UNIVERSITÀ DEGLI STUDI DI VERONA

Behavioral Differences Between Public and Private Not-For-Profit Hospitals in the Italian National Health Service

by

G. P. Barbetta, G. Turati, A. Zago

October 2004 - #12

WORKING PAPER SERIES

Dipartimento di Scienze Economiche

BEHAVIORAL DIFFERENCES BETWEEN PUBLIC AND PRIVATE NOT-FOR-PROFIT HOSPITALS IN THE ITALIAN NATIONAL HEALTH SERVICE[#]

Gian Paolo BARBETTA - Associate Professor[♣]

Gilberto TURATI – Assistant Professor[♦]

Angelo M. ZAGO - Assistant Professor [♠]

Abstract In this paper we attempt to identify behavioral differences between public and private not-for-profit hospitals, by exploiting the introduction of the DRG-based payment system in the Italian NHS during the second half of the Nineties. We estimate the technical efficiency of a sample of hospitals for the period 1995-2000 considering an output distance function, and adopting both parametric (COLS and SF) and non-parametric (DEA) approaches. Our results show a convergence of mean efficiency scores between not-for-profit and public hospitals, and seem to suggest that differences in economic performances between competing ownership forms are more the result of the institutional settings in which they operate than the effect of the incentive structures embedded in the different proprietary forms. We also observe a decline in technical efficiency, probably due to policies aimed at reducing hospitalization rates.

Keywords: ownership forms, technical efficiency, nonprofit organizations, hospital behavior, payment systems.

JEL Codes: I11, I18, L31.

[#] We wish to thank Sergio Destefanis, Nerina Dirindin, seminar participants at the 7th EWEPA (University of Oviedo, Spain), the XIII SIEP Conference (University of Pavia, Italy), the 3rd International Conference on Health Economics, Management and Policy (University of Athens, Greece), and the 2004 AIES Conference (Bocconi University, Milan) for insightful comments. Data for this research were kindly provided by the Italian Ministry of Health, and Chiara Albertini helped us in building the dataset. Financial support from the Italian MIUR is gratefully acknowledged. All remaining errors are our own.

[♣] Università Cattolica di Milano, Istituto di Economia e Finanza.

[♦] Università di Torino, Dipartimento di Scienze Economiche e Finanziarie.

[♠] Dipartimento di Scienze Economiche, Università di Verona, Polo Didattico Giorgio Zanotto, Viale dell'Università 4, 37129 Verona VR, Italy. Ph. +390458028414, fax +390458028529-524, E-mail: angelo.zago@univr.it.

1. Introduction

Over the last decades, having experienced an increase in public health care spending, the health systems of many Western European countries have been the subject of different reforms aimed at increasing efficiency and reducing the costs of service provision. Various efforts have been made to reach these objectives, for instance the creation of "quasi markets" for health services or the modification of public reimbursement schemes to public or private service providers. Effects of these reforms are highly debated. In particular, the impact of incentives created by public policies may vary significantly following the institutional settings of service provision in different countries. Moreover, organizations with different ownership structures (public, for-profit and nonprofit) could respond to incentives and adjust their behavior in different ways and with different speed.

The Italian hospital sector is a good example at hand of such a change in public policies aimed at increasing efficiency and reducing the costs of health services. Starting from 1995, the funding mechanism for hospitals operating in the National Health Service moved from a cost-based ex-post payment (for public hospitals) and bed-day rate (for private ones) to a Prospective Payment System (PPS) based on Diagnostic Related Groups (DRG) that applies to both types of hospitals. The aim of this paper is to use the opportunity offered by this radical change in public health policy in order to identify behavioral differences between public hospitals and private nonprofit ones. Considering a sample of more than 500 Italian hospitals over a 6 year time span (1995-2000), the paper investigates whether differences in technical efficiency of hospitals with different ownership structures were affected by the introduction of this reform. We also explore whether hospitals technical efficiency increased following the introduction of the DRG-based payment system.

The paper is organized as follows. In section 2 we discuss the relationship between ownership structure, payment systems, and performance in the hospital sector. We first analyze, from a theoretical standpoint, the potential impact of incentive structures on the performance of public, private for-profit and nonprofit organizations, and thus emphasize the inconclusive results emerging from the empirical evaluation of efficiency. We also investigate whether different types of organizations behave differently. In section 3 we briefly describe the Italian hospital industry, paying particular attention to public funding mechanisms. Section 4 includes the empirical part of the paper. We illustrate our testing strategy concentrating on two possible consequences of the policy reform that represent the basis of two testable hypotheses: (i) the potential reduction of the differences in technical efficiency between public and nonprofit hospitals, and (ii) the potential increase in technical efficiency as a consequence of the policy reform. We also describe the methodological approaches for measuring technical efficiency, estimating an output distance functions with standard non-parametric and parametric techniques, and present our sample of Italian hospitals. Some considerations for further research on these issues are discussed in section 5.

We find evidence of a convergence in the mean level of efficiency between the two types of producers, public and not-for-profit (NFP) hospitals, even though we find no conclusive evidence for the disappearance of the differences in mean efficiency scores, probably because of the relatively limited time span analyzed. We try to show that this convergence suggests that differences in economic performances between competing ownership forms are more the result of the institutional settings in which they operate - e.g., the reimbursement schemes - than the effect of the incentive structures embedded in the different proprietary forms. Moreover, our findings support the idea that the introduction of a DRG-based reimbursement scheme caused a decline in technical efficiency, more pronounced for private NFP hospitals than for public ones. One possible explanation for the observed decline in technical efficiency is related to the de-hospitalization policy pursued in Italy during the Nineties: excess capacity increased more for private nonprofit hospitals than for public ones. This point, however, deserves further research.

2. Ownership structures, reimbursement schemes and performance in the hospital sector

The impact of ownership. Ownership has a relevant role in explaining economic performance; in fact, different ownership structures create different incentives to economic actors. In general, private ownership characterized by the presence of residual claimants should represent a powerful incentive to economic efficiency and cost reduction; on the contrary, public ownership and/or the absence of any claimant of residual earnings (because of the presence of a non-distribution constraint, NDC from now on) may induce shirking and could decrease effort, consequently reducing efficiency (e.g., [1]). This simple rationalization for economic efficiency represents the basis of the wide "privatization movement" experienced by several western countries over the last twenty-five years. More recently, privatization policies involved sectors – such as health and social services – that for a long time had been considered emblematic of public provision and production, because of the presence of market failures. However, in these areas, privatization has often taken the form of transferring service provision to NFP rather than for-profit organizations, i.e., private organizations subject to a NDC.

