which improves their own payoff. Therefore, it must be that a managers' equilibrium offer $\left(h_i^*[M], n_i^*[M] \right)$

solves $\max \ \Pi_i[M]\left(h_i[M], n_i[M] \right) \ \ s.t. \ \Pi_i[D]\left(h_i[M], n_i[M] \right) = \delta V_i^*[D]$.

A necessary implication of this is that in equilibrium, managers choose a number of patients to be treated at which the *joint profits* of D and M are at the highest possible level. This must hold because any potentially attainable surplus share can a fortiori be reached by M when the total surplus to be divided between D and M is maximum. Indeed, the maximum of $\Pi_i[M]\left(h_i[M], n_i[M] \right)$ given the constraint $\Pi_i[D]\left(h_i[M], n_i[M] \right) = \delta V_i^*[D]$ is, by construction, equivalent to the maximum of the joint profit $\Pi_i[M]\left(h_i[M], n_i[M] \right) + \Pi_i[D]\left(h_i[M], n_i[M] \right)$ valued in $\Pi_i[D]\left(h_i[M], n_i[M] \right) = \delta V_i^*[D]$. Of course, such an argument assumes the second-order conditions are satisfied for the joint profit function, which we will shortly show to indeed be the case. Hence, we define the joint profit as $J\Pi \equiv \Pi_{Mi} + \Pi_{Di} = T_i - \dfrac{n_i^2}{2} - \dfrac{(n_i - K_i)^2}{2} + Q(n_i^*) + g_i n_i$ and denote n_i^* as the unique solution of $\partial J\Pi / \partial n_i = 0$, i.e. $(\alpha_i + 1/2) + g_i - n_i + K_i = 0$, namely,

$$n_i^* = \frac{(\alpha_i + \frac{1}{2}) + g_i + K_i}{3} \tag{A1.1}$$

Such solution turns out to be a global maximum of the joint profit function: in fact, the second-order condition is always verified as $\partial^2 J\Pi / \partial^2 n_i < 0$. Notice that n_i^* is increasing with case-mix α_i, i.e., given a fixed (g, K) pair, *type* 2 hospitals always treat a larger number of patients than *type* 1. Therefore, in equilibrium, it must be true that the managers always offer the doctors a proposed $n_i^*[M] \equiv n_i^*$ that maximises their joint profits.

Define $J\Pi[n_i^*] = T_i - \dfrac{(n_i^*)^2}{2} - \dfrac{(n_i^* - K_i)^2}{2} + Q(n_i^*) + g_i n_i^*$

the maximum attainable level of joint profits of doctors and managers, valued at the optimal number of admitted

patients n_i^*. Hence, the solution to the managers' programme is given by a equilibrium pair $\left(h_i^* [M], n_i^* [M] = n_i^* \right)$ such that:

$$\Pi_i [M] \left(h_i^* [M], n_i^* \right) = J\Pi \left[n_i^* \right] - \delta V_i^* [D] \qquad \text{(A1.2)}$$

However, if *doctors* have been selected to make offers, by a symmetric line of arguments, the equilibrium offer proposed by the doctors satisfies:

$$\Pi_i [M] \left(h_i^* [D], n_i^* [D] \right) = \delta V_i^* [M]$$

$$n_i^* [D] = n_i^* \qquad \text{(A1.3)}$$

$$\Pi_i [D] \left(h_i^* [D], n_i^* \right) = J\Pi \left[n_i^* \right] - \delta V_i^* [M]$$

The next step, then, is to look closer at the players' continuation payoffs $V_i^* [L]$, $L = M, D$. Consider the doctors, for instance. Whenever in equilibrium they access a further stage of negotiations, they expect to go through a new random selection of the agent entitled to make a proposal.

As shown above, a *PSSPN* equilibrium, at any node of the game at which they have been called to propose offers, doctors always propose an equilibrium offer $\left(h_i^* [D], n_i^* \right)$ that ensures a surplus of $\Pi_i [D] \left(h_i^* [D], n_i^* \right)$. In contrast, at any node of the game at which the managers have been selected to make proposals, they always offer a *PSSPN* equilibrium pair $\left(h_i^* [M], n_i^* \right)$, which, by construction, delivers to the doctors a payoff $\delta V_i^* [D]$. Therefore, we can immediately specify the doctors' continuation payoffs in a *PSSPN* equilibrium as $V_i^* [D] = \beta \Pi_i [D] \left(h_i^* [D], n_i^* \right) + (1 - \beta) \delta V_i^* [D]$ which, by direct substitution of $\Pi_i [D]$, is equivalent to:

$$V_i^* [D] = \beta \left[J\Pi \left[n_i^* \right] - \delta V_i^* [M] \right] + (1 - \beta) \delta V_i^* [D] = \frac{\beta J\Pi \left[n_i^* \right]}{1 - \delta (1 - \beta)} - \frac{\beta \delta V_i^* [M]}{1 - \delta (1 - \beta)} \quad \text{(A1.4)}$$

