THE UNIVERSITY of York

CENTRE FOR HEALTH ECONOMICS

Efficiency of Health Care Production in Different Hierarchically Structured Hospitals

Antonio Giuffrida
Fabrizia Lapecorella
Giacomo Pignataro

CHE Technical Paper Series 14

Efficiency of Health Care Production in Different Hierarchically Structured Hospitals

Antonio Giuffrida[a,*], *Fabrizia Lapecorella*[b], *Giacomo Pignataro*[c]

[a] Centre for Health Economics, University of York, UK
[b] Facoltà di Economia, Università di Bari, Italy
[c] Facoltà di Economia, Università di Catania, Italy

Abstract

The issue addressed in this paper is how efficiency of health care production in hospitals is affected by different organisational structures. More precisely, we consider *local public hospitals* integrated into a Health Authority, which also performs as purchaser of services in the NHS and *hospital trusts*, separated from the Health Authority. Is it true that, under asymmetric information, the need to provide doctors with appropriate incentives implies a loss of productive efficiency related to the different information structures that characterise the different organisations of supply? Using an agency model, we discuss how incentives to exploit exogenous informational advantages affect productive decisions taken by consultants in hospital trusts and in local public hospitals. The main finding is that in local public hospitals the optimal (second best) level of output is lower than in hospital trusts. We, then, employ data from an Italian region to test the theoretical results of the model. The productive efficiency of the different hospitals is estimated, using the Data Envelopment Analysis. The empirical analysis confirms that hospital trusts have, on average, higher efficiency rating than local public hospitals, with the difference in performance mainly due to the pure technical component, while the two groups do not show significant differences in scale efficiency. The results of the paper have important implications in designing the supply of health services in the NHS.

---

[*] Correspondence to: Antonio Giuffrida, Centre for Health Economics, University of York, YO10 5DD, York, UK. Tel: +44-1904-433718; e-mail ag19@york.ac.uk.
Contribution: A. Giuffrida, F. Lapecorella and G. Pignataro developed the study and wrote the paper jointly. However sections 5, 6 and 7 ought to be attributed to A. Giuffrida; sections 2 and 3.1 to F. Lapecorella; sections 3.1, 3.2 and 4 to G. Pignataro.
We wish to thank the Servizio Osservatorio Epidemiologico of Regione Lombardia - D. G. Sanita' – and Carlo Zocchetti for their kindness in providing the data used in the empirical section of this paper.

1. Introduction

Health care systems have been subject to many reforms in the last decade in many countries, urged by the financial concern about the growth of health expenditures in the eighties. In those countries with a predominant public sector control on the supply of health services, along with financial measures increasing the private share in total spending, it has also been planned a structural move towards competition in the provision of major services. The search for competition has been mainly justified by the attempt to reduce total health expenditures in the long run through incentives to cost containment in the production of health services as arising from competition. Competition has also been thought of as useful to induce an improvement in the quality of the services provided by producers. An appropriate implementation of a competitive system, however, has required structural changes in the organisation of the supply of health care, above all in those countries in which it was publicly provided. Usually, these countries were characterised by an integration of the funding and provision functions under the control of central and/or local governments. Therefore, the split between these two functions has represented one of the major reforms carried out in many countries. The idea behind this reform is that producers are made independent from those who pay for the services they provide, and with appropriate payment systems, they have an incentive to contain costs and to compete for patients. However, the separation has never been complete and it is, in some cases, under re-consideration. Italy and UK surely represent two cases whereas, even with significant differences, Local Health Authorities have retained production functions, alongside their purchasing role. Moreover, countries like UK and Sweden are trying to remodel their health care system moving from simply market based mechanisms of resources allocation, to look for an increasing role for planning and "co-operation" among the different actors of the system, without repudiating the previous reforms. Obviously, this implies a shift in the organisation of supply, from separation to integration. Therefore, it looks like there is an unresolved choice between integration and separation of different functions, which deserves to be analysed for a better understanding of their costs and benefits.

The purpose of this paper is to compare the different performance of an integrated organisation of supply of health services with a separated one. First of all, we will try to provide the comparison with a proper theoretical basis, which can be useful as a basis for empirical analysis. There are different theoretical approaches to the analysis of integrated and separated supply of health care. Recently, Halonen and Propper (1999) have approached this issue considering why self-interested politicians should introduce reforms of the public sector replacing a single public supplier, as it is the case when Local Health Authorities concentrate production and financing functions, with competing public suppliers, as it happens when the purchaser and the provider roles are separated. Our interest, however, is different since we want to look at the effects of the two structures of supply, whatever the reasons why they are implemented. The approach generally used for the analysis of this problem is based on the property rights theory. The results that can be drawn for the case of health services are reviewed in section 2. This approach, however, overlooks how the internal incentive problems of the production units (e.g. hospitals) are tackled within the different organisations of supply and, therefore, how integration/separation affects the performance of these units. In section 3, we present an agency model trying to study the different incentive problems of hospitals, separated by the purchaser of their services, referred to as Hospital Trusts, and of Local Public Hospitals, which are instead integrated into the purchasing authority. An empirical test of the main result of the model is carried out using Italian data. The Italian National Health Service, whose main features are briefly summarised in section 4, is an interesting case study, since it maintains a "mixed" presence of integrated and separated providers of services, especially as far as hospitals are concerned. We, therefore, use Data Envelopment Analysis, whose main characteristics are presented in section 5, to provide an empirical estimate of technical efficiency of the two types of hospitals studied in the theoretical model, as they stand in the Italian Health Service. The data used in the empirical analysis are shown in section 6, and they are related to all public hospitals operating in one Italian Region, Lombardy, for the year 1996. The results of the measurement of technical efficiency are presented in section 7, and the main conclusions of the paper are discussed in the final section.

