Hospitals: Teaming up*

Pedro Pita Barros Pau Olivella

v.3.2 – 4 April 2009

Contents

1 Introduction **2**

2 Technology: what does a hospital do? **4**

3 Preferences: what does a hospital want? **10**
3.1 Models of hospital behavior 10
3.2 The separation of governance and control 12

4 Incentives within the hospital **17**
4.1 General elements 17
4.2 Group incentives 19

5 The hospital and the outside world **27**
5.1 Contracting with payers 27
5.2 Referrals 32

6 Some empirical evidence **35**
6.1 Technology 35
6.2 Does ownership matter? 39

7 Discussion and challenges for future research **41**

*Acknowledgements: We have benefited from the comments and suggestions of Kurt Brekke, Pablo Colorado, Paula González, Inés Macho-Sadler, Xavier Martinez-Giralt, Andrew Street, Artur Vaz, and especially from the editors, Sherry Glied and Peter Smith. The usual disclaimer applies.

> "[T]he failure to recognize that doctors and hospitals are linked by a strong bond of joint production is at the basis of many hospital inefficiencies." Harris (1977, p. 475)

1 Introduction

Hospitals are one of the main institutions operating in health care markets. It is of no surprise that considerable research has been devoted to understand how they work, both in the economics literature and in other fields. Previous summaries of the more important findings are reported in the Handbook of Health Economics (Culyer and Newhouse, 2000) and in the Elgar Companion to Health Economics (Jones, 2006).

The present chapter presents two distinctive features with respect to these other surveys. The first one is that we emphasize the separation of ownership and control in hospitals. In other words, the day-to-day actions and decisions taken within a hospital are often far from the direct control of the hospital's ruling body. As we will see, this view is not new, and can be traced back to the seminal work of Pauly and Redisch (1973), who views the hospital as a "physician's cooperative." Once this separation is recognized, performance may be equally related to ownership as to internal organizational features like individual as well as group incentives or hierarchical design. This observation paves the way to the other distinctive feature of this chapter, namely, to look at hospitals through the lens of teams. Hospitals are entities that have an internal organization characterized by team work. Hospitals also have to "team up" with special partners, namely payers of health care provided to patients and patient-referring agents.

Some countries rely mainly on private hospitals, while in other countries a majority of public hospitals is present. In the first group, we have the United States and the Netherlands, for example, while in the later group,

we find countries with national health services (such as the U.K., Italy and Spain, among others). However, one implication of our view of hospitals is that the existing differences of the relative prevalence of each possible ownership structure may in fact have less importance than that deemed in the received literature. Hence, rather than provide an exhaustive treatment of these differences, our approach allows us to focus on common – internal organization – aspects.

The chapter is structured as follows. We will dedicate the next two chapters to introduce the reader to the more classical approach to hospital, first as a producer and then as an agent endowed with preferences. Hence, in Section 2 we address technology issues like output measurement, as well as economies of scale and scope. In Section 3 we briefly overview different theories on how the ruling body of the hospital chooses between the different technologically feasible options, and on how its preferences over these options are formed. These theories are based on the usual observation that many hospitals are non-profit organizations, and propose that the hospital maximizes some other "utility function". Also in this chapter, we depart from this view and explain the notion of the separation of ownership and control, through the approaches of physician agency (Pauly and Redisch, 1973), and the hospital as an internal market (Harris, 1977). In the latter, equilibrium of demand and supply of resources is achieved without resorting to explicit prices. Rather, explicit as well as implicit contracts are used to govern decision making and resource allocation, i.e., to govern incentives. We will however argue that the procedure by which contracts themselves are negotiated and chosen within the hospital, amongst those that are efficient, has received little attention. These approaches motivate our interest on incentives within the hospital, which we develop in Section 4. We focus on the framework where two or more agents interact. There, the issues of observability (the measurement of each team member's contribution), team

incentives, and optimal hierarchical design become crucial. In Section 5 we show that the idea of team interactions can be extended beyond the limits of the hospital in two main areas. We first address the relationship of hospitals with third-party payers. In most countries, we find health insurance mechanisms, be they private or public, which assume the responsibility to pay for care provided to patients. Most models addressing payment rules assume a single payer. Often, this is not the case, and the degree of economies of scale may matter to the payment rule. Secondly, we notice that patients arrive at hospital referred by other agents, namely general practitioners (GP). Therefore, hospitals also have "to team" with GPs, and hospital preferences will determine how these two parties interact. In Section 6, we turn to the existing empirical evidence on two important issues discussed in the previous sections. The first issue is that of the optimal size of the hospital, which is intimately related to economies of scale and scope. The second issue is whether ownership really matters in reference to observed performance, where we update the evidence provided by Gaynor (2006). Finally, In Section 7 we discuss and propose challenges for future research.

2 Technology: what does a hospital do?

According to the neoclassical view, a hospital uses resources (inputs) ranging from medical equipment to pharmaceutical products, nurses and doctors. Combining those resources according to current knowledge and technology, patients are treated (output). The utilization of resources entails costs and different combinations of resources may be used to treat similar clinical conditions. Treating patients, on the other hand, leads to revenues. Therefore, even if several particularities do exist, hospitals can be studied using the traditional economic apparatus. Namely, issues about size and diversification advantages can be addressed. The only difficulty lies in the precise

definition and measurement of hospital's output. Let us address this latter issue first, as it is one of the most challenging.

To answer the simple question "what does a hospital do?", we need to observe the sequence of events that take place: The patient arrives at the hospital either by advice of some health professional (directed in from outpatient visits, referred by primary care organizations, etc.) or from the patient's own initiative (mainly, through a decision to go to the emergency room services). When the patient arrives at the hospital (s)he brings with him a certain initial condition and likely evolution of his health status, which may depend on its own characteristics. Then, the hospital (the doctor) makes an assessment of the situation and decides on how many resources to spend on the patient. The resources the doctor uses include her own time, nurses and other non-medical personnel, medical equipment and pharmaceutical products. The hospital must be organized in such a way that these resources are available when needed.[1] Sickness evolves, given the initial condition and the resources dedicated to treatment, and a final health outcome emerges. In this respect, the most distinguishing characteristic of a hospital is the objective of providing *acute* illness treatment. This implies that hospitals have to deliver a service tailored to the specific condition of each patient. In order to tailor treatment, the doctor gathers information about the patient's condition, processes it and takes a decision about the therapeutic path to be followed. This makes the operation of a hospital quite distinct from that of traditional business units, where, like in the treatment of chronic illness, needs are identified well in advance. We believe that the existing literature has not addressed this distinction in a complete way.

From this brief description, several notions of output may be constructed. However, almost all studies consider physical measures of activity as the hospital output. This is done for practical reasons. A closer look at the hospital

[1]The hospital as an inventory in the presence of uncertain demand is addressed below.

shows that a different approach to empirical modeling may be appropriate, when feasible. Still, more often than not empirical studies are forced to rely upon the simple output measures. The most popular (and easy to measure empirically) is the number of patients treated, and we may refine that by considering different types of output. A common approach is to divide output into tree activities: inpatient treatments (or discharges),[2] outpatient visits (that is, visits to specialists) and emergency room episodes. In line with these observations, the classic view of the hospital sees it, naturally, as a multiproduct unit.

We may also define hospital output according to the type of illness involved in the case of admissions. The need to set a measure of output, essential to compare performance of hospitals, has led to the creation and development of patient classification systems, like the Diagnosis Related Groups (DRG), which are homogeneous episodes of sickness treatment that group similar clinical situations demanding roughly the same amount of resources.[3] They also provide a tool to organize payments to hospitals. The information system provided by DRGs could even allow for a characterization of hospital outputs that takes each DRG as a different output, although the number of outputs resulting from such a classification would make a full economic analysis too complex.

The main advantage of these hospital output definitions is the ability to measure them, even in a crude way. However, a more demanding definition of the hospital's output is to see it as a change in the probability distribution over possible health states of the patient. This measure, in contrast to the proceeding ones, conveys more a notion of quality than of quantity, and follows directly from the path the patient takes within the hospital. Its

[2]Technological change has reduced the hospital stay for several interventions, which in some cases means patients return home in the same day of intervention.

[3]Different countries use different names for patient classification systems. We keep here the term DRG for expositional purposes.

main difficulty is with observation, as detailed information regarding the health status of the patient when leaving the hospital is usually unknown, except for the coarse distinction of live discharge versus death, which leads directly to the notion of mortality rate as an output measure. When the patient enters the hospital with a clinical condition, (s)he may receive care at the hospital. Despite the efforts and resources used, the patient may not survive. However, the resources used by the hospital were not useless, as they have increased the probability of survival of the patient compared to doing nothing. They changed the odds of survival. Thus, the contribution of the hospital is this change in survival probability. The death of the patient is a possible outcome even after committing resources to treatment. Aggregating over patients, the higher likelihood of survival will translate into a lower mortality rate.[4]

Besides mortality rates, whenever appropriate data is collected, output of the hospital can be defined as the expected health gain resulting from treatment. Additional QALYs (quality-adjusted life-years) as a measure of output, for example, contemplates both extensions in life and improvements in quality of life.

The survival probability is, of course, one of the dimensions of interest. Ideally, one would like to measure the difference in health status with and without treatment. This is a hard issue, requiring further developments on both data available and analytic tools. Castelli et al. (2007) propose an approach taking into account increases in life, weighted by quality, and adjusting for waiting time until intervention. The approach is demanding in terms of the information required, however.

Another characteristic of hospitals' operation is randomness of demand.