The reasons why nonprofit organizations represent a relevant player in many "welfare sectors" (such as health) of western economies are well explained by a strand of economic literature that concentrates on "information asymmetries" [2] and "government failure" [3]. According to this literature, NFP organizations can be considered as an intermediate form between private and public firms: in fact, even though they are privately owned, nonprofit organizations cannot distribute profits to any residual claimant; this reduces their incentives to exploit information asymmetries - typical of many welfare industries, health included [4] - while it increases citizens reliance in nonprofit organizations as providers of welfare services [5].

Nonetheless, the impact of these characteristics on efficiency is far from being understood. While the absence of any owner, i.e., residual claimant *à la* Alchian-Demsetz, may produce a negative impact on efficiency by reducing managerial efforts, the NDC could represent a powerful device for controlling information asymmetries among different stakeholders [6], therefore increasing their efficiency by augmenting demand. Although public and private NFP firms share a common NDC, they may differ because they pursue different goals, and therefore they face different possibilities to attract particular inputs (such as, for instance, time and money donations). These differences, in turn, can explain differences in executives' compensation, and help explain different behavior and outputs produced by organizations subject to the same NDC [7]. Even if subject to the same NDC, public and private NFP differ from each other because only public organizations can be characterized by a *soft budget constraint*. As a consequence, private NFP share with for-profit firms a common incentive to comply with the *hard* budget constraint.

Ownership and performance in the hospital sector. In order to assess the impact of ownership on performance, one needs to develop a reliable system of performance measurement. In this paper we consider technical efficiency as a good proxy for the performance of a production unit. Efficiency is generally measured as the distance between a single unit and the "best practice" production (or cost) frontier, that can be estimated with several techniques (see, for instance, [8]). A vast literature deals with empirical analysis of technical efficiency and ownership structure in the hospital sector. In general, the empirical evidence is inconclusive. Indeed, "overall, the empirical evidence demonstrates no systematic differences in efficiency between for-profit and NFP hospitals" [9]. This statement is consistent with former research results, as in [10]. In fact, studies using different techniques to estimate efficient frontiers get different results. Wilson and Jadlow [11], using a linear programming technique, found that nonprofit hospitals were less efficient than for-profit hospitals but more efficient than public ones. Using stochastic frontier regressions, [12] could not find any relevant difference in efficiency between hospitals with different ownership structures. On the contrary, [13] and [4] found public and nonprofit hospital more efficient than for-profit ones. A more recent paper [15] compared public with private nonprofit and private for profit hospitals, and found that the main difference between the three types of producers is the *soft budget constraint* characterizing the public ones. In his analysis, [15] found that nonprofit and for-profit hospitals were equally responsive to changes in financial incentives (represented by an increase in state funding for services provided to indigent patients) and significantly more responsive than public hospitals; at the same time, both profit and nonprofit institutions tended to use growth in revenues to increase their financial assets, while public institutions did not.

The impact of reimbursement systems. Nonetheless, the ownership structure (public, for-profit or NFP) is not the only relevant factor explaining economic performance; in fact, economic incentives also matter. In particular, given the relevance of public funds in backing and financing hospitals, efficiency of institutions with different ownership structures should also be considered in the light of the rules used by public authorities to regulate and fund hospitals. In fact, differences in efficiency may originate both from incentive structures characterizing hospitals with different ownership forms, and from diversity in the regulatory rules and the funding mechanism used by public (as well as private) institutions to finance private and public hospitals.

As assessed by a wide strand of literature, mostly based on evidence from the USA, hospital's payment on the basis of reimbursement of incurred costs provides no effective incentives to both cost containment and price competition among hospitals, thus resulting in massive increase in health costs for both private and public purchasers of health services. In the USA, in the Seventies, this growth in health expenditures encouraged a vast movement toward public regulation of hospitals, in the form of rate regulation and capital expenditure controls on new hospitals as well as expansion of existing ones.

While the effects of these regulations on cost containment are highly debated, at the beginning of the Eighties a new PPS was introduced for hospitals providing services to population covered by the public Medicare program. Under this payment system, hospitals are paid a fixed amount correlated to the severity of patients treated, classified according to a set of DRG. The PPS is supposed to alter hospitals' incentives - when compared to reimbursement of incurred costs - in several ways. First of all, PPS should reduce hospitals' incentives to increase average length of stay (ALOS form now on) for their patients. In fact, "reductions in inpatients costs through shorter stays would now improve a hospital's bottom line" [16]. Moreover, under the assumption that both public and private hospitals respond to opportunities to increase their profits, the PPS should also induce hospitals to increase the number of admissions for cases whose costs are below reimbursement.

The effects of this new "regulation" on cost containment in the American context however are mixed. As widely shown in [17] in a review of studies dealing with the impact of PPS on American hospitals, the introduction of the new reimbursement system – consistently with expectations - decreased ALOS; this reduction was not limited to specific diagnoses, but interested - across the board - all patients and diagnoses. As far as admission is concerned, the authors report a decrease in admission as a consequence of the introduction of PPS, while an increase was expected because hospitals "theoretically could enhance their revenue and margins by increasing the number of admissions for all DRGs for which payment exceeded the marginal cost of care" [17]. This unexpected result could be attributed to the effectiveness of controls operated by Peer Review Organizations, whose activity prevented hospitals from increasing marginal admissions of profitable patients. As a second explanation, hospitals could have faced a "greater incentive to shift admissions to outpatients treatment rather than to increase inpatients admissions" [17].

Several people were concerned about the possible reduction in quality of care caused by the introduction of PPS and the correlated incentive of hospitals to save on costs. As far as "intensity of care"[1] is concerned, [17] report that "the studies just reviewed show PPS to have reduced the intensity of care or to have left the intensity of care unchanged (but,) (…) the best study (…) actually finds an improvement after PPS". Overall the authors assert that "the negative effect of PPS on quality are not so large and consistent as to register on commonly accepted measures of major patients outcome" [17].

The introduction of PPS had a noteworthy impact on case-mix of patients treated, whose significant – and costly - increase was at least in part attributable to hospital's incentive to "upcoding" and "DRG creep". As for different ownership structures, the empirical evidence shows that for-profit hospitals tend to upcode patients more than public and private nonprofit ones [18]. However, nonprofit hospitals behave more similarly to for-profit ones in markets dominated by this last type of hospitals (e.g., [19]). Moreover, distinguishing between "nominal" and "real" responses to price changes[2], [18] finds that hospitals did not change intensity or quality of care provided to patients under DRG subject to price increase, so that upcoding only resulted in a "nominal" change. These results are consistent with findings of prior research, showing that the introduction of PPS did not increase the number of patients treated. This effect is one of the most frequently cited reasons explaining why PPS was not completely effective in containing health costs.

3. The Italian hospital sector

[1] Measured by the number of laboratory tests, therapeutic procedures, medications, days of intensive care unit utilization, etc..

[2] Where "nominal" responses refer to hospital coding practices, while "real" responses refers to admission volumes and intensity of care provided.