2. Separation of functions and information structure

The issue of the separation of functions within the Health Service can be dealt with through the approach of the property rights theory. One of the basic components of this approach is the observation of contractual incompleteness, that is the existence of transaction costs will cause contracts to miss a full provision for all future contingencies. When there is need for renegotiations, because something happens not accounted for in the original contract, parties can have an opportunistic behaviour, above all if they are bounded in their trade. This latter circumstance is usually identified in the existence of the so called "ex-ante relationship-specific investments", that is "a prior investment, which creates value if the parties' economic relationship extends over time, but does not if the parties split up" (Hart 1996, 26). The economic implication is that there will be underinvestment for each party, so as to limit the potential opportunistic behaviour of the counterpart. Therefore integration, allowing for the exercise of residual control rights, gives an incentive for the realisation of a more efficient amount of relationship-specific investments. The property rights approach is then used, in industrial economics, to give important insights for determining the boundaries of firms.

This approach, however, has also been used to study the problem of integration/separation of purchasers and providers of health services[1]. The benefits of integrating the two functions depend on the degree of complementarity between the different assets. Belli (1997), for instance, uses the property rights approach to discuss the quasi-markets reform in Britain and argues that the assets bought by different segments of health care provision, even if not technical complements, can be regarded

---

[1] See Belli (1997) and Clerico (1997).

as such, since they are used for strictly complementary activities[2]. Whatever the recommendations arising from the application of the property rights approach to the problem at hands, it is clear that it can provide us with answers to the question of the optimality of an integrated /separated organisation of supply of health services. The theory treats the different units to be integrated as single entities, overlooking their internal incentive problems, arising from their specific information structure. It does not say anything about how the specific problems of each single unit are tackled within different organisations. In other words, if we consider the relative benefits of integrating a hospital with a health authority, we must consider not only the degree of complementarity of the different activities, but also how the internal incentives of the different agents operating within the hospital are changed by integration. It is clear that the integration of a hospital into (or its separation from) a health authority affects the decision-making hierarchy and, consequently, the information structure of the relevant relationships.

Now, if we focus on hospitals, there are some peculiarities related to their activity. Their managers are supposed to take the main decisions on the production to be carried out and on the inputs to be used, but these decisions, however, can only be partially controlled by managers since they are strongly influenced by doctors. The analysis of the role of doctors in the production of health care, not only within hospitals, has been one of the chief concerns of health economists and one of the widely used approaches is the theory of agency[3]. The relationship that has attracted the attention of many scholars is the one between doctors and patients. Doctors are regarded as informed agents of patients and, as long as they may pursue their own objectives, they will end up choosing treatments that are different from those which would have been chosen by patients, were they had the same information as doctors. Even if these models provide with an interpretation of what is called the "supplier induced demand", the welfare analysis of these results is not always clear cut, because such an analysis should also take the effectiveness of treatments into account[4]. Apart from these problems, however, the analysis of variations of informational problems within hospitals in different organisational structures requires to understand how the relationship between doctors and managers. It is crucial, in other words, to study how the different information structures, characterising integrated/separated hospitals, can impinge on the conflict over the use of inputs and the levels of output, which can arise between doctors and managers.

Here again, the theory of agency can provide useful insights into the analysis of the relationship between the hospital management (the principal) and doctors (the agents). Usually, this type of models offers results in terms of optimal payment schemes

---

[2] Belli (1997) also provides an example to show the benefits from the integration of the activities of doctors, general practitioners and insurance. However, as the author makes clear, no general conclusion can be derived unless one makes specific assumptions about the nature of the activities that can be integrated.

[3] See Mooney and Ryan (1993).

[4] It is well possible that doctors act as imperfect agents of their patients, i.e. they choose a treatment not consistent with their preferences, but this treatment is effective for them and, consequently, improves their health.

that principals should implement to give agents the appropriate incentives to act in their interest. What will be studied in this paper is, instead, how efficiency of health care production in hospitals is affected by different organisational structures (integration/separation), since these imply different information structures. More precisely, the main question, which will be addressed in the next two sections, is: is it true that, under asymmetric information, the need to provide doctors with appropriate incentives implies a loss of productive efficiency related to the different information structures that characterise the different organisations of supply?

3. Hospital Efficiency and Information Costs

3.1 The analytical setting

A hospital is an organisation comprising s different departments. Each department is a productive unit carrying out non-surgical and surgical activities and producing a level of output y. We assume that the department's total output is additively separable in the production related to each activity and denote y_i the total output produced by the ith department:

$$y_i = \sum_j y_i^j \qquad \begin{array}{l} j = activity \\ \forall i = 1,\dots,s \end{array}$$

Also, we assume that the hospital total output is simply the sum of the departments' output:

$$Y = \sum_i^s y_i$$

Under these assumptions we can treat the department's total output as a composite commodity and focus on y, the production of one representative department.

The representative department has an inherent productivity ϑ, known to the consultant, which determines the cost of producing y according to the function $C^D(y,\vartheta)$. For the purpose of our analysis it suffices to think that ϑ is the ability of the consultant, the doctor responsible of the department's activities. The hospital manager does not have any relevant private information and she does not know and cannot observe the ability of the consultant. However, she knows that ϑ is a binary random variable with strictly positive support over the closed interval $[\underline{\vartheta},\overline{\vartheta}]$ with $\underline{\vartheta}<\overline{\vartheta}$ and $p = prob(\vartheta = \underline{\vartheta})$. It is assumed that total and marginal costs are non-increasing in ϑ, i.e.:

A.1 $C_{\vartheta}^D(y,\vartheta) \leq 0$

A.2 $C_{y\vartheta}^D(y,\vartheta) \leq 0$

In addition, it is assumed that production costs are non-decreasing and convex in output for any ϑ:

$$\text{A.3} \quad C_y^D(y,\vartheta) \geq 0$$

$$\text{A.4} \quad C_{yy}^D(y,\vartheta) \geq 0$$

A.1 and A.2 are standard assumptions in hidden information models; A.2 is the "single crossing" property required to ensure that a solution that is locally incentive compatible will also be globally incentive compatible (see, for example, Guesnerie and Laffont, 1984). Finally, inequalities A.1 to A.4 are all strict inequalities for $y > 0$.

In the presence of asymmetric information on ϑ, the consultant who decides how much output is produced in his department acts always as an agent of the hospital manager (principal). The manager's role is, instead, different in different types of hospitals. We consider two types of public hospitals, hospital trusts and local public hospitals. The former are financially independent of the Health Authority that is the sole purchaser of the services provided to the patients; the latter "belong" to the Health Authority that produces in its hospitals the services provided to the patients. Accordingly, in hospital trusts the manager is the principal at the top of a two-layer hierarchy, but in local public hospitals he is the middle principal of a three-layer hierarchy where the top principal is the Health Authority.