[4]This approach also has the advantage of interpreting death as an output equal to live discharge - they are just different realizations from the same probability distribution - thus avoiding the somewhat ad-hoc procedure of treating death within a hospital episode as an undesirable output (Dismuke and Sena, 2001; Yawe and Kavuma, 2007).

Daily, weekly and monthly variance in demand for hospital services is considerable. Since hospital care should not be denied to those in need, excess capacity, on the one hand, and rationing/waiting lists, on the other hand, accommodate variations in demand. The natural response to demand variance is to hold some excess capacity to satisfy peak demand (Joskow, 1980; Gal-Or, 1994; Friedman and Pauly 1981, Carey, 1998).

Let us now turn to the classical issues of optimal size and diversification advantages. Size advantages can be summarized on the simple notion of having a lower unit cost the larger the hospital is. This leads to the economic notion of economies of scale. In simple terms, a hospital has economies of scale if, by increasing its activity level by some scalar larger than one (e.g., doubling activity), costs increase less than proportionately (less than double). If costs increase in line with (more than) activity, with constant (increasing) average costs, the hospital has constant (decreasing) returns to scale.

Another important concept is economies of scope. A hospital has economies of scope whenever producing several outputs entails lower costs than producing each of them alone. In the case of hospital activity, there are economies of scope if, for example, having inpatient treatment and emergency room episodes in the same organization has lower costs than providing the same levels of care through independent organizations.

The two concepts, economies of scale and economies of scope, are relevant to delineate the size and range of services provided by hospitals. If economies of scale were never exhausted, one should observe only a few and very large hospitals. In other words, analysis of economies of scale provides guidance with regard to the optimal size of the hospital. Since we observe that hospitals of very different size do survive over time in the same geographic area, this could mean that economies of scale vanish at some level of activity.

Similarly, economies of scope express the advantages of combining different types of services. Real world observation suggests that emergency room, outpatient visits and inpatient activities have economies of scope, as they are regularly performed by the same organization. Perhaps emergency room has less pronounced economies of scope with respect to the other two activities, since we do sometimes observe hospitals with only inpatient care and outpatient visits. On the other hand, primary care activities and/or long-term recovery seem to have diseconomies of scale with respect to inpatient care as they tend to be performed by separate organizations.

As for the empirical measurement of economies of scale and scope, which we review in subsection 6.1, this is usually done by considering the three broad aggregates that we outlined above: total number of discharges, outpatient visits and emergency room episodes. An alternative could be to estimate economics of scale and/or scope by taking each DRG as a distinct output. This could be quite a difficult task since, typically, a patient classification system may have more than 500 DRGs. In contrast, when our interest lies in the detail within the hospital, then focusing at the DRG level may be a fruitful way to go.[5]

Economies of scale and economies of scope are not a choice of the hospital. They are determined by medical knowledge and technology. However, an important role exists for "management" choices on which resources to use and how to combine them to achieve the desired outcome (activity level). The choices made can be seen at the light of three concepts of efficiency. First, a good decision process leads to technological efficiency, defined as a situation where no waste of resources exists. That is, reducing the amount of any of the inputs used results in not achieving the desired level of activity. Different combinations of resources for a given objective may have this characteristic of being efficient from a technological viewpoint. It now

[5]The recent work of Olsen and Street (2008) is a good example of this approach.

comes the second layer of efficiency: from the technologically efficiency combinations of inputs, valuation of resources at their (opportunity) cost will reveal which combination(s) have the minimum cost to achieve the desired objective for hospital activity. The crucial element added in this step is the set of input prices. The third, and final, step is the definition of the desired level of activity (output). This implies defining the objective function of the hospital: what is valued and which trade-offs are present. The definition of the objective function of the hospital is addressed in more detail in the next section.

Let us finally stress one of our main messages. The literature has failed to fully tackle the main distinctive feature of a hospital relative to other health care providers: the ability to treat acute illness in a way that accounts for the uncertainty in the treatment process.

3 Preferences: what does a hospital want?

The previous section addressed the definition of hospital technology. The next step is to define what the hospital wants to do. It implies defining the objective function of the hospital, an issue intertwined with ownership. We look first at utility-maximizing hospitals in a general way, and then at the implications of different ownerships.

3.1 Models of hospital behavior

The objective function of the hospital can be defined in several different ways and existing literature reflects those possibilities. Departures from simple profit maximization have been justified in various manners, including altruism in physician preferences (taking physicians to be the crucial decision makers), reduced-form representation of interaction between hospital managers (who care about costs and revenues) and physicians (who care

about benefits to patients), etc. Making explicit the objective function of the hospital allows for determination of the optimal/desired output level.

A traditional discussion about choosing the "right" (desired) hospital output is associated with the role of quality and the trade-off between lower costs and higher quality. This trade-off is based on the assumption that higher quality implies more costs. This is likely to be so in efficient hospitals. However, inefficient hospitals may have room to improve simultaneously in both dimensions.

Quality is also an issue when comparing public and private (for-profit) ownership structures, with the usual presumption being that private hospitals will degrade (non-observable) quality of care to save on costs. This view, however, ignores the role that quality may have in attracting patients. Whenever, due to health insurance arrangements (be it voluntary, mandatory, social or public insurance) patients become relative insensitive to price but are still reacting to quality differences, then private (for profit) hospitals will use quality as a competitive tool. Quality in for-profit hospitals may then be higher than in public hospitals, namely in those dimensions that influence more consumers' decisions. Differences in objective functions are not neutral to the desired level of activity for the hospital. Whenever a trade-off between quality and quantity exists, different preferences generate distinct outcomes.

The quantity-quality model of Newhouse (1970) focuses on the role of hospital administrators. As Newhouse (1970) noted, even if the profit maximization objective is removed, the hospital still faces a tradeoff between quantity and quality. Indeed, for a given budget, greater quantities of care can only come at the expense of a decrease in quality of care, and viceversa. We refer to this tradeoff as "the budget constraint." He then proposes that the hospital will pursue some balance between quality within this budget constraint, and that this balance will be determined by the preferences of

the ruling body of the hospital, to which we will refer simply as "the board". Phelps (2002) offers us an overview of the determinants of these preferences. The basic idea is that a median voter exists in the board, that is, a voter whose preferences are such that an equal number of board members are on the "more quality and less quantity side" and on the "more quantity and less quality" side. The median voter will win any vote over the choice within the budget constraint.

A deeper question is where a particular composition of the board in terms of these preferences comes from. Phelps propose that this depends on some (exogenous) initial conditions (the rules of the game and perhaps the preferences of donors), and on how new board members are elected (again the rules of the game). If new board members are chosen by the board itself, the composition of board preferences should show some stability.

Several other models of hospital behavior can be found in the literature. For example, Weisbrod (1988) describes non-profit hospitals, assuming decision-makers that aim at both quality and quantity. They are not fundamentally different from for-profit hospitals, as technological efficiency and allocative efficiency are equally important to achieve.

A problem with these views is that any conflict of objectives between the board and the many different groups that actually take the decisions within the hospital (managers, clinicians, and so on) is assumed away. We now describe the different attempts in the literature to address this important issue.

3.2 The separation of governance and control

Pauly and Redisch (1973) are the first authors to offer us a different view of the hospital. They assume that clinicians, who take many of the important decisions, will disregard the objectives of the board and will maximize their

own welfare.[6] This is the view of the hospital as a "physicians cooperative". Inasmuch as physicians may be able to capture a proportion of the hospital's profits (receipts of all kinds minus operation costs), the two views may seem truly antithetical. Indeed, the budget balance constraint assumed by Newhouse is tantamount to a zero profit condition, while the physicians cooperative will strive at maximizing profits. However, the advertent reader will quickly realize that the two views in fact converge in their predictions if (i) hospitals compete against each other and (ii) free entry is guaranteed by the authorities. Indeed, long run competition among profit maximizing hospitals (or among any for profit organization for that matter) will drive profits down to zero. It is well known, however, that this conclusion breaks down in industries where large economies of scale and scope are present, and where entry is heavily regulated. Hospital care is one such industry. Moreover, the time horizon for the usual decision maker in a hospital (not to speak of emergency admissions) and that of the board could be very different. In a nutshell, budget considerations are most surely ignored by a clinician taking life and death decisions.

This debate is, in any case, enlightening. It casts doubt into the difference between profit maximizing and non profit organizations, and redirects the attention from board politics to the separation between "ownership" (in the sense of governance) and control in organizations. It also cast doubts on whether it is relevant to perform the usual distinction between publicly owned, independent non-profit organizations, and for-profit hospitals. To put it bluntly, why should a physician who is paid on a fee per service basis in a publicly owned hospital and a physician who is paid on the same basis in a non-profit hospital behave any different? It then all boils down to the

[6]One direct way to perhaps reconcile the two views is to assume that the median voter in the Board is a clinician. However, the classical view looks at choices within the budget constraint, while Pauly and Redisch take a different view, where the quantity/quality pair falls short of the budget and the difference is pocketed by the clinicians.

choice of incentive structure and organizational design.

In light of this, one is faced with the difficulty of making sense of the observed diversity with respect to ownership structure. Indeed, within and across countries we often find coexistence of the following three ownership structures: public, private for-profit and private not-for-profit, with distinct shares in each case. For example, Belgium and the Netherlands have a majority of private hospitals, France has mainly public hospitals (2/3 of the total), Denmark and Sweden mainly public hospitals but owned locally, while the UK has a majority of public hospitals owned by the National Health Service. Germany has a slight majority (53%) of public hospitals, but financing is done at the regional level and rules apply equally to all. For-profit hospitals can be owned by corporations or by cooperatives (associations) of doctors. Not-for-profit hospitals may be owned by foundations or by churches, for example. Public hospitals belong to Governments, though in some countries this occurs at the city/county level while in others hospitals are controlled by the central Government.