The hospital sector in Italy represents a good example of a recent trend – common to other countries - for the creation of quasi-markets in welfare industries. In fact, services in the Italian hospital sector are supplied by a vast array of organizations that can be characterized as public, private for-profit and private NFP.[3] All these hospitals compete with each other for the provision of services, either directly (in Regions where patients are entitled to choose their provider of hospital health care services[4], or moving to a different Region), or indirectly (in Regions where Local Health Units[5] negotiate for their patients with public and private producers). While public hospitals are generally completely financed with public funds, private hospitals rely on a mix of public funds (counterpart of services provided to citizen covered by the National Health System, NHS from now on) and private ones (coming from citizens acquiring services that add to, or replace, those provided by the NHS).

In 2001, according to the most recent data made available by the Census of industries and services [22], in Italy operated more than 3000 hospitals[6]; the largest share of these hospitals was run by for-profit corporations (39%), followed by public (38%) and nonprofit institutions (22%). When considering employees[7], however, public hospitals gain the upper hand with about 79% of personnel, followed by nonprofit (11%) and for-profit hospitals (10%)[8]. Considering employees as a proxy variable for the number

[3] In particular, given their ownership structure, we consider as "public" the following categories of hospitals: Hospitals incorporated as public firms (*Aziende Ospedaliere)*, Local hospitals directly managed by Local Health Units (LHU), the local branches of the National Health System (NHS) (*Ospedali a gestione diretta ASL)*, Teaching and research hospitals incorporated as public bodies (*Istituti di Ricovero e Cura a Carattere Scientifico pubblici)* and University hospitals (*Policlinici Universitari)*. We classified as "not-for-profit" hospitals both Teaching and research hospitals incorporated as private bodies (*Istituti di Ricovero e Cura a Carattere Scientifico privati*) and Hospitals run by religious bodies (*Ospedali classificati*).

[4] The Health care system in Italy is mainly managed at the regional levels (see, e.g., [20] and [21] for a detailed discussion.)

[5] Local Health Units (LHU from now on) are vertically integrated public providers of health services at the local level. They can either produce directly these services, or purchase them from public and private hospitals.

[6] These figure include acute, psychiatric and long term care hospitals.

[7] Data include both paid personnel (679.322 units) and volunteers (15.500 units).

[8] Most unfortunately, the census data do not provide any information about the number of beds and the output produced.

of beds, data show quite clearly that public hospitals are much larger than private ones, while nonprofit institutions are larger than for profit ones.

Despite the large number of private (profit and nonprofit) institutions funded by the Italian NHS and providing health services to the population, side by side with public hospitals, the impact of ownership structure on performance and efficiency of these organizations has not been yet systematically analyzed. A few studies have been undertaken to measure technical efficiency of Italian hospitals. However, these studies are generally not interested in the relationship between efficiency and ownership structure, in terms of the distinction between public hospitals and private for-profit and NFP ones. For instance, [23] distinguish among five types of hospitals, but do not separate private nonprofit hospitals from private lucrative ones, a difference that economic theory deems to be important. An exception is a paper [24] that finds a weak impact of the nonprofit ownership structure on efficiency, considering a sample of hospitals located in Lombardia, the region that – at the end of the Nineties – moved toward the introduction of the highest level of competition in the hospital industry.

Moreover, almost no attempt has been made to establish the impact of changes in the funding mechanism on the performance of hospitals with different ownership structure. The funding mechanism of Italian hospitals changed significantly over time. Before 1978, when the NHS was introduced in Italy, both public and private hospitals were funded on a bed–day rate by patients. Rates were established at the hospital level and patients' costs were covered by mutual health funds. Not surprisingly, this funding mechanism created the incentive to increase both prices and the length of treatments, therefore jeopardizing the financial firmness of mutual health funds. When, at the end of the Seventies, most mutual health funds went through significant financial difficulties, the legislature passed the so called "first reform" of the Italian health care system. Mutual health funds were replaced by the NHS, a universal scheme providing free health care to all of the Italian citizens. The system was funded by the central government through a mixture of general and specific taxes, and was managed by regional authorities. Most public hospitals were incorporated into LHU, and did not enjoy any legal and financial autonomy, while private for-profit and NFP hospitals preserved their independence and many of them qualified as providers for the NHS. The 1978 reform established that public hospitals incorporated into LHU should get their funds from the budget of those units; in general, all of these hospitals had their expenses completely covered by public funds, regardless of their amount (*ex-post* payment). In a sense, this funding mechanism originated a problem of soft budget constraint, as hospitals costs were regularly reimbursed by the LHU, while expenses (and debts) of LHU were – at regular intervals of a few years – covered by regional and national governments (on this point, see [25]). On the contrary, private hospitals providing services to the NHS were to be reimbursed on a bed-day rate; rate levels were established at the national level and updated every 3 years.

These different funding mechanisms could induce different behavior in the two categories of hospitals. In principle, public hospitals did not have any incentive to keep costs under control, given the demonstrated propensity of the national government to bail out their debts. This situation, together with a real necessity to increase the amount and quality of health services provided to the population, could significantly contribute to decrease efficiency of hospitals and to increase public spending. On the other side, private hospitals had a clear incentive to boost the ALOS in order to raise their revenues; at the same time, they were encouraged to keep costs under control in order to raise profits. These incentives could result in high public costs and high private profits.

In 1992, difficulties in controlling public expenditures in the health sector gave way to a "second reform" of the Italian NHS. Public hospitals were given some level of autonomy (some of them – the largest – were in fact separated from the LHU and incorporated as self-governing public firms or *Aziende Ospedaliere*) in order to favour competition among different providers, and to create "quasi-markets" in some areas of the health care industry. Moreover, the funding mechanism of hospitals moved from ex-post payment (for public hospitals) and bed-day rate (for private ones) to a PPS that applies to both types of hospitals. The system was initially based on DRG with rates defined at the national level[9], but funding of hospitals was still in charge of the regional authorities that relied on funds transferred by the national government; the regional authorities were free to define rates different from – but not higher than – the national ones. The new reimbursement system – that had to be started between 1995 and 1997 – went through several changes in the following years: the reimbursement rates have been recalculated on the basis of the costs of a larger set of hospitals; moreover, in a general framework of devolution of tasks from the national to the local level, regional governments have been allowed to apply different regulations and, as a result, great differences in regional health services are now the rule more than the exception in Italy.