In general, motivating privately informed agent is costly and the basic theme in the vast literature on incentives in organizations is that the cost of providing appropriate incentives is different in different hierarchical structures. In this section we use an agency model developed by McAfee and McMillan (1995) to discuss how incentives to exploit exogenous informational advantages affect productive decisions taken by consultants in hospital trusts and in local public hospitals. To rule out any consideration of risk sharing and to focus on the cost of controlling a privately informed agent in hierarchies of different length we assume throughout that the Health Authority, the manager and the consultant are all risk-neutral.

3.2 Efficiency of Production in Hospital Trusts

To model the productive decision taken in this type of hospital we assume that the financial independence motivates managers to care about hospital profits.

The objective of the manager is to maximise hospital net returns given by the difference between total revenue, $R^H(y)$, and total costs, $T^D(y,\vartheta)$[5]. The hospital revenue is exogenous; $R^H(y)$ is the fixed fee at which hospital services are sold to the health authority, determined on the basis of the relevant DRG's. $T^D(y,\vartheta)$ is a transfer to

[5] This assumption can be regarded as a simplified version of a more general one, which considers the reward for the manager as dependent on the difference between total revenues and total costs.

the representative department that covers the costs of producing output y, $C^D(y,\vartheta)$, as well as the incentive compensation that is negotiated with the privately informed consultant.

Given $T^D(y,\vartheta)$, the resources allocated to the productive unit, the consultant is entitled to retain the difference between transfers and production costs (that include the fixed part of his own salary). The consultant's net benefit is endogenous and is given by the "profit" realized by the department: $R^C(y(\vartheta)) = T^D(y(\vartheta)) - C^D(y(\vartheta),\vartheta)$.

The manager provides the consultant with proper incentives if she chooses $y(\vartheta)$ and $T^D(y(\vartheta))$ that solve the following programme:

$$\text{MAX} \quad E_\vartheta\left[R^H(y(\vartheta)) - T^D(y(\vartheta))\right] \qquad \textbf{[MP]}$$

$$\text{s.t.} \quad T^D(y(\vartheta)) - C^D(y(\vartheta),\vartheta) \geq 0 \qquad \forall\vartheta \qquad (\text{IR}^\text{C})$$

$$T^D(y(\vartheta)) - C^D(y(\vartheta),\vartheta) \geq T^D(y(\tilde{\vartheta})) - C^D(y(\tilde{\vartheta}),\vartheta) \qquad \forall\vartheta,\tilde{\vartheta} \in [\underline{\vartheta},\overline{\vartheta}] \qquad (\text{IC}^\text{C})$$

Constraints (IR^C) and (IC^C) are standard individual rationality and incentive compatibility constraints respectively. With perfect in formation about ϑ, individual rationality requires that the consultant net benefit be non-negative for all values of ϑ. Incentive compatibility requires that the net benefit form producing y when ϑ is the true ability parameter be at least as great as the benefit that the consultant can get by pretending that his ability is $\tilde{\vartheta}$.

It is well known that in this type of agency model the solution is uniquely determined by the binding individual rationality constraint for the least productive agent, and the binding incentive compatibility constraint for the most productive agent (see, for example, Sappington, 1983). These constraints yield values of $T^D(y(\underline{\vartheta}))$ and $T^D(y(\overline{\vartheta}))$ such that for the uninformed manager, the expected transfer to the department is:

$$\begin{aligned} E_\vartheta T^D &= p\left[C^D(y(\underline{\vartheta}),\underline{\vartheta})\right] + (1-p)\left[C^D(y(\overline{\vartheta}),\overline{\vartheta}) - C^D_\vartheta(y(\vartheta),\vartheta)\right] = \\ &= E_\vartheta C^D(y(\vartheta),\vartheta) + (1-p)\left[-C^D_\vartheta(y(\vartheta),\vartheta)\right] \end{aligned} \qquad (1)$$

The last term in (1) is positive, since $C^D_\vartheta(.) < 0$ by assumption A.1; it is the expected value of the net benefit that has to be paid to the consultant, i.e. the expected informational rent. In fact, (1) can be rewritten as:

$$E_\vartheta T^D = E_\vartheta C^D(y(\vartheta),\vartheta) + E_\vartheta R^C(y(\vartheta)) \qquad (2)$$

and $E_\vartheta\left[C^D(y(\vartheta),\vartheta) + R^C(y(\vartheta))\right]$ is the agent "virtual cost" (see Myerson, 1981) that in expectation exceeds production costs by an amount equal to the informational rent, and is the relevant cost for the manager when, ex-ante, she chooses $y(\vartheta)$ that maximises:

$$E_{\vartheta}\left[R^{H}(y(\vartheta))-T^{D}(y(\vartheta))\right]=E_{\vartheta}\left[R^{H}(y(\vartheta))-C^{D}(y(\vartheta),\vartheta)-R^{C}(y(\vartheta))\right] \qquad (3).$$

The optimal (second best) level of output, $\hat{y}(\vartheta)$ satisfies:

$$R^{H\prime}(\hat{y}(\vartheta))=C_{y}^{D}(\hat{y}(\vartheta),\vartheta)-C_{y\vartheta}^{D}(\hat{y}(\vartheta),\vartheta) \qquad (4)$$

and, since $C_{y\vartheta}^{D}(.)<0$ by assumption A.2, $\hat{y}(\vartheta)<y^{*}(\vartheta)$, where $y^{*}(\vartheta)$ is the full information level of output that satisfies:

$$R^{H\prime}(y^{*}(\vartheta))=C_{y}^{D}(y^{*}(\vartheta),\vartheta) \qquad (5).$$

3.3 *Efficiency of Production in Local Public Hospitals*

The Health Authority owns the local public hospital and, given its budget constraint, wants to maximise hospital net returns given by the difference between total revenue, $R^{H}(y)$, and total costs, $T^{H}(y,\vartheta)$. In this case $R^{H}(y)$ is the exogenous value of the hospital output determined on the basis of the relevant DRG's and assigned to the hospital budget. $T^{H}(y,\vartheta)$ is a transfer from the budget of the Health Authority to the budget of the hospital that covers the transfer to the representative department, $T^{D}(y,\vartheta)$, as well as the incentive compensation for the manager. As before, $T^{D}(y,\vartheta)$ is the sum of the department's production costs, $C^{D}(y,\vartheta)$, and of the incentive reward for the consultant $R^{C}(y,\vartheta)$.