In any case, the central question is whether the performance of each of these structures is any different. We will address this empirical question in subsection 6.2. If there is a difference, is it due to differences in ownership structure or to the way in which the diverging objectives (physicians', nurses', managers') are aligned though implicit and explicit contracts?

A further advance concerning the separation between ownership and control within the hospital is that of Harris (1977). He introduces the idea that we should see the hospital as a non-market way of organizing demand and supply of a certain type of health care. On the demand side, we have medical staff whose objective is to acquire the necessary resources to treat patients. On the supply side, we have the administration who is in charge of the management decisions, making sure that these resources are available and also balancing them across the diverse internal demands within the

hospital. Another important contribution from Harris (1977) is the explicit recognition that a double hierarchy exists within the hospital, each one associated to one side of the market. More importantly, he argues that the equilibrium in this market is reached not by a price mechanism but by negotiation. This original view has had only a partial impact on the economic analysis of the hospital. We propose a possible explanation for this below.

Once the market ceases to be the mechanism to allocate resources, one needs to study what is the alternative procedure by which the actual players within the organization choose amongst the different options. In this respect, economists have distinguished between two phases. The first phase is the determination of the set of efficient alternatives. The second is the choice of alternatives within this set. Notice that, by the first fundamental theorem of welfare economics, both phases can be achieved, by price magic, in a completely decentralized fashion. However, when the first fundamental theorem is not applicable due to externalities, economies of scale and scope, or asymmetric information then one needs to resort to contracts and negotiation. In this sense, one could view Phelps' approach to board politics as one attempt to address this second phase of the analysis. However, notice that the rules of the game in the day-to-day interaction among hospital staff are much less clear than in the board. Defining such rules for each possible such interaction becomes a difficult task, given the specificity of each patient's case. Hence, on the spot negotiation and a general lack of commitment will dominate these interactions.

As for the first phase, the determination of optimal contracts in the hospital, agency and team theory have made important contributions, which we review in Section 4. On theory grounds, we identify there several attempts to address the specific nature of the hospital as a decision making unit.

In contrast, the second phase – how and which allocation is chosen among

the efficient – has received little attention. This is despite the fact that there exist well-established tools in cooperative game theory and in its non-cooperative underpinning (bargaining theory) that could in principle be applied. Why is it then that we seldom see this done? In our opinion, the fundamental reason is the fact that certain aspects of the environment can only be observed by one of the parties. For instance, consider a clinician asking a hospital administrator to purchase some high-tech piece of equipment. The administrator may be more knowledgeable about the financial costs, about whether this equipment is available in other hospitals in the area, or about the true magnitude of its cost offsets; whereas the doctor may have privileged information on the health benefits and diagnostic accuracy that the equipment would bring in for each particular patient. Unobservable differences in objectives (e.g., altruism, career concerns, and so on) among the negotiating parties would play the same role. Such a situation is what Game Theory refers to as an asymmetric information setting. Unfortunately, available theory (or we should say theories) provides very partial insights on how such a manager and clinician would (or should be allowed to) arrive to a decision. The reason is that current theories of bargaining under asymmetric information are plagued by the curse of multiple equilibria. In other words, even if one agrees that certain theory is the appropriate one, that single theory is in may case unable to pin down a unique prediction. What is worse, the particular set of equilibria that may be obtained depends on the bargaining procedure itself (who starts making offers, the existence of outside options, and so on).

4 Incentives within the hospital

4.1 General elements

Before we enter the discussion of incentives, let us point out the three most important factors that give rise to a (non-trivial) incentive problem in the hospital context. The first factor is the existence of a conflict of objectives.

Here are some examples that will allow us to set some terminology. As pointed out above, even if the (median voter of) the board may have decided on a particular balance between quantity and quality, the agents actually taking decisions may have conflicting objectives with that particular mandate. The doctor may be more inclined to either save on her own diagnostic effort or to improve the evidence about her skills by ordering extra tests, or by overstaffing. Similarly, hospital managers may want to secure good results by installing state of the art technical equipment that may be unneeded given the existence of similar resources nearby. In this examples, the board acts as "the principal" in the relationship, while we refer to doctors and managers as "agents". As another example, the institution that pays for hospital care but does not enjoy it directly (be the government, an insurance company, or even a donor) may desire to contain costs more than the board. In this case the board is the agent of the paying institution, who acts now as the principal. These two examples point out that the actual chain of command from the payer to the doctor is much longer than just a two-tier principal-agent relationship. Hence, the hierarchy will have several concatenated principal-agent relationships, where the agent of given pair is the principal of the next. The payer is the principal vis-á-vis the board and the board is the principal vis-á-vis the doctor. In fact, the actual structure of the hierarchy constitutes an important subject of study as will be seen below.

The second factor is the presence of asymmetric information. Asymmet-

ric information may arise for two reasons. One is the impossibility that the principal observe and scrutinize all the actions taken by its agents. Doctors –and not the board– decide over admission, testing and treatment while managers –and not the board– decide over equipment and resource inventories. The other is the impossibility that the principal perfectly observe the environment in which these actions were taken. For instance, the doctor may have privileged information on the present health conditions of the potential population of patients.

The third and final factor is the existence of risk. Risk blurs the link between what it is observable by the principal and either the actions of the agents or the environment. The health status of patient may worsen despite the fact that the correct course of action was taken by the doctor in charge. The average length of stay may increase because of reasons that lie beyond the responsibility of the team in charge of the floor. Although procedures exist to gain information about the agent's actions or the environment, these procedures will at most yield partial information. Examples of such procedures are external and internal audits and reports or inquests filled in by patients.

Such asymmetric information problems as well as how remuneration systems can palliate these problems, are issues treated elsewhere in this volume (on Agency, see chapter 26; and on Provider Payment and Incentives, see chapter 27). However, we will argue here that the hospital environment adds new aspects to the usual single-principal/single-agent relationship. Most of these aspects arise from our view of the hospital as a team (e.g., doctors and nurses, hospital administrators and specialists). This view naturally leads to issues like the availability of measures of each team member's performance, the convenience to provide group incentives, and hierarchical design (should some decisions be delegated? who contracts with whom?).

4.2 Group incentives

The first question that arises when looking at team performance and remuneration is with what accuracy can one observe each individual's contribution. This depends on the particular measure of performance that one uses. Take the example of a doctor and a nurse, an example that we will exploit further below. If one assesses the performance of the nurse as simply the amount of visits that the nurse makes to any given patient, then such performance measure could be deemed to be the nurse's individual responsibility. In contrast, if one looks at the difference between the patient's initial and final health status, then this will be a complex function of the efforts of both the doctor and the nurse. We will formalize this distinction below. Another issue is whether the pay of these two agents should be based on such complex signals, i.e., on whether each agent's pay should be linked to the other's actions and decisions. One more example illustrates this point. Suppose, as it is typicality in US hospitals, that the doctor is paid on a fee-for service basis whereas the nurse is salaried.[7] Even if the improvement in patient's health can be regarded as joint production, neither remuneration system is taking into account the fact that such improvement is in fact the result of the interaction between doctor and nurse. If instead one would reward these professionals on the basis of some measure of health improvement, we would indeed be observing team pay.

This example notwithstanding, it is true that team incentives are more the exception than the rule in the real world. This is despite the fact that paying individuals according to individual performance measures only may lead to uncooperative behavior and inefficiencies. To better understand where and when these inefficiencies arise, we provide a more formal analysis of this issue.

[7] In US hospitals, medical staff holds a loose contactual relationship with the hospital while nurses are hospital employees.

The theory of team incentives has dealt with two different frameworks. Whether we are in one or the other framework depends on whether individual signals carrying information on each and all parties' decisions are available or not. The simplest example of each in the context of a hospital would be the following. Suppose that one keeps track of most of the actions that a nurse and a physician have taken when treating an inpatient. This information would include number of bedside visits, medication dispensed, or even questionnaires filled in by the patient on his appraisal of the performance of each. We then would be in a context of individualized signals. Consider now a situation in which we know only the number of days before the patient is discharged, or his or her condition at the time of discharge. We then have a single signal that is the result of the interaction of nurse and physician.[8]

The issues involved in each of these two frameworks are very different. In the first framework one wonders whether the contract of one agent should be based not only on his own signal but also on the other agent's signal as well. In the second framework one is concerned about free riding. Notice also that whether we are in one or the other framework is not exogenous, as one can design protocols, monitoring and auditing, or even adjust the physical location of each player, in order to better distinguish one player's contribution from the other's.

As for the individualized signal's framework, the main insight is offered by Holmstrom (1982). It is only in very special circumstances that the third-party payer should *not* make one team member's contract depend on the other's performance. In fact, the needed mathematical property is somewhat obscure.[9] It includes, though, an intuitive case: the situation

[8]For an empirical analysys of the effects of financial rewards on team incentives in the collection of indirect tax in the UK, see Burgess et al. duggna (2007).

[9]The property is that team member A's signal (say signal a) be a sufficient statistic for team member B's signal (say signal b) when one is trying to infer A's actions. In that

where team members performances (signals) are completely independent. Notice, however, that such a restrictive assumption leads to an extremely degenerate case of a "team". It is very likely that useful indicators of a nurse's performance will indeed depend on the actions taken by the team's doctor. Take as a rather extreme example, the number of visits by the nurse will increase as a consequence of a shirking doctor.