Studies concerning the impact of the introduction of PPS based on DRG are rather scarce in Italy. Following the introduction of the new reimbursement scheme, [21] register a decrease in ALOS, together with a fall in the number of inpatients, as a result of a sharp decrease in admissions in public hospitals and a mild growth in admissions in for-profit hospitals. Moreover, the incentive effect of PPS has been reduced by the simultaneous attempt to contain hospitalization rates and public hospital expenditure. In this sense, the introduction of tight budget ceilings for hospitals may have reduced (technical) efficiency, generating productive capacity in excess. At the same time, working on a small regional sample of hospitals, [26] observe both an increase in day-hospital and a decrease in ordinary hospital admissions. In addition, the authors find a growth in severity of illness among hospitalized patients, partly due to upcoding and DRG creep. Similar results can be found in [27]. Apart from differences in the reduction of ALOS, nothing is known with respect to the behavioral response of hospitals to PPS-DRG

[9] With rates based on the full costs of a very limited set of 8 medium to large hospitals.

introduction. This is the focus of our attention in the empirical part of the paper.

4. The empirical analysis

4.1. Methodology

Testing strategy. The main aim of the paper is to investigate the behavioral differences in public and private nonprofit hospitals using the opportunity offered by the change in the reimbursement scheme occurred after 1995. As discussed above, before the introduction of the DRG-based payment system, public and private hospitals faced different reimbursement schemes (hence, different incentives in terms of pursuing productive efficiency). On the contrary, with the new scheme, both private and public hospitals receive their funds on the basis of the amount and the nature of services provided, and they should share a common incentive to decrease ALOS, thus the average cost of treatment with respect to the reference hospitals (i.e., those hospitals whose costs were used as a basis for setting reimbursement rates). In fact, not being able to perform such a task would jeopardize the economic performance of the hospital. Differences in efficiency would then be due solely to differences in behavior characterizing the two ownership structures. Therefore, in the remainder of the paper we are interested in testing the following hypothesis:

- *$H_0(A)$: differences in hospitals technical efficiency characterizing different ownership structures disappear after the introduction of the DRG-based payment system*.

Failing to reject $H_0(A)$, i.e., observing a convergence of efficiency between NFP and public hospitals, would give empirical support to the hypothesis that differences in efficiency between public and private hospitals were mainly due to different reimbursement schemes, and not to different incentives embedded in the diverse ownership structures.

The new reimbursement system is obviously aimed at keeping public spending for health under control, as a result of the expected improvement in productive efficiency; its effects could depend on capacity utilization of the reformed hospitals. Considering excess capacity, two alternative hypothesis should be considered. Suppose first that a hospital is currently producing health care services *without* excess capacity, i.e., the hospital produces the maximum amount of potential inpatient days given its beds: in this case, the hospital can reduce ALOS only by increasing beds' turnover, hence the number of discharged patients. Suppose now that the hospital is producing *with* excess capacity: reduction of ALOS (and of excess capacity) can be obtained by increasing the number of discharged patients in this case too. Thus, in both cases, following the introduction of the DRG-based payment system, we should observe an increase in the number of patients, hence an increase in efficiency measured by considering the number of patients among outputs. Therefore, the second hypothesis to be tested is:

- *$H_0(B)$: hospitals technical efficiency, measured by considering the number of patients, increases following the introduction of a DRG-based payment system*.

Rejecting *$H_0(B)$* would give empirical support to the hypothesis that the change in the reimbursement scheme did not produce beneficial effects in terms of cost containment.

Both tests are based on a measure of technical efficiency. We obtain scores of technical efficiency by estimating an output distance function (see below), and then testing for differences in means. In particular, the null hypothesis (A) will be rejected if we detect statistically significant differences in the mean levels of efficiency after the change in the reimbursement scheme between public and private NFP producers. The null hypothesis (B) will be rejected if we detect no statistically significant differences in the mean levels of efficiency before and after the change in the reimbursement scheme.

Of course, one must recognize that the DRG-based payment system has some potential drawbacks: the very first one is the incentive to discharge patients earlier than necessary, influencing negatively the quality of care; a second related problem is the practice of discharging and readmitting patients, simply to increase revenues; a third problem is the practice of upcoding, i.e., the hospitals' practice of registering patients in more severe - and costly – DRG; a fourth problem is cream-skimming, that is the incentive to provide only the more lucrative - and less severe - services, the ones the hospital has a cost advantage for, making it quite difficult any planning of service provision at the local level. In this paper, because of data availability, we cannot control for these problems. However, [18] and [19], for instance, show that for-profit hospitals in the U.S. upcoded more than public and nonprofit ones, the two ownership structures on which our empirical analysis is based.

Non-Parametric estimation of the output distance function. The literature on the measurement of productive efficiency has produced several techniques to estimate the efficiency characterizing production units. Here we concentrate only on three of these: Data Envelopment Analysis (DEA), Corrected Ordinary Least Squares (COLS), and Stochastic Frontiers (SF).

DEA has been used in management science to evaluate *ex-post* the efficiency of achieving an objective from a given level of inputs [28]. Its applications in economics build on the work of Debreu [29], Koopmans [30] and [31] and Farrell [32]. DEA employs linear programming techniques to measure efficiency as the distance of each firm from a non-parametric production frontier constructed from convex combinations of observed input-output pairs. Let $x \in \Re_{+}^{N}$ be a vector of inputs and $y \in \Re_{+}^{M}$ be a vector of outputs. Feasible input-output combinations are represented by the production possibilities set, $T \subset \Re_{+}^{N} \times \Re_{+}^{M}$:

$$T = \left\{(x, y):\ x \text{ can produce } y\right\}. \tag{1}$$

We assume that T satisfies standard axioms listed, for instance, in [33] or [34]. For a given input-output vector (x, y), the output distance function

[35] is the minimum proportional expansion of all outputs such that the output combination can still be produced from the original input vector:

$$D_O(y,x) = \inf\left\{\alpha : \left(x, \frac{y}{\alpha}\right) \in T\right\}. \tag{2}$$

The output distance function[10] is a non-decreasing measure of efficiency, homogeneous of degree 1 and convex in y, decreasing in x and ranging between 0 and 1, where a value of 1 represents technical output efficiency.
The calculation of the output distance function requires the solution of a nonlinear programming problem, but an easier approach is available. Indeed, the reciprocal of the output distance function is the Farrell's measures of output efficiency defined by $F_O(y,x) = \sup\{\gamma : (\mathrm{x}, \gamma\ \mathrm{y}) \in T\}$ where $\mathrm{F}_O(x,y) \geq 1$. This measure is easily obtained as a solution to a linear programming problem [36].
Alternative assumptions on returns to scale characterizing production units are available. We start here by computing the Farrell's measure of output technical efficiency for a Variable Returns to Scale (VRS) technology by solving the following linear program:

$$\begin{aligned} & F_O\left(y,x \middle| V\right) = \max \gamma : \\ & \sum_{j=1}^{J} z_j y_{jm} \geq \gamma y_{jm}, \quad m = 1,2,..,M, \\ & \sum_{j=1}^{J} z_j x_{jn} \leq x_{jn}, \quad n = 1,2,..,N, \\ & \sum_{j=1}^{J} z_j = 1. \end{aligned} \tag{3}$$

where the z's are the DEA weights to be estimated, and the other variables remain defined as before. Färe et al. [36] suggest an informative decomposition of the most restrictive Constant Returns to Scale (CRS) technical efficiency measure into components based on scale efficiencies