Given $T^{H}(y,\vartheta)$, the resources allocated to the hospital, the manager's compensation is given by the difference between the transfer received from the Health Authority and the transfer assigned to the department. For the sake of simplicity, we assume that this payment is entirely endogenous:

$$R^{M}(y(\vartheta))=T^{H}(y(\vartheta))-T^{D}(y(\vartheta))=T^{H}(y(\vartheta))-R^{C}(y(\vartheta))-C^{D}(y(\vartheta),\vartheta).$$

The manager is responsible of negotiating the compensation required to control the privately informed consultant. For this reason, she needs to be given appropriate incentives even though she does not have any relevant information or any productive role. The specification of the set of incentive constraints for the manager reflects her double role of agent of the Health Authority and principal of the privately informed consultant.

With imperfect information about ϑ the manager accepts the job if the compensation negotiated ex-ante with the Health Authority grants her a non-negative expected return even though the actual benefit from running the local public hospital will depend ex-post on the true value of ϑ. However, she cannot be asked ex-post to bear any loss resulting from the hospital activity; accordingly, $T^{H}(y(\vartheta))$ and $T^{D}(y(\vartheta))$ must be such that $R^{M}(y(\vartheta))$ be non-negative for all values of ϑ. Having agreed on running the hospital the manager must be motivated in negotiating an incentive feasible compensation with the consultant. This is ensured by requiring that her net return if the consultant produces $y(\vartheta)$ when ϑ is the true ability parameter be at least as great as the net return she can get allowing the consultant to produce $y(\tilde{\vartheta})$.

Accordingly, to provide the manager with proper incentives the Health Authority must choose $y(\vartheta)$, $T^H(y(\vartheta))$ and $T^D(y(\vartheta))$ that solve the following programme:

$$\text{MAX} \qquad E_\vartheta\left[R^H(y(\vartheta))-T^H(y(\vartheta))\right] \qquad \textbf{[LHAP]}$$

$$\text{s.t.} \quad E_\vartheta\left[T^H(y(\vartheta))-T^D(y(\vartheta))\right]\geq 0 \qquad (\mathrm{IR}^M)$$

$$T^H(y(\vartheta))-T^D(y(\vartheta))\geq 0 \quad \forall\vartheta \qquad (\mathrm{LL}^M)$$

$$T^H(y(\vartheta))-R^C(y(\vartheta))-C^D(y(\vartheta),\vartheta)\geq T^H(y(\tilde{\vartheta}))-R^C(y(\tilde{\vartheta}))-C^D(y(\tilde{\vartheta}),\vartheta)$$
$$\forall\vartheta,\tilde{\vartheta}\in\left[\underline{\vartheta},\overline{\vartheta}\right] \qquad (\mathrm{IC}^M)$$

$$T^D(y(\vartheta))-C^D(y(\vartheta),\vartheta)\geq 0 \qquad \forall\vartheta \qquad (\mathrm{IR}^C)$$

$$T^D(y(\vartheta))-C^D(y(\vartheta),\vartheta)\geq T^D(y(\tilde{\vartheta}))-C^D(y(\tilde{\vartheta}),\vartheta) \qquad \forall\vartheta,\tilde{\vartheta}\in\left[\underline{\vartheta},\overline{\vartheta}\right] \ (\mathrm{IC}^C)$$

Constraints (IR^M), (LL^M) and (IC^M) are needed to ensure incentive feasibility of the reward to the manager; in particular, (IR^M) guarantees individual rationality ex-ante and (LL^M) guarantees limited liability. Constraints (IR^C) and (IC^C) ensure incentive feasibility of the reward to the consultant and are identical to those in problem **[MP]**.

A quick inspection of the constraint set shows that additional cost of operating local public hospitals through a longer hierarchy can only result if at the optimum of **[LHAP]** the individual rationality constraint of the manager is slack. In fact, given limited liability, this implies that for some value of ϑ, $T^H(y(\vartheta))>T^D(y(\vartheta))$.

As usual the problem is solved backwards. At the optimum of **[LHAP]** the binding constraints for the consultant are the same as in problem **[MP]** and, similarly, the binding constraints for the manager are the limited liability constraint of the manager who runs the least productive hospital and the incentive compatibility constraint of the manager who negotiates with a consultant of type $\overline{\vartheta}$. These constraints yield values of $T^H(y(\underline{\vartheta}))$ and $T^H(y(\overline{\vartheta}))$ such that for the Health Authority the expected transfer to the hospital is:

$$E_\vartheta T^H = p\left[C^D(y(\underline{\vartheta}),\underline{\vartheta})\right]+(1-p)\left[C^D(y(\overline{\vartheta}),\overline{\vartheta})-C^D_\vartheta(y(\vartheta),\vartheta)+R^C(y(\overline{\vartheta}))\right]=$$

$$=E_\vartheta C^D(y(\vartheta),\vartheta)+(1-p)\left[-C^D_\vartheta(y(\vartheta),\vartheta)+R^C(y(\overline{\vartheta}))\right] \qquad (6)$$

The last term in (6) is positive, since $C_\vartheta^D(.) < 0$ by assumption A.1. This term is the expected value of the incentive reward that has to be paid to the manager, i.e. her expected informational rent that includes the rent of the productive consultant. In fact, (6) can be rewritten as:

$$E_\vartheta T^H = E_\vartheta C^D(y(\vartheta),\vartheta) + E_\vartheta R^M(y(\vartheta)) + E_\vartheta R^C(y(\vartheta)) \qquad (7)$$

and $E_\vartheta\left[C^D(y(\vartheta),\vartheta) + R^M(y(\vartheta)) + R^C(y(\vartheta))\right]$ is the "virtual cost" of the manager that in expectation exceeds production costs by an amount equal to the sum of her informational rent and the informational rent of the consultant.