In the absence of any informational asymmetries amongst team members and the third-party payer, the previous discussion implies that it is optimal to centralize contracts. We return to this issue in the last subsection. To sum up, we should see team incentives that are based on every team member's performance.

A second contribution of Holmstrom (1982) is to formalize the very intuitive but often forgotten idea that one should not punish a team member for some other member's inappropriate action. This is not a fairness issue only. Take the example of a team composed of an oncologist, who chooses her diagnostic effort, and a manager, who decides upon the quality of a computed tomography scanner to be installed. Suppose that if the manager installs a very low quality CT scanner, the performance measure of the doctor (say, length of stay) becomes independent of the doctor's diagnostic effort. Then the contract of the doctor should not depend on the length of stay. Suppose on the contrary that, if the manager has installed a high quality scanner, then the length of stay becomes a good signal of doctor's effort.[10] In other words, the probability of a long length of stay diminishes when the doctor exerts the needed level of care. In that case the doctor's contract should indeed depend on the length of stay.

---

case, A's contract is independent of B's signal. One says that a is a sufficient statistic for b if one can express the joint distribution of a and b as a product of two functions, one that depends only on a and b, and one that depends only on a and on A's actions.

[10] We are assuming that the advantages of that signal together with the improved efficacy of the scanner more than compensate the higher cost of the better quality scanner.

To sum up, it should be the case that the incentives contained in the doctor's contract depend on the manager's performance. The doctor's pay should depend heavily on the length of stay when there is indication that the manager performed well, whereas it should not depend on the length of stay otherwise. This not only provides the right incentives to the doctor, but also leads to the right risk-sharing between manager and doctor. Nothing in this discussion relates to fairness.

Admittedly, this points toward a quite sophisticated system of incentives. Indeed, it is also perhaps the case that one cannot directly observe the manager's actions and one needs to look at, say, running costs as a proxy for these actions. But these running costs surely depend on doctors' actions, and so on. Failure to recognize these interdependencies would lead to an incentive system that does not promote the right decisions, makes team members bear unnecessary compensation risks, and may even undermine overall morale.

As a *caveat* for the reader, one should not think only of monetary incentives. It has been shown that non-pecuniary rewards are as important: improved facilities and amenities, promotion, recognition by the other members of the team, and so on are good examples.[11] In general, the literature has converted these incentives into money equivalents, a simplification that is not free of criticism. It has been argued that psychological and behavioral factors like team morale or altruism could be undermined when financial incentives come in.[12]

Let us turn to the single team signal framework. Although there are ways to organize a hospital so that individual accountability to the board is

[11] Ratto et al. (2001) discuss how non-pecuaniary rewards may foster team incentives in the context of the National Health System in the UK.

[12] As Seabright (2002) poses it, "civic virtue may, on this view, be crowded out by the introduction of explicit incentives." See this work for a way to endogenize this result by means of a standard screening model under asymmetric information.

possible, there are some instances where it is just impossible to disentangle each team members' performance. Take for instance emergency admissions, where there is just no time to produce individual indicators. Perhaps only extremely imperfect indicators can be obtained ex-post. Then, the only signal available is the health condition of the patient in subsequent periods. How should the team members' performance be influenced? As mentioned above, it is the free-rider problem that is central in this context. The general literature has produced quite a negative result here: there exists no mechanism that solves the free-riding problem, unless some outside party "breaks the budget." In other words, if the team enjoys the net benefit of the relationship and shares it among its members, there does not exist a sharing rule that provides the right incentives. The idea is the following. In order to motivate each member of the team, it is necessary that, in case of underperformance, some group penalty be levied. This group penalty then implies a reduction in whatever net revenue has been accomplished. This in turn will decrease the shares of the net revenue enjoyed by each of the team members. Provided that the group penalty is sufficiently harsh, individuals will behave. This mitigates the free rider problem that originates when the full net benefit is to be shared. Take for example a health authority that provides incentives to a hospital by allocating extra funds. This health authority acts as the outside party that breaks the budget balance. If the hospital performs under some standard, it will not enjoy that extra fund. More importantly, the health authority keeps the undistributed funds. This serves as a commitment device. It is easy to see why it is difficult to sustain threats of group penalties that would be purely wasteful in the absence of such a third party. Once team performance is low, the team members have all the incentives to renegotiate the contract. Why self-inflict a wasteful punishment? However, if team members foresee such a renegotiation before deciding what to do, the free-rider problem reappears –full force.

The main lesson that springs from the previous discussion is that one should ensure that no such renegotiation is possible. Namely, one should keep distance between the payer (whoever bears the performance bonuses, be it the hospital manager, the insurer, or the Board of Trustees), and the team members. Put more bluntly, the payer should not be one more team member. He or she should benefit from any reduction in pay that is levied on the underperforming team.

Another possibility would be to introduce competition between teams. This would be tantamount of having the winning team collecting the bonus. Some experiments in this area have been performed, but never within the hospital, where each team provides such a different service from the other that any comparison becomes impossible. In other words, competition could be beneficial only if teams are whole hospitals that compete in providing a similar set of services to similar populations.

Doctors and nurses

One of the issues addressed in Chapter 22 is that of the shifting of skills between professions. Tavares (2005) analyses the particular case of the shift of responsibilities from doctors to nurses. Her main insight is that such a shift may in fact require a larger overall budget for incentive compensation. This author uses a model of the doctor-nurse team in which there is a single, overall indicator (patient is either cured or requires new treatment). The main intuition is that free-riding incentives on the part of doctors are reinforced if nurses' effort becomes more productive. She uses this model to discuss how failure by the health authorities to recognize this effect can lead to equilibrium actions within the team that not only harm overall welfare, but also may lead to resentment on the part of nurses. Therefore, she can also use this model to explain both the demand by doctors for higher pay and the demand by nurses of a recognition of their change in

status.

Centralization versus decentralization and the provision of incentives

When we consider the hospital as a cascade of contracts we are already assuming a special hierarchical form. Indeed, one could imagine a hospital where all contracting is centralized, so that a single decision maker – possibly the third-party payer – chooses the contracts of all team members. Abstracting from communication costs, several authors have studied the advantages of delegated contracting. Jelovac and Macho-Stadler (2002) propose a model of a hospital where two agents, let us refer to them as a manager and a doctor, decide on separate production inputs. Let us say that the first input is investment in equipment and the second one is diagnostic effort. These authors illustrate a subtle idea: when two agents interact, the contract for one of the agents can also be used to discipline the other. With this in mind, assume that the third-party payer can observe neither the manager's investment nor the doctor's effort. Consequently, an important piece of information is lost when all contracting is in the hands of the third-party payer. More specifically, the third-party payer cannot make the contract for the doctor contingent on the manager's investment decision, because the third-party payer cannot observe this decision. Hence, that contract, which could be used to govern the manager, loses this functionality. In contrast, if we allow the manager to design the doctor's contract, this specific function of the doctor's contract is restored. Of course, delegating contracting also has a disadvantage. The manager can manipulate the doctor's contract to his own benefit. When is it then that decentralization fares better from the third-party payer's point of view? Take the case when investment and diagnostic effort are complements of production. Then decentralizing is better, since the manager can, by providing incentives to the doctor, enhance his own incentives. Therefore, it all depends on the degree

of complementarity. For instance, it could be argued that, for treatments that are not high-tech, the doctor's productivity is quite independent of the hospital equipment. Centralization would be the better choice there.

Admittedly, and despite the theoretical appeal of the proceeding discussion, it not easy to find real world instances where these issues have been pondered in the context of hospital organization. In contrast, they have originated much debate in the managerial science literature.

Macho-Stadler and Perez-Castrillo (1998) study the delegation of contract design in the presence of collusion amongst team members. Although their analysis is very general, their insights can be readily applied to hospital management. Take the same example as before. If the third-party payer delegates to the manager the contracting of a specialist, this is equivalent to the third-party payer centralizing all contracts and then the manager and the specialist colluding in their decisions. Formally, if the third-party payer is to find the optimal coalition-proof set of contracts, this is equivalent to finding the best manager's contract from the point of view of the payer whilst taking into account that the manager will behave opportunistically when providing incentives to the doctor.

This has important implications for the organization of a hospital. Suppose that, in the absence of coalition formation, it was better to centralize contracts. Then, the policy of separating management and medical decisions (the double hierarchy, as seen by Harris, 1977) is indeed beneficial if such separation limits the extent to which managers and doctors can collude.

Finally, Boadway et al. (2004) also recognize the distinctive role that managers and doctors play in the hospital. The main issue addressed is the asymmetry of information: doctors have better information about patient needs (high versus low tech treatments and equipment) than hospital managers do. Managers, in turn, have better information on case mix than the payer (which is taken here to be a public authority, like the NHS). They

show that a cascade of contracts should be aimed at eliciting the private information of these agents at the least cost. The main implication is that high-tech equipment and high tech services are over-promoted in comparison to a situation where information is symmetric.

5 The hospital and the outside world

5.1 Contracting with payers

Recent decades have witnessed important changes in the way hospitals relate with payers of health care. The managed care trend, initiated in the early 1970s in the United States, and the movement toward activity based payments in many other countries have changed the way hospitals are financed (see chapter 27 in this volume for a discussion of provider payment rules). We take here a different view, looking at other issues, like the existence of multiple payers and the (potential) role of economies of scale in their presence.

The creation of DRGs has made possible the use of more sophisticated payment schedules. From simply invoicing payers for the health care provided, many hospitals around the world now have to negotiate in advance how payments will be made and how much it will receive for each type of care.