[10] We consider here an output orientation, instead of calculating an input-oriented distance function, because in Italy – as in other countries – the problem of waiting lists for hospitals has become a priority for health policies at the national level. For instance, the Budget Law for the year 2006 requires hospitals to respect national defined standards for waiting lists;

and the least restrictive VRS technical efficiency measure. The construction of this measure enables the decomposition of the CRS output distance function measure into sources of output scale and technical efficiency under VRS, and this latter is also referred as a measure of the pure technical efficiency. The decomposition can be written as:

$$F_O(x^j, y^j \mid C) = F_O(x^j, y^j \mid V) \times S_O(x^j, y^j), \tag{4}$$

with $j=1, \dots, J$, where $S_O(x^j,y^j)$ is the scale efficiency measure, which can be computed for each observation j by calculating the ratio $F_O(x^j,y^j|C)/F_O(x^j,y^j|V)$[11].

In addition, by running a further DEA problem with Non-Increasing Returns to Scale (NIRS) imposed[12], it is possible to discern whether an hospital is operating in an area of increasing or decreasing returns to scale. Indeed, when the NIRS efficiency score is equal to the VRS score, then decreasing returns to scale holds locally. In the other case, when the NIRS and VRS efficiency scores are unequal, increasing returns to scale applies.

Parametric Estimation of the output distance function. We also model the multi-input-multi-output production frontier of hospitals as an output distance function in a parametric setting by following, for instance, Coelli and Perelman [37]. This allows us to avoid output aggregation that can bias efficiency scores estimates. In particular, we specify Eq. (2) using the translog functional form. Hence, our general empirical model can be written as:

$$\begin{aligned} \ln D_{Oft} = \alpha_0 + \sum_{i=1}^{M} \alpha_i \ln y_{ift} + \frac{1}{2}\sum_{i=1}^{M}\sum_{j=1}^{M} \alpha_{ij} \ln y_{ift} \ln y_{jft} + \sum_{h=1}^{K} \beta_h \ln x_{hft} + \\ + \frac{1}{2}\sum_{h=1}^{K}\sum_{k=1}^{K} \beta_{hk} \ln x_{hft} \ln x_{kft} + \sum_{h=1}^{K}\sum_{i=1}^{M} \delta_{hi} \ln x_{hft} \ln y_{ift} \end{aligned} \tag{5}$$

---

hence, they need to increase the number of patients treated, holding constant the amount of resources consumed.

[11] Notice that the returns to scale estimates with an input and output orientation are equal only when the technology is characterised by CRS.

[12] With NIRS, the last constrain of equation 3 becomes $\sum_{j=1}^{J} z_j \leq 1$.

where *f* is an index for hospitals and *t* is an index for the years from 1995 to 2000. Since homogeneity of degree 1 in output implies $D_O(x, \omega y) = \omega D_O(x,y),\ \forall \omega>0$, we choose $\omega=1/y_M$ and normalize Eq. (4) with respect to the *m-th* output:

$$\ln\left(\frac{D_{Oft}}{y_{Mft}}\right) = \alpha_0 + \sum_{i=1}^{M}\alpha_i \ln\left(\frac{y_{ift}}{y_{Mft}}\right) + \frac{1}{2}\sum_{i=1}^{M}\sum_{j=1}^{M}\alpha_{ij}\ln\left(\frac{y_{ift}}{y_{Mft}}\right)\ln\left(\frac{y_{jft}}{y_{Mft}}\right) + \\ + \sum_{h=1}^{K}\beta_h \ln x_{hft} + \frac{1}{2}\sum_{h=1}^{K}\sum_{k=1}^{K}\beta_{hk}\ln x_{hft}\ln x_{kft} + \sum_{h=1}^{K}\sum_{i=1}^{M}\delta_{hi}\ln x_{hft}\ln\left(\frac{y_{ift}}{y_{Mft}}\right) \tag{6}$$

Clearly, symmetry of cross-partial derivatives entails further restrictions, namely $\alpha_{ij}=\alpha_{ji}$ and $\beta_{hk}=\beta_{kh}$. In order to ease estimation, we rewrite Eq. (6) as:

$$-\ln y_{Mft} = \alpha_0 + \sum_{i=1}^{M}\alpha_i \ln\left(\frac{y_{ift}}{y_{Mft}}\right) + \frac{1}{2}\sum_{i=1}^{M}\sum_{j=1}^{M}\alpha_{ij}\ln\left(\frac{y_{ift}}{y_{Mft}}\right)\ln\left(\frac{y_{jft}}{y_{Mft}}\right) + \\ + \sum_{h=1}^{K}\beta_h \ln x_{hft} + \frac{1}{2}\sum_{h=1}^{K}\sum_{k=1}^{K}\beta_{hk}\ln x_{hft}\ln x_{kft} + \\ + \sum_{h=1}^{K}\sum_{i=1}^{M}\delta_{hi}\ln x_{hft}\ln\left(\frac{y_{ift}}{y_{Mft}}\right) + R_f + Y_t - \ln D_{Oft} \tag{7}$$

where the output distance measure D_O is interpreted as an error term that satisfies standard OLS assumptions. We also include in the model additional controls for year (*Y*) and regional (*R*) fixed effects; the former pick up the impact of factors common to all hospitals in a given year, e.g., a change in national regulation affecting all producers, while the latter control for the effect of factors common to all hospitals in a given region, e.g., the provision of different DRG tariffs across regions. We estimate Eq. (7) by initially using the COLS methodology. Consequently, we first estimate Eq. (7) by OLS; then, by using $-\varepsilon_{max}$ (the largest negative OLS residual), we correct the intercept parameter so that the function envelope all the observations as a frontier. The distance measure for the *f-th* hospital is thus defined as:

$$D_O = \exp\{-\varepsilon_{\max} - \varepsilon_f\} \tag{8}$$

It is clear from Eq. (8) that D_O=1 for the observation with the largest negative residual, that represents the most efficient hospital.
We also reinterpret Eq. (7) by considering a composed error term; this is obtained by assuming D_O is distributed as a half-normal random variable, and by adding a normally distributed error term v, such that $\varepsilon = v + D_O$. The distance measure estimates in this SF setting are then obtained using the methodology proposed in [38], as $E[D_O|\varepsilon]$[13].