This is the relevant cost for the health authority when choosing, ex-ante, $y(\vartheta)$ that maximises:

$$E_\vartheta\left[R^H(y(\vartheta)) - T^H(y(\vartheta))\right] = E_\vartheta\left[R^H(y(\vartheta)) - C^D(y(\vartheta),\vartheta) - R^M(y(\vartheta)) - R^C(y(\vartheta))\right] \quad (8).$$

Comparison of (8) with (3) shows that in local public hospitals the optimal (second best) level of output is lower than in hospital trust and, hence, is more distorted away from the full information level.

4. The supply of hospital services in Italy

The National Health Service in Italy has been exposed to two subsequent reforms, in 1992 and 1999. We will briefly summarise here the main contents as far as hospital services are concerned. Before these reforms, the health care system was an integrated system where the Local Health Authorities – USL, Unità Sanitarie Locali – acted as both purchasers and providers of the services, and public hospitals were incorporated into USLs. After the reforms, major hospitals, those that come close to our definition of Hospital Trusts, were delegated responsibility for their own budget, finances, management and technical functioning. Consequently, they were separated from the USLs. The hospitals that may eventually become self-governing include all teaching hospitals, other public hospitals providing specialised services within university-related facilities, nationally important hospitals and those emergency wards. Their administration will conform to the principle of economic and financial independence, with budgets drafted by their managers. The service they provide are reimbursed according to a payment-per-case system based on DRGs. All other public hospitals, not reorganised as independent entities, and which resemble our definition of Local Public Hospitals, are still governed by USLs, even if they are accorded economic and financial autonomy, with separate accounts within the overall budget of their USL.

5. Efficiency estimation

Two main techniques have been developed for the estimation of production efficiency. There is the econometric approach, based on parametric estimation of production functions by fitting a regression plane through the centre of the data.

The second approach is Data Envelopment Analysis (DEA) which was developed by Charnes, Cooper and Rhodes (1978). Using observed output and input data, and without making any assumptions on the nature of the behavioural objective (e.g. cost maximisation or profit maximisation) or the functional form of the underlying technology, DEA calculates a measure of the efficiency of each observation. This is accomplished by estimating linear programming models to construct an empirically based frontier, and by evaluating each observation against all others included in the data set. The only assumption made in DEA is that the piece-wise linear envelopment frontier is convex. Another advantage of DEA is that it can handle easily multiple outputs and multiple inputs technologies.

We consider that there are $k = 1,...,K$ observations that use a vector of N inputs, $\mathbf{x} \in \Re_+^N$, to produce a vector of M outputs, $\mathbf{y} \in \Re_+^M$. The original DEA model assumes that the production process presents Constant Returns to Scale (CRS) and defines the efficiency of the observation k' as the solution of the following linear programming problem

$$\begin{aligned} \text{CRS efficiency} = \min\Bigg\{ \theta &: \sum_{k=1}^{K} \lambda_k y_{mk} \geq y_{mk'}, m = 1,\ldots, M, \\ &\sum_{k=1}^{K} \lambda_k x_{kn} \leq \theta x_{mk'}, n = 1,\ldots, N, \lambda_k \geq 0, k = 1,\ldots, K \Bigg\} \end{aligned} \tag{9}$$

Banker, Charnes and Cooper (1984) extended the original DEA model and showed that is possible to assume that the reference technology variable returns to scale (VRS) adding the constraint $\sum_{k=1}^{K} \lambda_k = 1$ to the linear programming model which imposes that hyperplanes for each observation do not pass through the origin. The DEA efficiency with VRS of the observation k' is defined as

$$\begin{aligned} \text{VRS efficiency} = \min\Bigg\{ \theta &: \sum_{k=1}^{K} \lambda_k y_{mk} \geq y_{mk'}, m = 1,\ldots, M, \\ &\sum_{k=1}^{K} \lambda_k x_{kn} \leq \theta x_{mk'}, n = 1,\ldots, N, \lambda_k \geq 0, k = 1,\ldots, K, \sum_{k=1}^{K} \lambda_k = 1 \Bigg\} \end{aligned} \tag{10}$$

Banker, Charnes and Cooper showed that the efficiency estimated by the DEA with CRS is a measure of *overall technical efficiency* which can be broken up into measures of *scale efficiency* and *pure technical efficiency*, where the latter is the efficiency measured by the DEA with VRS.

Numerous examples now exists in which DEA has been successfully applied to the study of health care organisation. Papers by Sherman (1984) and Banker et al. (1986) were among the first to apply DEA to the study of hospitals efficiency. Particularly interesting are the studies that have evaluated ownership and performance across hospital types using DEA. We mention the work of Grosskopf and Valdmanis (1987) and Valdmanis (1992) that, using a sample of hospitals operating in Michigan, found that public (government-owned) hospitals were consistently more efficient that non-for-profit hospitals. DEA has been used to compare the performance of public hospitals with for-profit hospitals, but while in some studies public hospitals appeared to be more efficient than for-profit hospitals (Ozcan et al 1992; Ozcan e Luke 1993) other studies reached opposite conclusions (Ferrier and Valdmanis 1996).

6. Data

The data in this study consist of all public hospitals operating in one Italian Region, Lombardy, and relate to the year 1996.[6] The sample contains 16 hospital trusts and 93 local public hospitals. Following the literature on hospital efficiency, we define four inputs: the number of physicians, the number of nurses, the number of other personnel and capital, measured as the number of staffed beds.[7]

The output vector measures hospital activity based on the Diagnosis Related Groups (DRGs). The DRGs provide information about the number of in-patient admissions, in-patient days and day cases. The Italian DRG system has 492 categories, which must be aggregated into broader composite outputs. This raises the issues of the choice of the aggregation weights, the choice of the unit of measurement and the choice of the aggregation criteria. When weighting, we follow the approach of Ozcan and Luke (1993) and Magnussen (1996) and we use the relative costs attached to the DRGs as aggregation weights.[8] In relation to the unit of measurement, previous analysis of hospital efficiency have defined the output either using the number of in-patient admission (Ozcan and Luke 1993) or the number of in-patient days (Grosskopf e Valdmanis 1987; Valdmanis 1992). To see the effect of changing the unit of measurement, we perform this aggregation using both in-patient admission and in-patient days as unit of measurement.[9] Finally, in relation to the aggregation criteria we based the output vector on type of activity and we choose to aggregate the DRGs in medical and surgical activities (see Grosskopf and Valdmanis 1987; Magnussen 1996).