Analogous arguments allow characterising the managers' continuation payoffs in a *PSSPN* equilibrium as:

$$V_i^* [M] = \frac{(1 - \beta) J\Pi \left[n_i^* \right]}{1 - \beta \delta} - \frac{(1 - \beta) \delta V_i^* [D]}{1 - \beta \delta}$$

These equations clearly have a unique solution: in equilibrium, managers and doctors expect the continuation values:

$$V_i^* [M] = \frac{J\Pi \left[n_i^* \right] (1 - \beta)^2}{\beta (\delta + 1)(\beta - 1) + 1}, \quad V_i^* [D] = \frac{J\Pi \left[n_i^* \right] \beta^2}{\beta (\delta + 1)(\beta - 1) + 1}$$

The equilibrium continuation values depend on both β and δ. In the limit scenario, when the game converges to the case of perfectly patient agents, i.e., when $\delta \to 1$, the equation payoffs tend to: $V_i^* [M] = J\Pi \left[n_i^* \right] (1 - \beta)^2 / \left(\beta^2 + (1 - \beta)^2 \right)$ and $V_i^* [D] = J\Pi \left[n_i^* \right] \beta^2 / \left(\beta^2 + (1 - \beta)^2 \right)$, which, crucially, depend on the bargaining power β. When managers and doctors are equally likely to be selected to make offers, i.e., $\beta = 1/2$, they expect to split the surplus of the hospital equally, each obtaining $J\Pi \left[n_i^* \right] / 2$. However, when doctors (managers) all have the bargaining power, $\beta = 1$ ($\beta = 0$) doctors (managers) earn the whole hospital surplus. This last case represents a *take-it-or-leave-it* scenario.

34

Finally, the equilibrium *fee for service* offers can be easily worked out as the $h_i^*[L]$, $L = D, M$ that, given the above continuation values, solves the system:

$$\Pi_i[M]\left(h_i^*[M], n_i^*\right) = J\Pi\left[n_i^*\right] - \delta V_i^*[D]$$

$$\Pi_i[D]\left(h_i^*[D], n_i^*\right) = J\Pi\left[n_i^*\right] - \delta V_i^*[M]$$

This returns the two asymmetric equilibrium *fee for service* offers by managers and doctors:

$$h_i^*[M](\delta, \beta, \alpha_i; K_i, T_i, g_i) \neq h_i^*[D](\delta, \beta, \alpha_i; K_i, T_i, g_i)$$

with, $h_i^*[M] = \dfrac{3\beta\delta Z - 4H}{12(1 + 2(\alpha_i + g_i + K_i))}$ and $h_i^*[D] = \dfrac{3\delta(\beta - 1)Z - 1 + 4Q}{12(1 + 2(\alpha_i + g_i + K_i))}$ where

$$Z = 1 + 4\left[g_i(1 + g_i) + \alpha_i(1 + \alpha_i) + K_i\right] + 8\left[\alpha_i(g_i + K_i) + K_i(g_i - K_i)\right]$$
$$+ 24T_i$$

$$H = 1 + K_i + 2K_i(\alpha_i + g_i - K_i) + 4\left[\alpha_i(\alpha_i + 1) + g_i(g_i + 1)\right] + 8\alpha_i g_i$$

$$Q = \left[2(K_i - \alpha_i g_i) + 4K_i(\alpha_i + g_i + K_i) + 18T_i - g_i - \alpha_i(1 - \alpha_i)\right].$$

(A1.5)

In the limit case of perfectly patient agents, $\delta \to 1$, both $h^*(M)$ and $h^*(D)$ converge to the identical expression $\overline{h}_i^*$

$$h_i^*(M)_{\delta \to 1} \equiv h_i^*(D)_{\delta \to 1} \to \overline{h}_i^* = \dfrac{3\beta Z - 4H}{12(1 + 2(\alpha_i + g_i + K_i))}, \text{ where } Z \text{ and } H \text{ are as in (A.15).}$$