4.2. Data description

We consider data provided by the National Ministry of Health on all Italian hospitals providing health care services for the NHS, and excluded observations with missing data and hospitals devoted to Long Term Care (with ALOS > 15 days). Our final sample consists of a balanced panel of 531 hospitals, both public and NFP, observed during the years 1995-2000[14]. Our sample period can be ideally divided into two sub-periods, one from 1995 to 1997, in which the DRG-based payment system was introduced, and the other from 1998 to 2000, in which the new reimbursement scheme should have started to produce its desired effects. Data include information on different inputs and outputs usually considered in the studies on hospitals' efficiency. Input variables comprise data on staff and a rough measure of capital (the number of beds). Output variables include data on the number of discharged patients and the number of inpatient days, the number of Day Hospital (DH) and emergency room treatments.
The empirical literature on the estimation of technical efficiency in the health care sector strongly suggests the number of discharged patients to be the most reliable measure of output, since the number of inpatient days could reflect a productive choice of hospitals. For the three methodologies

[13] In the SF model we also dropped regional and year fixed effects. See the results section below.
[14] The sample represents approximately one third of all Italian acute care hospitals, which in turn constitutes about half of all Italian hospitals as described in section 3. Private for-

(DEA, COLS, and SF), we then consider a distance function model that includes as outputs the number of discharged patients, the number of emergency room cases, and the number of DH treatments.[15] As input variables are concerned, we consider six inputs, namely the number of beds for ordinary hospitalization, beds for DH treatments, physicians, nurses, teaching staff and other employees. An F-test strongly rejected further aggregation of inputs for all the years considered in the sample[16].

Unfortunately, since DRG weights are unavailable for all the years 1995-2000, main output measures are not adjusted to take into account cross-sectional and time series variation of case-mix; this could represent a severe limitation of the present analysis, that needs to be discussed in future research[17]. There are however three elements that may reduce the impact of such a limitation on our results. First, while it is true that - at a national level – an increase in the severity of illness mix was observed, [27] suggest both "real" and "nominal" causes as an explanation. We do not have evidence that real causes are the most relevant; and real causes are the ones that could mostly affect our results. Among nominal causes that may have increased severity of illness mix, we could mention both the introduction of new DRG weights (in 1997), and a major improvement in the quality of data collected (as a consequence of the learning process in DRG management). No evidence is available concerning differences among hospitals in adapting to the new system.

---

profit hospitals could not be included since complete data are not available at the National Ministry of Health.

[15] In a previous version of the paper, we also considered an additional model by substituting the number of discharged patients with the number of inpatient days. All main findings are substantially unchanged. These additional results are available in the working paper version of this work; see [39].

[16] In particular, we considered an alternative model with only two inputs, capital (sum of all available beds) and labor (sum of all employees).

[17] DRG weights for the entire Italian sample for all the relevant years after the introduction of the new payment system were unavailable. We then collected available data for the Veneto Region (77 hospitals). DEA scores calculated on this reduced regional sample confirm the analysis undertaken at the national level with a less refined output specification. Notice indeed that ALOS is included in the DRG weights, since it is anyway related to the complexity of treatment. These additional results are available from the authors upon request.

Moreover, among real causes, the increase of case-mix complexity could be explained as a result of the tendency toward the increase of DH treatments; and in our models we control for this variable. Second, [27] recall that a further factor impacting on case-mix is the incentive to "upcode" created by the introduction of PPS-DRG. While we do not have any evidence on this point in Italy, the international empirical literature shows that for-profit hospitals tend to upcode patients more than public and private nonprofit ones [18], these two latte being the ownership structures we are dealing with.

Finally, since our main goal here is to uncover differences in behavior between the two groups of hospitals, unadjusted output does not constitute a major problem if variability *within each sub-sample* coincides. In fact, if public hospitals and private NFP hospitals belong to the same distribution, then one can argue that there are no detectable differences in terms of output case-mix *with respect to the ownership structure of hospitals*; in other words, case-mix differences are not related to ownership. Indeed, Kolmogorov-Smirnov tests on the distribution of the number of patients for each year of our sample found no evidence to reject the null hypothesis of common distribution for public and private nonprofit hospitals, suggesting that case-mix variability is the same within each sub-sample. We also run Kolmogorov-Smirnov tests on the distribution of the ratio between discharged patients and the number of beds, obtaining the same conclusion.

Table 1 collects descriptive statistics for the main variables considered in the empirical analysis. On average, NFP producers appear larger than their public counterparts, both considering inputs and main outputs. This conclusion is reversed if one considers the number of DH and emergency rooms treatments.

[TABLE 1 ABOUT HERE]

Table 2 considers ALOS by years and type of hospitals. As expected, ALOS shows a steady decline from 1995 to 1999, although we observe a minor increase in the last year of our sample period. When looking at the two sub-samples, however, there appear to be strong differences between NFP and public producers, with the former able to respond more promptly and modify ALOS more significantly than the latter.

[TABLE 2 ABOUT HERE]

3.3. Results

In this section we present our efficiency score estimates, concentrating first on testing hypothesis $H_0(A)$, and distinguishing between non-parametric and parametric estimates. DEA results are presented in Table 4 and 5, and refer to a technology specified as an *intertemporal* frontier with VRS[18]; we rejected the null hypothesis of CRS at a 1% confidence level by using a Banker's test [40] (see Table 3). As described for instance in [41], since all the data are pooled, to estimate efficiency scores an assumption of time invariant technology is made when using an intertemporal frontier model[19].

Table 3. Banker (1996) Tests for Returns to Scale

	Exponential Distribution		***Half-Normal Distribution***	
H_0: CRS vs H_1: VRS	*Test Stat.*	*Critical F (1%)*	*Test Stat.*	*Critical F (1%)*
All	1.25	1.06	1.4	1.09
Public	1.25	1.06	1.39	1.09
NFP	1.95	1.34	2.26	1.52

[TABLE 4 ABOUT HERE]

[18] Using a VRS technology specifications means that the efficiency scores refer to pure technical efficiency. See below Table 8 for the decomposition of overall technical efficiency in pure technical efficiency and scale efficiency.

[19] We also experimented with contemporaneous frontiers, i.e., estimate efficiency scores for each year separately (on this point, see [38]). The main results are unchanged and the correlation between the two sets of scores is 0.87 for the model with the number of patients (0.86 for the model with the number of inpatient days).

Overall, NFP hospitals appear to be more efficient than their public counterparts (table 5). Moreover, for both types of producers, we observe a declining trend in efficiency, more marked for private NFP producers. In particular, in 1995 the average score is 0.73 for NFP and 0.64 for public hospitals; in year 2000, average scores are respectively 0.65 and 0.63. The efficiency scores of the different ownership forms thus appear to be converging (see also figure 1).