The outputs defined so far measure the activity of the hospital. Activity is only an intermediate output, as services are delivered to patients in order to produce health improvement, which may be considered as the final output of the hospital production process. Health improvement depends also on the quality of the health care and it is reasonable to assume that quality is the result of differences in the organisation of the hospital, skills and motivation of the personnel. Therefore, it is important to try to include quality in the analysis of hospital efficiency at least for the following two reasons: firstly, because quality in one of the determinants of patients' health improvement. Secondly, because it is reasonable to suppose that there is a trade off

[6] Data were obtained from the Servizio Osservatorio Epidemiologico of Lombardy and the Italian Minister of Health.

[7] Similar inputs were, for instance, used by Burgess and Wilson (1995) and Magnussen (1996).

[8] We have used the cost published in the Gazzetta Ufficiale D.M. 30.06.1997.

[9] Depending on the unit of measurement, the weights reflect the relative cost per in-patients admissions, in-patient days or day-cases.

between efficiency and quality, in other words, higher quality can be achieved only using of more resources.

As measures of the quality of hospital activity we use the mortality rate among patients admitted for cerebrovascolar diseases (DRG 14) and cardiovascular diseases (DRG 121 e DRG 122). Operationally, quality is embodied in the DEA model, including among the inputs the number of patients admitted with the relevant diagnosis (DRGs 14, 121 and 122), and among the outputs the number of those patients discharged alive. The descriptive statistics of each of the variables used in this study are given in Table 1. Hospital trusts are, on average, larger than local public hospitals and show higher mean values of all inputs used and outputs produced.

7. Results

The DEA approach has various valuable properties, as we discussed in the previous sections, which makes it amenable to study efficiency in hospitals. Problems may arise, however, in terms of the choice of variables. Smith (1997) showed that misspecification can have serious implications on DEA results, but whereas in econometrics model specification may be assessed using the R^2 statistics and using statistical tests based on the residuals of the regression, no corresponding tests exist in DEA. To evaluate the validity of DEA results, the criterion we use in this paper is that of robustness. In other words, for a finding to be considered robust, it must be shown that minor changes in the list of variables cannot alter fundamentally the conclusion of the DEA analysis.[10]

A total of 18 different DEA models are evaluated and the specification of each model is presented in Table 2. Differences regard three main aspects: firstly the orientation of the DEA model. Even if the DEA model contains exactly the same vector of inputs and outputs, efficiency may vary if it is defined in term of minimisation of the input used of maximisation of the output produced. We also evaluate three different specifications outputs vector. In the first specification we define hospital activity using the number of in-patient admissions; in the second specification we use the number of in-patient days; in the third specification we included both the number of in-patients admissions and the number of in-patient days regarding them as two separate activities.[11] Finally, the models estimated may or not account for the quality of hospital activity.

Table 3 provides the results of all 18 variations, distinguishing between overall technical efficiency and its two components, pure technical and scale efficiency. Statistical significance is tested by the Mann-Whitney test. The null hypothesis is that the distribution of the efficiency measure is the same for hospital trusts and local public

---

[10] Applications of this rule in the analysis of hospital efficiency can be found in Valdmanis (1992), Parkin and Hollingsworth (1997).
[11] In the first two specifications the weights used are, the relative cost per in-patients admissions and in-patient days, respectively. In the third specification we weighted in-patients admissions using the relative cost of one day in-patient admission.

hospitals. The mean overall efficiency measure for hospital trusts ranges from 87.836 to 100%, whereas the mean efficiency for local public hospitals ranges from 80.777 to 97.128%. In all specifications, hospital trusts are more efficient relative to local public hospitals and the difference is statistically significant in three out of 9 specifications.

Relaxing the assumption on returns to scale, we observe that the differences in mean pure technical efficiency between hospital trusts and local public hospitals are more marked and they are statistically significant in 16 out of 18 specifications. Finally, the two types of hospitals do not show clear differences in the level of scale efficiency. In two specifications hospital trusts appear, on average, significantly more scale efficient than local public hospitals, on the other hand in the remaining 16 specifications we observe that local public hospitals are the most efficient, but the results are not statistically significant.

8. Conclusions

This paper has dealt with the issue of how efficiency of health care production in hospitals is affected by different organisational structures (integration/separation), such as those currently characterising the Italian hospital system. The main question to be answered was whether the need to provide doctors with appropriate incentives implies a loss of productive efficiency, related to the different information structures that characterise the different organisations of supply.

Therefore, we model two different types of hospitals: *hospital trusts*, whose manager is regarded as the principal at the top of a two-layer hierarchy; *local public hospitals*, in which the manager is the middle principal of a three-layer hierarchy where the top principal is the Health Authority. Using an agency model developed by McAfee and McMillan (1995) the paper shows how incentives to exploit exogenous informational advantages affect productive decisions taken by consultants in hospital trusts and in local public hospitals. The main implication is that in local public hospitals the optimal (second best) level of output is lower than in hospital trust and, hence, is more distorted away from the full information level.

The empirical analysis confirms the findings of the theoretical model. Hospital trusts have, on average, higher efficiency rating than local public hospitals, with the difference in performance being mainly due to the pure technical component.

References

Banker, R.D., Chang, H., Cooper, W.W. (1984) Some models for estimating technical and scale inefficiency in data envelopment analysis. *Management Science*, 30, 1078-1092.

Banker, R.D., Conrad, R.F., Strauss, R.P. (1986) A comparative application of data envelopment analysis and translog methods: An illustrative study of hospital production. *Management Science*, 32, 30-44.