Third-party payers, be they commercial health insurance companies, sickness funds or national health services, faced with ever-increasing health costs look at ways to induce more efficient health care delivery. Over time, payment rules to hospitals became a tool to third-party payers to influence hospital behavior, not just a mere transfer of funds. It is worth pointing out the most important transformation that took place: payment mechanisms are now seen as a way to influence indirectly hospitals' organization. From a traditional cost reimbursement system, in which hospitals invoiced their

costs to third-party payers, payment mechanisms evolved to prospective payment, with prices set for treatment episodes. This change resulted from the recognition that cost reimbursement gives too little incentives for efficiency. Moving to prospective payments, characterized by more aggregate units of care provided (treatment episode, for example) leads to stronger incentives for efficiency. The downside of this change is that is also creates incentives for risk selection and for misclassification of patients.

The change in payment rules, and the existence of asymmetric information between payers and hospitals, opens the door for hospitals to "game" against payment schedules. The traditional, simple, relationship of cost invoicing of cost reimbursement has been replaced. The new relationships have now to balance two different objectives: provide incentives for efficiency of health care delivery within the hospital and avoid incentives for cream-skimming (risk selection) by hospitals. The basic trade-off that needs to be solved in the definition of payment schedules to hospitals has given rise to an extensive literature, which is reviewed in chapter 27.

Some related issues are less well understood: the role of economies of scale (and scope), and the role of large third-party payers, both of which are associated with an often neglected feature – hospitals can be funded by more than one payer. This is generally true in health systems where a multiplicity of health insurers coexist. Also in countries where a national health service is the preferred mode of health insurance, it is often the case that voluntary health insurance covering hospital care is present.

The existence of multiple payers leads to new issues in the definition of payment schedules, especially if one of the payers is significantly larger than the others. These large payers negotiate with hospitals, and are usually able to obtain different prices. Volume typically gains lower prices to the payer.

Suppose a hospital is already operating at a volume range where diseconomies of scale exist, after contracting an activity level with a third-party

payer. Then, another third-party payer wishes to contract with the hospital. The basic question is: how should the price be set?

In this simple sequential description, the price set for the second, smaller, third-party payer must reflect its impact on costs, and therefore, will exceed the average cost (per patient treated). If there are diseconomies of scale, the incremental average cost of extra patients treated will be higher than the average costs in the main contractual relationship. Under economies of scale, the reverse possibility occurs: the marginal third-party payer benefits from dilution of fixed costs, and the incremental average cost will be less. Therefore, for a constant batch of potential patients that it brings to the hospital, the exact payment schedule that the hospital is willing to accept can be quite different according to whether economies, or diseconomies, of scale prevail.

A related question is whether in the presence of a strong component of fixed costs, should the smaller third-party payers contribute to cover the fixed cost? How should this contribution be set? For example, one may think that smaller third-party payers should pay only the incremental cost that the care provided to their beneficiaries creates. Another approach would be to set a payment by each third-party payer equal to the average cost. The difference between the two alternatives may be significant. Firstly, in the presence of diseconomies of scale, having the smaller third-party payer paying according to the average cost shifts some cost to the larger payer. Secondly, in a more subtle way and following from the first effect, third-party payers may offer a more generous coverage to their beneficiaries and induce higher use of hospital services. Both effects suggest that marginal third-party payers should pay more than the average cost, though the question of how to determine the identity of the marginal payer is not easy to answer.

Another (disputed) effect in the relationship of the hospital with large payers is the cost shifting behavior. Cost shifting has been defined as the

practice of hospitals increasing prices to one group of third-party payers following price decrease associated with payments from another group enforced by some large payer, typically large Government-based programmes (like Medicare or Medicaid, in the US). The existence of cost-shifting behavior naturally requires the ability of the hospital to set prices for the care it provides. An initial explanation for cost shifting behavior required market power of hospitals and that prior to a change in prices paid by large payers such market power was not fully exploited by the hospital (although no rational is presented why market power was not exercised to full extent before, and why that changes). Morrisey (2003) addresses the issue from the perspective of market power as the cause of cost-shifting, to dismiss it as a relevant issue.[13] When faced with a price decrease forced by a large payer, hospitals face an alternative course of action, rather than cost shift: to look for savings resulting from economies of scale and/or from economies of scope, and to reduce inefficiencies.[14] Morrisey's looks at the role of non-profit organizations (and provision of charitable care) as the reason why market power is not fully exploited in the first place, just to discard it. Under Morrisey's argument, having other objectives than pure profit maximization does not imply that payers do not negotiate to obtain the lowest price possible from providers. Providers also have an advantage in diversifying payers, to avoid becoming dependent on a single one. Improvement in the mix of patients from different payers can also be attempted by providers, namely through carefully crafting services offered. Thus, other strategies available to hospitals are superior to cost-shifting. For public policy action, Morrisey (2003) proposes measures that increase market competition.

On the empirical side of cost-shifting behavior, most works find small effects of cost-shifting, and not particular pattern emerges. In some cases

[13] See Morrisey (1994) for an earlier view on the issue.

[14] Note that prospective payment mechanisms aim at reducing inefficiencies as well.

for-profit hospitals do not engage in cost-shifting, while for-profit do. In other cases, both type of hospitals do cost-shifting.[15] Cost shifting across different services is only possible if they are paid for by distinct third-party payers. More recently, Santerre (2005) performs an empirical analysis of cost shifting behavior, finding only a very minor efficiency loss, less than 1% of total private hospital expenditures (for 1992, US data). This corroborates an earlier discussion by Ginsburg (2003), who highlights that welfare costs are associated with distortions with inelastic demands tend to be small. Only distribution issues remain.

A related problem is addressed in Ma and McGuire (1993), where the way to pay for fixed costs (joint costs, in their terminology) by third-party payers is discussed.[16] Their analysis takes into account that the level of fixed costs is endogenous with payment structures in the case of hospitals. The way the payment rules are set will determine the incentives for investment and technology adoption.[17]

In this context, the possibility of a reimbursement transfer to pay for the fixed cost (called pass-through in Ma and McGuire's (1993) discussion) may improve overall efficiency. In particular, if one of the third-party payers has the opportunity to use cost reimbursement and commit first to its payment schedule, it may use the payment system strategically. By defining first its payment rule, it forces the other providers to cover the remaining costs of the provider. In doing so, it also increases aggregate social welfare, although some fixed cost share is shifted from the largest to the smallest third-party payer.

In their paper, Ma and McGuire (1993) identify the larger third-party

[15] See Clement (1999), Dranove (1988), Duggan (2002), Friesner and Rosenman (2004), Hadley and Feder (1985), Rosenman and Friesner (2002), Zwanziger et al. (2000).

[16] Further discussion of hospital decisions in the presence of cost shifting is done in Glazer and McGuire (1994).

[17] Typically, payers refrain from paying innovation and tend to shift the innovation adoption risk to providers.

payer with a public payer, the federal Medicare programme. Their most interesting result is that an efficiency allocation of resources (investment decisions) is achieved in equilibrium if and only if the large payer moves first and commits to a reimbursement level of the joint costs. This commitment mitigates the incentive the large payer has to shift costs to small payers (through the different prices the hospital has to charge each payer).

There are, of course, other reasons for differential treatment of third-party providers, namely variation in bargaining power, which may be linked to the size of demand each third-party payer brings to the hospital.[18]

The great majority of discussions about payment schedules to hospitals implicitly assume a single third-party payer or identical third-party payers. We argue here that variation in third-party payer size, in terms of patients they bring to the hospital, can easily lead to different payment schedules, with non-trivial impact on the design of such payment schedules. Elements such as the degree of economies of scale, interaction amongst third-party payers mediated by the need to finance the hospital and bargaining positions will create variation in payment schedules across third-party payers.

This is potentially relevant for all health systems, irrespective of their funding organization being based on tax-funded national health services, sickness funds or private health insurance. Further research is needed to fully detail how these considerations are to be incorporated into activity-based payment rules and other contractual approaches using DRGs as the basic unit of payment.

5.2 Referrals

Typically, the analysis of hospitals starts with patients showing up at the door of the hospital, and not much is said about demand for hospital services.

[18]For a summary of questions related to bargaining between hospitals and third-party payers, see Barros and Martinez-Giralt (2006) and the references therein.

Either it is assumed to be randomly determined from the point of view of the hospital or that demand for hospital services is determined by quality of care and prices, both time and monetary ones (including here the prices resulting from agreements between third-party payers and the hospital).

Future demand is discussed usually with reference to waiting lists for surgery, which nonetheless result from hospital decisions. This common approach neglects the role of the interactions between the hospital and other health care providers that may refer patients for hospital treatment.

In this respect, a common observation one hears from hospital managers, specialists, and policymakers is that a large fraction of cases they see could have been resolved in primary care.[19] This issue is closely related with the discussion on payment structures. Whenever hospitals are funded by case payments they prefer to receive more patients for treatment while hospitals funded by capitation (to treat people in a defined catchment area) will invest more in keeping patients treated at primary care level when clinically feasible. In other words, these professionals are demanding a larger role of primary care in gatekeeping. In fact, several health systems make referral by a GP a necessary condition to visit a specialist. This is indeed the case in Italy, the Netherlands, Norway, Spain, and the United Kingdom. In contrast, the gatekeeping role of the GP is very limited in Belgium, Finland, France and Germany. The issue of whether referral by a GP should be made compulsory or not has deserved quite a lot of attention recently. Gatekeeping determines to a considerable extent the demand faced by the hospital. Moreover, referrals to the hospital depend on both the incentives faced by GPs and on the formal relationship between primary care and hospitals.[20]

This is true not only for a national health system where hospitals are

[19]This concern was already voiced in the survey by Scott (2000), who mentioned the large differences observed in referral rates across GPs. See also Malcomson (2004).