[TABLE 5 ABOUT HERE]

Indeed, the efficiency scores between the two types of hospitals are statistically different by using both a Mann-Whitney (MW) and a Kolmogorov-Smirnov (KS) test for the initial years of our sample period (1995, 1996 and 1997, even though in this last year only with the MW test). On the other hand, in the final years, in particular 1999 and 2000, NFP and public hospitals do not have significantly different efficiency scores. These conclusions do not vary substantially when considering all the years together but divided into the two sub-periods, before and after the introduction of the new DRG-based payment system. In both periods, statistically significant differences in mean efficiency levels are observed between NFP and public hospitals, but these differences tend to decrease over time. In sum, by using DEA we find some evidence that, as the impact of the DRG-PPS kicks in, the differences between ownership forms decrease; however, they become statistically insignificant only in the last year of the sample.

[FIGURE 1 ABOUT HERE]

To summarize, the analysis based on DEA shows two main findings[20]. First, there appear to be a puzzling (and unexpected for policy makers) declining trend in efficiency. Second, the results seem to show support for the $H_0(A)$, i.e., there is convergence in efficiency between NFP and public hospitals as soon as the common DRG-based payment system is implemented.

Parametric estimates of the output distance function are presented in Table 4 and 6. For the two parametric versions of the model (COLS and SF), we control for the presence of outliers by using the leverage value [42]. Results seem to show a reasonable fit to observed data: adjusted R-squared are always in excess of 97%. Estimated coefficients on output first-order terms are always statistically different from zero and present the expected sign. Most of the estimated coefficients on input first-order terms are also statistically significant, even in the presence of multicollinearity among the six inputs[21].

Mean efficiency scores derived from the output distance function estimated with COLS (including both time and regional fixed effects, see Table 4), appear always lower than those estimated with DEA for both types of hospitals and for all the years in the sample. In this case too, as a robustness check, we experimented with alternative models excluding time and regional fixed effects[22]. On the contrary, as expected, mean efficiency

[20] Notice however that we have few data points for the NFP hospitals, and thus one might question the construction of the frontier under the convexity assumption. When we run two separate DEA frontiers for public and NFP hospitals, results partially change (available from the authors upon request). We do not observe any decline in efficiency for NFP hospitals before and after the introduction of PPS; moreover, with respect to common frontier, average efficiency scores are higher. We do not observe convergence in the mean level of efficiency for public and private NFP producers. On the contrary, coherently with previous results, we do not observe any significant change in the efficiency of public hospitals. In fact, while the correlation for efficiency scores of public hospitals estimated with the common (to NFP) and separate frontier is 0.99, it is only 0.38 for NFP. This last result is due to the fact that public hospitals defines the common "best practice" frontier. However, running two separate DEA frontiers is equivalent to assuming that NFP and public hospitals have in fact different technologies, which we believe it is not very plausible an assumption. In addition, the DEA results we present in the paper are confirmed by the parametric estimation with both COLS and SF.

[21] Regressions results are not included here for brevity; all tables are available from authors upon requests.

[22] The correlation between efficiency scores obtained with the model including fixed effects and the model without fixed effects is 0.86. For more details, see [39].

scores obtained with SF are always higher than those obtained with DEA; however, correlation with estimates obtained with COLS is 0.87.

As before, NFP hospitals appear to perform better than public hospitals at the beginning of our sample period (Table 6). However, mirroring the DEA analysis, after controlling for regional and time fixed effects we observe a convergence between NFP and public hospitals, due to a declining trend in technical efficiency mainly for private NFP producers (figure 1). Indeed, according to MW and KS tests, efficiency scores appear to be statistically different only for the years between 1995 and 1997. Thus, the technical efficiency of the two types of producers is different only before the introduction of the new DRG-based payment system; this last conclusion is valid for both SF and COLS. In addition, contrary to the DEA results, the disappearing differences between NFP and public hospitals are supported by the statistical tests of the two sub-periods for both set of results[23].

[TABLE 6 ABOUT HERE]

In sum, the results based on the COLS and SF analysis broadly confirm those obtained with DEA. We thus fail to reject $H_0(A)$ with both methodologies. In essence, the different efficiency performances of NFP and public hospitals, with these latter less efficient especially in the first years considered, seem to be the consequence of the different payment systems they had to face before the reform. Once a common payment system is introduced, the efficiency of the different proprietary forms under consideration tends to converge.

The last two columns of Table 4, 5 and 6 provide a first test for our second hypothesis $H_0(B)$. When looking at DEA and SF estimates, MW and KS tests confirm a statistically significant decline in mean efficiency after the introduction of the new payment scheme (Table 4). On the contrary, the same tests do not detect differences between the two sub-periods when comparing COLS results. Very clearly, figure 1 shows indeed that overall the mean efficiency scores based on COLS did not change over the period 1995-2000[24]. The main explanation for this result relies on the role of year dummies, that pick up the impact of the new payment system; we will further explore this issue in the second stage analysis below.

Looking at the different ownership forms, results are more varied. With DEA, there appear to be differences between the two sub-periods, before and after the DRG introduction, even when distinguishing between different ownerships. Both the MW and the KS tests show that mean efficiency are different, but only at low significance levels. Overall, however, it seems reasonable to argue that the DEA analysis shows a decrease in mean efficiency levels, for hospitals of both proprietary forms, after the introduction of the DRG payment system. With COLS the impact of the DRG introduction affects only the NFP hospitals: indeed, public hospitals performance remains stable over time, whilst that of NFP ones decreases (Table 6). On the contrary, reinforcing the DEA results, scores obtained with the SF methodology suggest that, for both types of hospitals, there is a statistically significant change in performance following the introduction of the new payment system; but reaction of NFP seems on average to be greater than that of public producers[25]. Thus, our findings provide evidence to reject $H_0(B)$, since efficiency levels either remain stable, i.e., for public hospitals with COLS, or decrease, in all the other cases. This is unsurprising considering previous results on the US experience (e.g., [17]), and concerns expressed in [21] on the combined effect of PPS, de-hospitalization policies, and budget ceilings.

3.4. Discussion

---

[23] See the results of KS and MW tests in the columns titled "before DRG" and "after DRG" of Table 6.

[24] Since the vast majority of hospitals is public, the mean efficiency score for all hospitals and for the sub-sample of public hospitals tend to be very close.

Taken as a whole, our findings support the idea that the introduction of a DRG-based reimbursement scheme caused a decline in technical efficiency, more pronounced for private NFP hospitals than for public ones. In this sense, the former appear to respond more promptly than the latter to the change in incentives caused by the change in the payment scheme. This may depend on the quite different power of incentive of the reimbursement scheme used before and after the reform implementation. However, the new payment system seems to be the cause of the convergence in the observed mean level of efficiency between the two types of producers (even though a longer panel would provide more conclusive evidence for the disappearance of any differences in mean efficiency scores).

Our findings thus seem to emphasize the private nature of NFP hospitals, mirroring results in [15], that identifies the main difference between public, private for-profit and private NFP hospitals in the *soft budget constraint* of government-owned institutions. In other words, our results suggest that different performances found among the various ownership structures may be due more to the different payments systems they face than to their intrinsic incentive structures.