Belli, P. (1997) *La Rilevanza della Teoria dei Diritti di Proprietà per lo Studio della Riforma Sanitaria Inglese*, in Petretto ed. (1997), pp.127-158.

Charnes, A., Cooper, W.W., Rhodes, E. (1978) Measuring the efficiency of decision making units. *European Journal of Operational Research*, 2, 18-39.

Clerico, G. (1997) *Isitituzioni, Incentivi ed Efficienza in Sanità*, Cedam, Padova.

Ferrier, G.D., Valdmanis,V. (1996) Rural hospital performance and its correlates. *Journal of Productivity Analysis*, 7, 63-80.

Grosskopf, S., Valdmanis, V. (1987) Measuring hospital performance: A non-parametric approach. *Journal of Health Economics*, 6, 89-107.

Halonen, M., Propper, C. (1999) The organisation of gvoernment bureaucracies: the choice between competition and single agency, *Centre for Market and Public Organisation*, University of Bristol, Working Paper 99/010.

Hart, O. (1996) *Firms, Contracts and Financial Structure*, Clarendon Press, Oxford.

Kim, S.K. (1997) Limited liability and Bonus Contracts. *Journal of Economics and Management Strategy*, 6, 899-913.

Magnussen, J. (1996) Efficiency measurement and the operationalization of hospital production. *Health Services Research*, 31, 21-37.

McAfee, R.P., McMillan, J. (1995) Organizational diseconomies of scale. *Journal of Economics and Management Strategy*, 4, 399-426.

Mooney, G., Ryan, M. (1993) Agency in health care: getting beyond first principles. *Journal of Health Economics,* 12, 125-135.

Myerson, R. (1981) Optimal Auction Design. *Mathematics of Operations Research*, 6, 58-73.

Ozcan, Y.A., Luke, R.D., Haksever C. (1992) Ownership and organizational performance: A comparison of technical efficiency across hospital types. *Medical Care*, 30, 781-784.

Ozcan, Y.A., Luke, R.D. (1993) A national survey of the efficiency of hospitals in urban markets. *Health Services Research*, 27, 719-739.

Parkin, D., Hollingsworth, B. (1997) Measuring production efficiency of acute hospitals in Scotland, 1991-94: Validity issues in data envelopment analysis. *Applied Economics*, 29, 1425-1433.

Petretto, A., (1996a) *Regionalizzazione della Sanità e Controllo della Spesa Pubblica*, Nota n. 5, Ministero del Tesoro, Commissione Tecnica per la Spesa Pubblica.

Petretto, A., (1996b) *Razionalizzazione e Contenimento della Spesa Ospedaliera*, Raccomandazione n. 5, Ministero del Tesoro, Commissione Tecnica per la Spesa Pubblica.

Petretto, A., (1997) *Economia della Sanità*, Il Mulino, Bologna.

Sappington D.,(1983) Limited liability contracts between principal and agent. *Journal of Economic Theory*, 29, 1-21.

Sherman, H.D. (1984) Hospital efficiency measurement and evaluation. Empirical test of a new technique. *Medical Care*, 22, 922-38.

Smith, P. (1997) Model misspecification in data envelopment analysis. *Annals of Operational Research*, 73, 233-252.

Valdmanis, V. (1992) Sensitivity analysis for DEA models: An empirical example using public vs. NFP hospitals. *Journal of Public Economics*, 48, 185-205.

Table 1: Descriptive statistics: Mean values (standard deviation)

Variables	Entire Sample	Hospital Trusts	Local Public Hospitals
Inputs			
Physicians	111.688	344.500	71.634
	(128.027)	(162.163)	(62.498)
Nurses	285.284	815.875	194.000
	(291.181)	(361.614)	(144.316)
Other personnel	293.165	926.250	184.247
	(334.278)	(452.539)	(126.843)
Beds	322.835	894.563	224.473
	(322.735)	(415.889)	(165.663)
Outputs			
Medical in-patients admissions	7643.184	20677.813	5400.667
	(9037.627)	(16899.168)	(3832.727)
Surgical in-patients admissions	3569.853	9789.188	2499.860
	(3837.409)	(5265.550)	(2209.490)
Number of medical day cases	4330.697	17287.688	2101.538
	(7982.831)	(14734.199)	(2278.554)
Number of surgical day cases	488.431	1946.375	237.602
	(918.589)	(1709.286)	(284.478)
Number of medical in-patients day	55978.615	153492.125	39202.097
	(56130.454)	(77813.202)	(27805.075)
Number of surgical in-patients day	29761.468	93955.125	18717.398
	(35339.803)	(43547.593)	(17802.946)
Quality variables (deaths/admissions)			
Mortality index for the DRG 14 [a]	0.126	0.148	0.112
	(0.095)	(0.047)	(0.100)
Mortality index for the DRGs 121 and 122 [b]	0.120	0.144	0.107
	(0.166)	(0.044)	(0.179)
Mortality index for the DRG 14, 121 and 122	0.124	0.147	0.110
	(0.097)	(0.032)	(0.104)

a: DRG 14: cerebrovascolar diseases.
b: DRG 121 and DRG 122: cardiovascular diseases.

Table 2: Specification of the DEA models estimated

	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18
Orientation of the DEA model																		
Input	×		×		×		×		×		×		×		×		×	
Output		×		×		×		×		×		×		×		×		×
Input	×	×	×	×	×	×	×	×	×	×	×	×	×	×	×	×	×	×
Beds	×	×	×	×	×	×	×	×	×	×	×	×	×	×	×	×	×	×
Physicians	×	×	×	×	×	×	×	×	×	×	×	×	×	×	×	×	×	×
Nurses	×	×	×	×	×	×	×	×	×	×	×	×	×	×	×	×	×	×
Other personnel	×	×	×	×	×	×	×	×	×	×	×	×	×	×	×	×	×	×
Admissions for DRG 14 [d]					×	×					×	×					×	×
Admissions for DRG 121 e 122					×	×					×	×					×	×
Admissions for DRG 14, 121 e 122 [e]			×	×					×	×					×	×		
Output																		
Medical in-patients admissions [a]	×	×	×	×	×	×												
Surgical in-patients admissions [a]	×	×	×	×	×	×												
Number of medical day cases [a]	×	×	×	×	×	×	×	×	×	×	×	×	×	×	×	×	×	×
Number of surgical day cases [a]	×	×	×	×	×	×	×	×	×	×	×	×	×	×	×	×	×	×
Number of medical in-patients day [a]							×	×	×	×	×	×	×	×	×	×	×	×
Number of surgical in-patients day [a]							×	×	×	×	×	×	×	×	×	×	×	×
Medical in-patients admissions [b]													×	×	×	×	×	×
Surgical in-patients admissions [b]													×	×	×	×	×	×
Patients discharged alive, DRG14					×	×					×	×					×	×
Patients discharged alive, DRG 121 e 122					×	×					×	×					×	×
Patients discharged alive, DRG 14, 121 e 122			×	×					×	×					×	×		