[20]For the reader interested in the issue, see, for example, Brekke et al. (2007), Garcia-Marinoso and Jelovac (2003), and Gonzalez (2006).

public but also for managed care institutions. The later also tend to value primary care as a screening device before people get admitted for hospital treatment. Therefore, the scarcity of literature regarding how this relationship is established and what its implications are is surprising. For example, in national health services, primary care centers are run separately from hospitals, while in health maintenance organizations a more integrated view of hospital care and primary care exists.

In many circumstances, referral decisions are the only connection and entail no financial transfer. The fund-holding GP's experience in the UK introduced a financial element into this relationship (GPs that were fund holders had to pay for the care to their patients out of the budget they received). Whenever no money is involved, we often see calls for cooperation between physicians at the hospital and at the primary care centre. At the other extreme, we see managed care institutions vertically integrating the role of GPs in terms of referral to hospital care.

Some interesting questions arise in this relationship, since the decision to refer a patient to hospital treatment, and thus to create the hospital's demand, lies in the hands of the GP at primary care institutions. The relationship that is established between them will influence the referral flows and the overall efficiency of the health system.[21]

From the point of view of the hospital, a closer relationship with primary care may be more or less valued according to the way its financing is set.

The more aggregated way of paying a hospital with a well defined catchment area, namely within a national health service, is to pay by capitation – a fixed amount per year per member of the population in the relevant geographic area.[22]

[21] For further discussion on this, see Barros and Martinez-Giralt (2003) for a theoretical approach and Kripalani et al. (2007) for a review of existing literature on the relationship between primary care and hospitals.

[22] Of course, capitation payments can also be defined with reference to groups formed

Under this capitation system, the hospital has a clear incentive to detect health problems early and to treat patients at primary care facilities whenever feasible. On the other hand, if the hospital is paid mainly on the basis of how many patients it treats, it has the opposite incentive.

The balance between improving internal efficiency and improving referral decisions from primary care is clearly different to the hospital according to the way it is paid. Recent empirical research provides some information on the relevance of these mechanisms. Dusheiko et al. (2006) find that fundholding in the UK reduced referral to elective surgery. Thus, an integrated view of hospital and primary care incentives may prove useful to better understand how the demand for the hospital is actually determined. Still, further work on the issue is welcome.[23]

6 Some empirical evidence

6.1 Technology

Since data on hospitals is often routinely collected, it is of no surprise that many studies related to scale and scope economies in hospitals have been produced, as well as studies addressing productive efficiency, using several statistical techniques (the most popular ones being stochastic frontier analysis and DEA - Data Envelopment Analysis).[24]

Typically, concern with economies of scale (and/or scope) arises from the search for the most efficient scale for a hospital to operate at.

Knowledge about economies of scale and economies of scope is important to define the size of new hospitals, and range of services offered, when the hospital care is delivered directly by the Government (as in the case

in different ways, say, the beneficiaries of a given sickness fund or health insurer.

[23]Windmeijer et al. (2005) have one of the few studies that address the relationship between hospitals and general practice, though they do not address the role of referral.

[24]See Chapter 27 in this volume for a more in-depth review of efficiency studies. See also the general reviews in Hollingsworth (2003) and Jacobs et al. (2006).

in countries where a national health service exists) and to assess hospital mergers, in countries where hospital provision is mainly private. The literature on hospital cost structures attempts to define the optimal hospital size, with most estimates putting it in the range 200-300 beds. The discussion of the optimal size and scope of hospitals has somewhat faded away in recent years.

From the recent studies, Preyra and Pink (2006) estimate a short-run cost function and extend from it to the long-term cost function, finding an optimal value of about 180 beds, using data from Ontario, for the period 1994/1996. This estimate is lower than previous conventional wisdom suggested, though it is in line with a general trend toward shorter hospitalization spells, which require fewer beds.[25]

One of the main issues in measuring economies of scale (productivity, in general) in hospitals is the role of quality. More efficient hospitals are more likely to have a lower marginal cost of providing quality, and accordingly they may supply a higher quality level in equilibrium (which is likely to raise costs and mask their efficiency advantage).

The way we look at the quantity-efficiency trade-off pointed out earlier needs to be reformulated in the presence of inefficiency. An inefficient hospital may actually improve on both dimensions, quality and quantity, when it moves closer to the efficiency frontier.

The question of how the efficient scale and the degree of economies of scope change with quality has not been addressed in recent literature. The explicit consideration of quality choice by the hospital has other implications as well, since the competitive environment of the hospital will affect such choices.

On the issue of quality, the recent review by Gaynor (2006) makes clear

[25]For other recent evidence, see, among others, Dranove and Lindrooth (2003), Gaynor and Vogt (2003), Ho and Hamilton (2002), Ray (2003).

the ambiguities resulting from theory and concludes for the existence of mixed evidence. The regularities identified by Gaynor (2006) allow a distinction according to whether a regulated price regime or a free price regime prevails. In the first case, covering basically studies involving the US Medicare system, more competition seems to be associated with higher quality. In contrast, when considering privately insured patients and prices set by hospitals, no relationship between competition and quality exists. Both findings are in line with economic theory.

Under regulated prices, quality is the main "competitive tool" of hospitals and it is used intensively. Whenever both price and quality are available instruments to the hospital, the effort to attract patients is spread over both of them.[26] More competition may then be associated with either higher or lower quality, also depending on the intensity to which the hospital uses the price in its policy. Since the marginal gain from using price or quality to attract patients is also determined by the cost structure, not only is competition in the marketplace relevant, but also cost structures. Economies of scale will favour the use of the price vis-á-vis quality, for example.

In more recent research, the discussion on economies of scale has been set aside. Most empirical studies about hospital cost functions concentrate on efficiency measurement, in particular technological efficiency. They seldom report information on the optimal efficient scale of hospitals.

The issue of scale is being recovered somewhat at a different level. Given the increasing availability of micro data, exploration of scale effects becomes possible. An example is provided by Gaynor et al. (2005), who, in a simple but effective way distinguish between economies of scale and learning by doing as motives for a better performance (meaning a lower rate for in-hospital mortality). There is a growing literature on the link between higher

[26]This is true also when patients have insurance coverage, as most health insurers negotiate prices with providers.

volume and higher quality for many surgical procedures.

Scale also brings bargaining power gains. Burgess et al. (2005) and Melnick and Keeler (2007) find that a group of hospitals is in a position to charge higher prices. Since location of care provided is not a main force, it is bargaining power that is driving the results. Thus, scale brings gains, compared to single units, though related to prices rather than technological opportunities. These gains are not social gains as prices just distribute surplus between agents.

A final area where scale does matter is teaching. Several empirical works have addressed the extra costs of teaching activities, usually by inclusion of a dummy variable for teaching hospital status (capturing different levels in costs in teaching hospitals, compared to non-teaching ones) or by separating samples according to teaching/non-teaching status. This is hardly satisfactory, as it ignores how teaching activities impact the cost function. Several studies found higher costs for teaching hospitals.[27] Teaching increases costs because it diverts effort from doctors and leads to use of more resources, but it can also reduce costs because medical students can be a cheaper substitute for some medical labour work. A better understanding of how teaching hospitals operate should be sought in future research agendas.

The issue of random demand and its impact on hospital costs has received attention in several works.[28] More recently, Sharma et al. (2008) show that response to an unexpected demand surge for hospital services is more likely to be met by early discharges to free up capacity than to ration admissions.

[27] See, among others, Linna (1998), Grosskopf et al. (2001), Rosko (2001, 2004), Farsi and Filippini (2007).

[28] See Gaynor and Anderson (1995), Keeler and Ying (1996), Carey (1998) and Baker et al (2004).

6.2 Does ownership matter?

Most of the debate about ownership has been centred on empirical analysis of performance differences, namely cost efficiency. The main empirical hypothesis tested is based on the idea that as not-for-profit hospitals have a less well defined residual claimant,[29] we should expect them to be less efficient, although related problems are that for-profit hospitals may more easily induce demand and charge higher prices. Differences in several performance indicators have been addressed: financial results, quality variables and "upcoding" behavior. By "upcoding" we have in mind that one way for hospitals to play against the payment system is to code patients in a payment category higher than the correct one (in order to obtain a higher transfer from the third-party payer).

The legal distinctions associated with each ownership form are reviewed in Sloan (2000) and will not be repeated here. The early evidence points to no systematic difference in either cost efficiency or quality levels between for-profit and not-for-profit hospitals, though most literature on the issue is related to US experience. The ambiguity resulting from earlier studies has not been solved by recent works.

Several explanations have been advanced to explain the inconclusive comparisons across ownership forms, namely measurement issues (especially the questions associated with the appropriate definition of cost of capital in not-for-profit hospitals) and the hypothesis that market discipline is not strong enough to make for-profit hospitals sufficiently different from not-for-profit hospitals.

The ownership of a hospital is an issue only if it implies different objective functions or different constraints on behavior. With regard to the latter, it is often the case that not-for-profit organizations typically have

[29] On this see Section 4 on the consequences of the absence of a residual claimant.

constraints on how to distribute accounting surpluses. In particular, any surplus generated cannot be used to pay dividends, and it should, according to common provision in statutes, be re-invested in the activity (or in a closely related one). Surpluses may also be distributed to stakeholders (managers, employees, physicians, etc.), though usually not openly discussed as such.

The differences in objective functions may result from choice of boards. Whilst in for-profit institutions, shareholders select who runs the hospital, in not-for-profit institutions other means must be found, and self-perpetuation is easily achieved by entrenched directors. Moreover, exerting accountability is more difficult, presumably opening more room for managerial discretion. To control it, to some extent at least, compensation plans (involving stock options and bonuses) to executives in not-for-profit institutions are usually ruled out.