B Verification of Single Crossing conditions

Consider, for instance, the contracting scenario. Under perfect information, the optimal menu of contracts is given by (11). Let $v_{H_i} + T_i$ stand for the hospital indirect utility function. Given managers and doctors solutions,

$$v_{H_i}(K_i, g_i, T_i; \alpha_i) \equiv T_i - \dfrac{\left(n_i^*\right)^2}{2} - \dfrac{\left(n_i^* - K_i\right)^2}{2} + Q\left(n_i^*; \alpha_i\right) + g_i n_i^*.$$

Computing the total differential and the partial derivatives of $v_{H_i}(K_i, g_i, T_i; \alpha_i) + T_i$, it is possible to write the marginal rate of substitution between any pair of K_i, g_i and T_i. For instance, take g_i and T_i by the implicit function theorem $\dfrac{\partial(v_{H_i} + T_i)}{\partial T_i} dT_i + \dfrac{\partial(v_{H_i} + T_i)}{\partial g_i} dg_i = 0$, then:

$$\dfrac{dT_i}{dg_i} = -\dfrac{\partial\left(v_{H_i} + T_i\right)/\partial g_i}{\partial\left(v_{H_i} + T_i\right)/\partial T_i}$$

(B.1)

Analogously:

$$\dfrac{dT_i}{dK_i} = -\dfrac{\partial\left(v_{H_i} + T_i\right)/\partial K_i}{\partial\left(v_{H_i} + T_i\right)/\partial T_i}, \quad \dfrac{dK_i}{dg_i} = -\dfrac{\partial\left(v_{H_i} + T_i\right)/\partial g_i}{\partial\left(v_{H_i} + T_i\right)/\partial K_i}$$

(B.2)

Computing the partial derivatives,

$$\partial\left(v_{H_i}+T_i\right)/\partial g_i=\frac{g_i+\alpha_i+2K_i+1/2}{4},\quad \partial\left(v_{H_i}+T_i\right)/\partial T_i=1,\quad \partial\left(v_{H_i}+T_i\right)/\partial K_i=\frac{g_i+\alpha_i+1/2}{2}$$

Plugging these expressions into (B.1) and (B.2) and simplifying, we find:

$$\frac{dT_i}{dg_i}=-\frac{g_i+\alpha_i+2K_i+1/2}{4},\qquad \frac{dT_i}{dK_i}=-\frac{g_i+\alpha_i+1/2}{2},\qquad \frac{dK_i}{dg_i}=-\frac{2g_i+2\alpha_i+4K_i+1}{4\left(g_i+\alpha_i+1/2\right)}$$

Differentiating with respect to α_i,

$$\frac{d}{d\alpha_i}\frac{dT_i}{dK_i}\bigg|_{v_{H_i}+T_i}=-\frac{1}{2}\qquad \frac{d}{d\alpha_i}\frac{dT_i}{dg_i}\bigg|_{v_{H_i}+T_i}=-\frac{1}{4}$$

$$\frac{d}{d\alpha_i}\frac{dK_i}{dg_i}\bigg|_{v_{H_i}+T_i}=\frac{K_i}{\left(\alpha_i+1/2+2g_i\right)^2}>0$$

it turns out that the marginal rates of substitution are monotonic in α_i. Therefore, the Single Crossing Conditions are verified. The proof for the bargaining follows by analogous arguments.

C Comparison: bargaining vs. contract

C.1 Rent comparison
Substituting the optimal contract for the bargaining scenario given by (8) and (9) and the optimal contract for the contracting scenario as described in (11) and (12), we obtain the informational rents under both scenarios, R^b and R^c. Computing their difference:

$$R^b-R^c=\frac{\alpha_1-\alpha_2}{48}\left[r-\left(\alpha_2-\alpha_1\right)\frac{3\lambda\left(\lambda+4\right)+2}{12\left(\lambda+2\right)\left(1+\lambda\right)}\right]$$

As $\alpha_2>\alpha_1$ then $R^b\geq R^c$ if and only if $r\geq\left(\alpha_2-\alpha_1\right)\frac{3\lambda(\lambda+4)+2}{12(\lambda+2)(1+\lambda)}$.

C.2 Welfare comparison
Under asymmetric information, plugging the optimal government contract variables for the bargaining scenario ((8) and (9)), and for the contracting scenario ((11) and (12)) in for W^b and W^c, respectively (both given by (3)), we find that:

$$W^b-W^c=\frac{\alpha_2-\alpha_1}{48}\left[\left(\alpha_1-\alpha_2\right)\left(1+3\lambda\right)+2r\left(1+6\lambda\right)\right]$$

Solving $W^b-W^c=0$ we find that if $r\geq\dfrac{\left(\alpha_2-\alpha_1\right)\left(1+3\lambda\right)}{2\left(1+6\lambda\right)}$, then $W^b\geq W^c$.