Depending on the unit of measurement, the weights reflect the relative cost per in-patients admissions, in-patient days or day-cases. See Appendix 1. Gazzetta Ufficiale D.M. 30.06.1997.
The weight used in this specification is the relative costs of one day in-patient admission.

Table 3: Efficiency results

Model	Type of efficiency	Entire sample		Hospital Trusts		Local Public Hospitals		M-W statistic	Prob. (2-code)
		Mean	SD	Mean	SD	Mean	SD		
1 e 2	Overall	81.813	16.635	**87.836**	14.122	**80.777**	16.880	-1.703	0.089
1	Pure technical	86.470	13.933	89.055	13.917	86.025	13.962	-1.125	0.261
1	Scale	94.458	9.968	**98.598**	2.434	**93.746**	10.592	-2.573	0.010
2	Pure technical	85.163	15.210	89.341	13.520	84.445	15.434	-1.299	0.194
2	Scale	95.812	7.681	**98.204**	2.719	**95.400**	8.178	-1.875	0.061
3 e 4	Overall	94.162	6.364	95.541	5.790	93.924	6.457	-1.118	0.264
3	Pure technical	96.385	5.344	**99.009**	2.241	**95.933**	5.595	-2.229	0.026
3	Scale	97.688	3.573	96.456	4.613	97.900	3.347	-0.288	0.773
4	Pure technical	96.416	5.364	**99.086**	2.098	**95.957**	5.622	-2.275	0.023
4	Scale	97.658	3.575	96.382	4.689	97.877	3.330	-0.306	0.760
5 e 6	Overall	96.207	5.298	96.326	5.131	96.186	5.352	-0.248	0.804
5	Pure technical	97.802	4.298	**99.767**	0.933	**97.464**	4.556	-2.491	0.013
5	Scale	98.356	2.872	96.539	4.802	98.668	2.292	-0.714	0.475
6	Pure technical	97.837	4.325	**99.798**	0.808	**97.500**	4.590	-2.532	0.011
6	Scale	98.323	2.892	96.509	4.840	98.635	2.308	-0.714	0.475
7 e 8	Overall	88.332	12.204	**94.307**	7.504	**87.304**	12.586	-2.224	0.026
7	Pure technical	90.709	11.225	**97.106**	6.845	**89.608**	11.486	-3.078	0.002
7	Scale	97.256	4.642	97.135	3.816	97.277	4.788	-0.237	0.812
8	Pure technical	90.509	11.875	**97.211**	6.639	**89.356**	12.215	-3.113	0.002
8	Scale	97.570	3.530	97.015	3.854	97.666	3.485	-0.056	0.955
9 e 10	Overall	94.551	6.238	**96.999**	4.119	**94.130**	6.457	-1.762	0.078
9	Pure technical	96.633	5.281	**99.823**	0.710	**96.084**	5.530	-3.215	0.001
9	Scale	97.822	3.024	97.165	3.905	97.935	2.857	-0.359	0.719
10	Pure technical	96.674	5.444	**99.842**	0.634	**96.129**	5.716	-3.243	0.001
10	Scale	97.789	2.913	97.147	3.924	97.899	2.714	-0.351	0.726
11 e 12	Overall	96.441	5.367	98.153	3.100	96.147	5.626	-1.220	0.222
11	Pure technical	97.893	4.453	**100.000**	0.000	**97.530**	4.729	-2.668	0.008
11	Scale	98.494	2.462	98.153	3.100	98.553	2.351	-0.173	0.863
12	Pure technical	97.868	4.677	**100.000**	0.000	**97.501**	4.975	-2.668	0.008
12	Scale	98.530	2.398	98.153	3.100	98.595	2.270	-0.168	0.867
13 e 14	Overall	91.582	11.051	95.764	5.668	90.862	11.600	-1.590	0.112
13	Pure technical	93.687	9.869	**98.600**	4.023	**92.842**	10.334	-2.202	0.028
13	Scale	97.634	4.543	97.120	3.992	97.722	4.644	-0.147	0.883
14	Pure technical	93.518	10.490	**98.660**	3.911	**92.634**	11.014	-2.258	0.024
14	Scale	97.884	3.513	97.058	4.033	98.026	3.420	-0.071	0.943
15 e 16	Overall	95.930	5.736	97.452	3.894	95.668	5.974	-1.046	0.296
15	Pure technical	97.520	4.576	**99.823**	0.710	**97.124**	4.839	-2.617	0.009
15	Scale	98.337	2.779	97.618	3.646	98.461	2.607	-0.128	0.898
16	Pure technical	97.522	4.719	**99.842**	0.634	**97.123**	4.998	-2.617	0.009
16	Scale	98.342	2.718	97.600	3.668	98.470	2.523	-0.101	0.920
17 e 18	Overall	97.325	4.837	98.469	2.839	97.128	5.087	-0.540	0.589
17	Pure technical	98.487	3.825	**100.000**	0.000	**98.227**	4.088	-2.325	0.020
17	Scale	98.796	2.330	98.469	2.839	98.853	2.244	-0.058	0.954
18	Pure technical	98.487	3.913	**100.000**	0.000	**98.227**	4.184	-2.325	0.020
18	Scale	98.800	2.289	98.469	2.839	98.857	2.195	-0.077	0.938

Numbers in bold indicate that the differences between the two type of hospitals are statistically different at 90% confidence level.