Distinct methodological approaches have been followed to make the comparisons, with unclear implications of the ambiguous results. Whether paired comparisons, regression analysis or stochastic frontier estimation is used, there is no clear supremacy of one ownership form over the other. Shen et al. (2005) conduct a meta-analysis for US hospitals (since 1990), finding that there are only minor differences between for-profit and not-for-profit hospitals.[30] Moreover, their analysis indicates that differences found in previous studies were attributed to methodological issues, such as functional forms, model specifications and definition of dependent variables in the empirical application.[31]

The debate about ownership has been less intense outside the US, although evidence has been gathered over the years in other countries. The

[30] See also Wong et al. (2005).

[31] See also, among others, Cremieux et al. (2005), Dafny (2005), Duggan (2000), Eggleston et al. (2008), Leone and van Horne (2005), Lindrooth and Weisbrod (2007), Picone et al. (2002), Rosenau and Linder (2003), Rosko (2001), Silverman and Skinner (2001), and Wright (2007).

focus has moved to management practice, with separation of ownership on the one side, and management rules on the other side. This is especially visible in countries with national health services that own and operate hospitals. Introduction of different governance and management rules, namely the use of trust or foundation status, is seen as a way to improve management efficiency without touching the usually delicate issue (at least, on political grounds) of ownership. Again, no strong differences in performance seem to be associated with ownership status.[32]

Sloan (2000) sets a research agenda about unsolved ownership issues: the role of other outputs associated with ownership (such as local control of the hospital), the need to have evidence from other countries, to look on differences other than performance (cost efficiency) alone, deeper knowledge of quality differences, better understanding of what constraints exist in each case, and how patient choice may affect the definition of ownership form. Except for more evidence from outside the US, little progress has been made in the other topics. The ambiguities regarding the comparison over ownership forms are still far from settled.

7 Discussion and challenges for future research

For a long time, economic analysis treated hospitals differently in theory and in empirical applications. In the former, issues of physician agency, contracting of hospital services and team production have been at the heart of many studies. In contrast, empirical studies have treated the hospital as a black box, addressing mostly issues about efficiency and implications of different ownership forms (Government-owned, private for-profit and private not-for-profit). Bridging the gap of theory to empirical work is a major challenge for current empirical research. A potential line of future developments

[32] See Marini et al. (2008) on the UK, Herr (2008) on Germany, Milcent (2005) on France, Barbetta et al. (2007) on Italy.

is to explore the empirical implications stemming from new theories on the internal organization of the hospital.

Even at a theoretical level, there are still many open questions. Many of them await the development of new tools. As mentioned in previous sections, a major shortcoming of present analyses is that the negotiation process that leads to the allocation of risk and expected surplus amongst the parties within the hospital has been overlooked. One advantage of focusing on hospitals in addressing this issue will be that technological aspects of the medical decision will constrain the set of possible bargaining procedures. As mentioned at the outset, Harris (1977) pointed out that medical care at the hospital has some peculiarities, urgency and specificity, that require a tailored decision-making process marked by a mixture of protocols and needed improvisation. Perhaps these peculiarities will guide the theorist in the specification of the "right" bargaining process.

Although it is true that the literature has studied the negotiation of the hospital with the third-party payer, the issue of how to incorporate economies of scale and scope (an therefore size), as well as the different objectives that ownership structure implies into these negotiations, has only recently received attention. Ownership issues alone, as envisaged by Sloan's (2000) agenda, present many opportunities for fruitful future research.

Still at the theoretical level, a thorough analysis of behavioral rather than fully rational and selfish decision making is needed. Issues of altruism have only recently been introduced in the neoclassical paradigm of selfish utility maximization.[33]

Finally, many studies have addressed the issue of hospital efficiency. Typically, efficiency scores are estimated and sometimes researchers attempt to find regularities between high efficiency scores and hospital characteristics.

[33]Biglaiser and Ma (forthcoming) and Ma and Rochet (2007) are examples of how one could integrate negotiation and altruism in the classical incentive model.

However, the internal organization of the hospital is not a common variable of interest, most likely due to obvious measurement problems. This may hinder development of a more thorough knowledge on how hospitals work.

One of the messages from our review is that hospital activities are better seen as performed by "teams". We interpret teams in a very broad sense, to include the joint effort of several economic agents, in order to achieve an outcome in the absence of a clear market mechanism (i.e. price) and where individual accountability is hard to obtain. If one considers the production of health and well-being of the system as a whole, this idea of team, or interaction, trascends the limits of the hospital. The interaction between acute treatment at the hospital and chronic treatment outside the hospital, and the interaction of primary care at the general practitioner and secondary care at the hospital specialist through referrals, are clear examples of this.

References

Baker, L., C. Phibbs, C. Guarmo, D. Supina and J. Reynolds, 2004, Within-year variation in hospital utilization and its implications for hospital costs, Journal of Health Economics, 23: 191–211.

Ballow, J. and B. Weisbrod, 2003, Managerial rewards and the behavior of for-profit, governmental and nonprofit organizations: evidence from the hospital industry, Journal of Public Economics, 87: 1895–1920.

Barbetta, G.P., G. Turati and A. Zago, 2007, Behavioral differences between public and private not-for-profit hospitals in the Italian National Health Service, Health Economics, 16: 75–96.

Barros, P.P. and X. Martinez-Giralt, 2006, Models of negotiation and bargaining in health care, pp. 242–249, in A. Jones, 2006, The Elgar Companion to Health Economics, Edward Elgar.

Barros, P.P. and X. Martinez-Giralt, 2003, Preventive health care and payment systems, Topics in Economic Analysis & Policy, Berkeley Electronic Press, 3(1): Article 10.

Biglaiser, G. and A. Ma, 2007, Moonlighting: Public Service and Private Practice, Rand Journal of Economics, 38(4): 1113-1133.

Boadway, R,. Marchand, M., and M. Sato, 2004, An optimal contract approach to hospital financing, Journal of Health Economics, 23(1): 85 - 110.

Brekke, K.R., Nuscheler, R. and Straume, O.R., 2007, Gatekeeping in Health Care, Journal of Health Economics, 26(1): 149-170.

Burgess, S., C., Propper, M., Ratto, S. Scholder, and E. Tominey, 2007, Smarter Task Assignment or Greater Effort: What Makes a Difference in Team Performance?, mimeo.

Burgess Jr, J.F., K. Carey and G. Young, 2005, The effect of network arrangements on hospital pricing behavior, Journal of Health Economics, 24(2): 391–405.

Carey, K., 1998, Stochastic demand for hospitals and optimizing "excessÓ bed capacity, Journal of Regulatory Economics, 14: 165–187.

Carey, K., 2000, A panel data design for estimation of hospital cost functions, Review of Economics and Statistics, 79(3): 443–453.

Castelli, A., D. Dawson, H. Gravelle, and A. Street, 2007, Improving the measurement of health system output growth, Health Economics, 16: 1091 – 1107.

Clement, J. , 1999, Dynamic cost shifting in hospitals: evidence from the 1980$\tilde{O}$s and the 1990$\tilde{O}$s, Inquiry, 34: 340–350.

Crémieux, P.-Y., P. Ouellette and F. Rimbaud, 2005, Hospital cost flexibility in the presence of many outputs: a public – private comparison, Health Care Management Science, 8(2): 111–120.

Culyer, A.J. and J.P. Newhouse, editors, 2000, Handbook of Health Economics, vol. I-II, Elsevier Science, B.V.

Dafny, L., 2005, How Do Hospitals Respond to Price Changes?, American Economic Review, 95(5): 1525–1547.

Dismuke, C. and V. Sena, 2001, Is there a trade-off between quality and productivity? the case of diagnostic technologies in Portugal, Annals of Operations Research, 107: 101–116.

Dranove, D., 1988, Pricing by non-profit institutions: the case of hospital cost shifting, Journal of Health Economics, 7(1): 44–57.

Dranove, D. and R. Lindrooth, 2003, Hospital consolidation and costs: another look at the evidence, Journal of Health Economics, 22(6): 983–997.

Duggan, M.G., 2002, Hospital market structure and the behavior of not-for-profit hospitals, evidence from responses to California's disproportionate share program, Rand Journal of Economics, 33: 433–446.

Duggan, M.G., 2000, Hospital ownership and public medical spending, Quarterly Journal of Economics, CXV(4): 1343–1373.

Dusheiko, M., H. Gravelle, R. Jacobs and P. Smith, 2006, The effect of financial incentive on gatekeeping doctors: evidence from a natural experiment, Journal of Health Economics, 25(3): 449–478.

Eggleston, K., Y-C Shen, J. Lau, C. Schmid and J. Chau, 2008, Hospital ownership and quality of care: what explains the different results in the literature, Health Economics, 17(12): 1345–1362.

Farsi, M. and Filippini, 2008, Effects of subsidization and teaching activities on hospital costs in Switzerland, Health Economics, 17(3): 335–350.

Friedman, B. and M. Pauly, 1981, Cost functions for a service firm with variable quality and stochastic demand, Review of Economics and Statistics, 63: 610–624.

Friesner, D. and R. Rosenman, 2004, Inpatient – outpatient cost shifting in Washington hospitals, Health Care Management Science, 7(1): 17–26.

Gal-Or, E., 1994, Excessive investment in hospital capacities, Journal of Economics and Management Strategy, 3(1): 53–70.

García-Marinoso, B. and I. Jelovac, 2003, GP's payment contracts and their referral practice, Journal of Health Economics, 22: 617–635.

Gaynor, M., 2006, Competition and quality in health care markets, Foundations and Trends in Microeconomics, Vol 2(6), NowPublishers.

Gaynor, M. and G. Anderson, 1995, Uncertain demand, the structure of hospital costs and the cost of empty beds, Journal of Health Economics, 14: 291–317.

Gaynor, M., H. Seider and W. Vogt, 2005, The value-outcome effect, scale economies, and learning-by-doing, American Economic Review, 95(2): 243–247.

Gaynor, M. and W. Vogt, 2003, Competition among hospitals, Rand Journal of Economics, 34(4): 764–785.

Gerdtham, U.-G. and Bengt Jonsson, 2000, International comparisons of health expenditure: theory, data and econometric analysis, chapter 2, pp. 12–53, in A.J. Culyer and J.P. Newhouse, editors, Handbook of Health Economics, vol. I, Elsevier Science, B.V.

Ginsburg, P., 2003, Can hospitals and physicians shift the effects of cuts in Medicare reimbursement to private payers?, Health Affairs, W3: 472–479.

Glazer, J. and T.G. McGuire, 1994, Payer competition and cost-shifting in health care, Journal of Economics and Management Strategy, 3(1): 71–92.

González, P., 2006, The Gatekeeping Role of General Practitioners. Does Patients' Information Matter? Working paper ECON 06.09, Universidad Pablo de Olavide.

Grosskopf, S., D. Margaritis and V. Valdmanis, 2001, Comparing teaching and non teaching hospitals: a frontier approach, Health Care Management Science, 4(1): 83–90.

Hadley, J. and J. Feder, 1985, Hospital cost-shifting and care for the uninsured, Health Affairs, 4: 67–80.

Harris, J., 1977, The internal organization of hospitals: some economic implications, The Bell Journal of Economics, 8(2): 467–482.

Herr, A., 2008, Cost and technical efficiency of German hospitals: does ownership matter?, Health Economics, 17: 1057–1071.

Ho, V. and B. Hamilton, 2002, Hospital mergers and acquisitions: does market consolidation harm patients? Journal of Health Economics, 2000, 19(5): 767–791.

Holmstrom, B., 1982, Moral Hazard in Teams, Bell Journal of Economics, 13: 324–340.

Hollingsworth, B., 2003, Non-parametric and parametric applications measuring efficiency in health care, Health Care Management Science, 6(4): 203–218.

Jacobs, R., P. Smith and A. Street, 2006, Measuring efficiency in health care – analytic techniques and health policy, Cambridge University Press.

Jelovac, I.and I. Macho-Stadler, 2002, Comparing Organizational Structures in Health Services, Journal of Economic Behavior and Organization, 49: 501–522.

Jones, A., editor, 2006, The Elgar Companion to Health Economics, Edward Elgar.

Joskow, P., 1980, The effects of competition and regulation on hospital bed supply and the reservation quality of the hospital, Bell Journal of Economics, 11: 42–447.

Keeler, T. E. and J. Ying, 1996, Hospital costs and excess bed capacity: a statistical analysis, Review of Economics and Statisticas, 78: 470–481.

Kripalani, S., F. LeFevre, C. Phillips, M. Williams, P. Basanah, D. Baker, 2007, Deficits in communication and information transfer between hospital-based and primary care physicians - implications for patient safety and continuity of care, Journal of the American Medical Association, 297(8): 831–841.

Leone, A. and R. van Horne, 2005, How do nonprofit hospitals manage earnings?, Journal of Health Economics, 24(4): 815–837.

Linna, M., 1998, Measuring hospital cost efficiency with panel data models, Health Economics, 7(5): 415–427.

Lindrooth, R., A. LaSasso and G. Bazzoli, 2003, The effect of urban hospital closure on markets, Journal of Health Economics, 22(5): 692–672.

Lindrooth, R. and B. Weisbrod, 2007, Do religious nonprofit and for-profit organizations respond differently to financial incentives? The hospice industry, Journal of Health Economics, 26(2): 342 – 357.

Ma, A. and J.-C. Rochet, ??? to be completed ?? check line 1408

Ma, A. and T. McGuire, 1993, Paying for joint costs in health care, Journal of Economics and Management Strategy, 2(1): 71–95.

Macho-Stadler, I. and D. Pérez-Castrillo, 1998, Centralized and decentralized contracts in a moral hazard environment, The Journal of Industrial Economics, 46(4): 489–510.

Malcomson, J.M., 2004, Health Service Gatekeepers, RAND Journal of Economics 35: 401–421.

Marini, G., M. Miraldo, R. Jacobs, M. Goddard, 2008, Giving greater financial independence to hospitals: does it make a difference? the case of the English NHS trusts, Health Economics, 17(6): 751–775.

McKay, N., M. Deily and F. Dorner, 2003, Ownership and changes in hospital inefficiency 1986 – 1991, Inquiry, 39: 388–389.

Melnick. G. and E. Keeler, 2007, Journal of Health Economics, 26(2): 400–412.

Milcent, C., 2005, Hospital ownership, reimbursement systems and mortality rates, Health Economics, 14(11): 1151–1168.

Morrisey, M., 2003, Cost-shifting: new myths, old confusion, and enduring reality, Health Affairs, W3: 489–491.

Morrisey, M., 1994, Cost-shifting in health care: separating evidence from rethoric, The American Enterprise Institute Press, Washington.

Newhouse, J.P., 1970, Toward a theory of nonprofit institutions: an economic model of a hospital, American Economic Review, 60(1): 64–74.

Olsen, K. and A. Street, 2008, The analysis of efficiency among a small number of organizations: how inferences can be improved by exploiting patient-level data, Health Economics, 17(6): 671-81.

Pauly, M. and M. Redisch, 1973, The Not-For-Profit Hospital as a Physicians' Cooperative, American Economic Review, 63(1): 87–99.

Phelps, C.E., 2002, Health Economics, Addison-Wesley.

Picone, G., C. Shin-Yi, and F. Sloan, 2002, Are for-profit conversions harmful to patients and to Medicare? Rand Journal of Economics, 33(3): 507–523.

Preyra, C. and G. Pink, 2006, Scale and scope economies through hospital consolidations, Journal of Health Economics, 25(6): 1049–1068.

Ratto, M., S. Burgess, B. Croxson, I. Jewitt, and C.Propper, 2001, Team-Based Incentives in the NHS: An Economic Analysis, CMPO Working Paper Series No. 01/37.

Ray, S., 2003, Measuring scale efficiency from the translog multi-input, multi-output distance function, Dept. of Economics, University of Connecticut.

Rosenau, P. and S. Linder, 2003, Two decades of research comparing for-profit and nonprofit health provider performance in the United States, Social Science Quarterly, 84(2): 219–241

Rosenman, R., and D. Friesner, 2002, Cost shifting revisited: the case of service intensity, Health Care Management Science, 5(1): 15–24.

Rosko, M., 2004, Performance of US teaching hospitals: a panel analysis of cost inefficiency, Health Care Management Science, 7(1): 7–16.

Rosko, M., 2001, Cost efficiency of U.S. hospitals: a stochastic frontier approach, Health Economics, 10(6): 539–551.

Santerre, R.E., 2005, The welfare loss from hospital cost-shifting behavior: a partial equilibrium analysis, Health Economics, 14(6): 621–626.

Scott, A., 2000, The economics of general practice, chapter 22, pp. 1176–1202, in A.J. Culyer and J.P. Newhouse, editors, Handbook of Health Economics, vol. II, Elsevier Science, B.V.

Seabright, P. 2002, Blood, Bribes and the Crowding-Out of Altruism by Financial Incentives, mimeo IDEI, Université de Toulouse-1.

Sharma, R., M. Stano and R. Gehring, 2008, Short-term fluctuations in hospital demand: implications for admissions, discharge, and discriminatory behavior, Rand Journal of Economics, 39(2): 586–606.

Shen, Y.-C., K. Eggleston, J. Lau and C. Schmid, 2005, Hospital ownership and financial performance: a quantitative research review, NBER Working Paper 11662.

Silverman, E. and J. Skinner, 2001, Are for-profit hospitals really different? Medicare upcoding and market structure, NBER Working Paper 8133.

Sloan, F., 2000, Not-for-profit ownership and hospital behavior, chapter 21, pp. 1142–1174, in A.J. Culyer and J.P. Newhouse, editors, Handbook of Health Economics, vol. I, Elsevier Science, B.V.

Tavares, A.I., 2005, The Change of the Doctor-Nurse Paradigm: Some Conditions and Implications, mimeo, Department of Economics, Universitat Autònoma de Barcelona.

Weisbrod, B., 1988, The nonprofit economy, Harvard University Press.

Windmeijer, F., H. Gravelle, P. Hoonhout, 2005, Waiting lists, waiting times and admissions: an empirical analysis at hospital and general practice level, Health Economics, 14(19): 971–985.

Wong, H., C. Zhan and R. Mutter, 2005, Do different measures of hospital competition matter in empirical investigations of hospital behavior?, Review of Industrial Organization, 26(1): 61–87.

Wright, D., 2007, Specialist payment schemes and patient selection on private and public hospitals, Journal of Health Economics, 26(5): 1014–1026.

Yawe, B. and S. Kavuma, 2007, Technical efficiency in the presence of desirable and undesirable outputs: a case study of selected district referral hospitals in Uganda, mimeo.

Zwanziger, J., G. Melnick and A. Banezai, 2000, Can cost shifting continue on a price-competitive environment?, Health Economics, 211–225.