Of course, one can claim that our results on the different behaviour for public and private hospitals are not entirely due to the introduction of the new payment system based on DRG weights. For instance, it is possible that – besides the newly introduced system – regional policies have changed with respect to the relative role given to public and private producers.

To be sure that the observed impact on efficiency is due to change in incentives derived from reimbursement schemes, we consider also a second stage analysis based on the following general Tobit model:

$$\begin{gathered} EFF_i^* = \mathbf{X}_i^{'}\beta + u_i \\ EFF_i = EFF_i^* \quad \text{if } 0 < EFF_i^* < 1 \\ EFF_i = 0 \quad \text{if } EFF_i^* \leq 0 \quad EFF_i = 1 \quad \text{if } EFF_i^* \geq 1 \end{gathered} \tag{9}$$

[25] Notice that the decrease in (average) performance experienced by NFP hospitals seems to be due to the increase in the variance of efficiency scores.

where *EFF* are the efficiency scores estimates derived from both COLS and SF, and **X** = (*DRG*, *PREXP*, R_f), where *DRG* is a dummy variable equal to 0 from 1995 to 1997, and equal to 1 from 1998 onwards; *PREXP* measures the share of regional public health expenditure devoted to private producers; R_f are regional fixed effects aimed at capturing different regional policies and structural differences constant across years.
Our results on the different behaviour induced by the change in hospitals' reimbursement would be reinforced, if – after controlling for some of the factors that can affect efficiency – the coefficient on *DRG* still remains significant. Table 7 contains three sets of estimates for our Tobit model, differing for the specifications of the output distance function from which efficiency scores are derived (namely SF and COLS with and without regional and year fixed effects). All the three sets of regressions tell us fairly the same story: the *DRG* dummy coefficient is negative and statistically significant even after controlling for *PREXP* and R_f, the regional fixed effects[26], thus reinforcing our conclusions that the change in incentives caused a reaction in hospitals' behaviour.
Moreover, the reaction was sharper for private than for public hospitals, as interacting the *DRG* dummy with a dummy for public (*PUB*) and for private NFP (*NFP*) hospitals, we obtain coefficients that are both negative, but different in magnitude. This latter conclusion is strengthened by a LR test, that confirms the difference between the two coefficients. Finally, the coefficient on *PREXP* is always positive and statistically significant, suggesting a positive role of private producers in influencing hospitals efficiency, presumably through a more competitive environment.

[TABLE 7 ABOUT HERE]

One point that deserves further discussion concerns the interpretation of our results, especially in order to derive policy implications. The new DRG-based payment system was introduced to increase the level of efficiency of all producers in the Italian hospital industry. However, opposite to expectations and to the observed decline in ALOS, our findings seem to indicate that (technical) efficiency levels started to decline after the change in the reimbursement scheme, and this affected more private NFP hospitals than public producers. One possible explanation relies on the fact that the change in the reimbursement scheme went together with a significant reshaping of the hospital industry, guided by policies aimed at reducing hospitalization rates (see, e.g., [21]).

Indeed, when looking at aggregate data at the national level, the observed decline in ALOS is driven by the decrease in the number of inpatient days more marked than the reduction in the number of patients. However, while the number of beds started declining in 1998, the number of medical staff did not, thus resulting in a growth of workers per bed. This effect was more pronounced for private hospitals than for public ones. In other words, hospitals reduced their "capacity" while maintaining their personnel. Indirect evidence on this point is available by looking at capacity utilization rates: while private hospitals recorded an increase in excess capacity, public hospitals showed a decline. All these trends are replicated in our sample. Of course, the decline in outputs coupled with only a partial reduction in inputs offers an explanation to the estimated declining trend in (technical) efficiency.

A partial support to this argument – which however deserves a more specific and thorough investigation – can be obtained by studying the nature of the returns to scale. The observed change in the input-mix stems from a reduction of productive capacity, in terms of the number of beds. Holding the sample of hospitals fixed, this should have caused more producers to be operating in the region of increasing returns to scale. As for DEA, in the previous analysis we have been discussing pure technical efficiency results,

---

[26] It should be noted that year fixed effects cannot be included in the model since the DRG dummy is a linear combination of year fixed effects.

i.e., efficiency scores based on a VRS specification. We now decompose these scores according to Eq. (4) above; results are in Table 8.

Overall technical efficiency is now lower because of scale inefficiency. Indeed, scale efficiency, even if quite high (almost 0.9 overall), is lower than 1, and thus it lowers overall efficiency compared to pure technical efficiency. Notice as well that the hospitals operating in the area of increasing returns to scale are almost one sixth overall. However, they are more frequent in public hospitals and increasing over time, going from 14.8% before to 17% after the introduction of the new reimbursement scheme. On the other hand, in the case of NFP, we find a lower incidence of hospitals operating under increasing returns, and they are anyway decreasing over time: from 7.9% to 4.8% after the DRG-based payment mechanism introduction. Controlling for size, as expected all of the hospitals operating in the area of increasing returns to scale are the smallest ones[27].

[TABLE 8 ABOUT HERE]

Scale efficiency estimates can be derived also in the parametric setting, as shown for instance in [43] and [44]. Let us define $e^x = \sum_i \frac{\partial \ln D^O}{\partial \ln x_i}$ and $a = \sum_i \sum_j \alpha_{ij}$. In [44] it is illustrated that, at the Most Productive Scale Size, the statistic:

$$SE(x_0, y_0) = \exp\left[\frac{-(1 + e^x(x_0, y_0)}{2a}\right]$$

is equal to unity, since constant returns of scale hold locally. Table 9 shows our estimates obtained by considering the SF model. Confirming the general picture emerging from DEA estimates, very modest increasing returns

[27] The distribution of size is made according to the number of beds: small hospitals are those in the first quartile, medium hospitals are those in the second and third quartile, large hospitals are in the fourth quartile.

seems to emerge in the industry, with most hospitals operating at constant returns to scale. Moreover, we observe no change in mean scale efficiency before and after the introduction of the new payment system.

[TABLE 9 ABOUT HERE]

In conclusion, given the mild presence of scale economies, the reshaping of the hospital industry seems to have contributed to the decline in (technical) efficiency, because of the induced modification in the input-mix. From this point of view, one needs to discuss the technological properties of hospital services production function, before suggesting any policy. If labour and capital tend to be complements, then the input-mix change can be justified only by some (unmeasured here) quality increase of the services produced. Available evidence (see, e.g., [26]) seems to suggest that, at best, we did not observe a decline in quality. And this raises additional doubts about the effectiveness of the recently implemented policies toward hospitals in improving (technical) efficiency of producers. On this point, further analysis is needed.

5. Concluding remarks

In this paper we try to identify behavioral differences between public and private NFP hospitals, by exploiting the introduction of a PPS-DRG based in the Italian NHS in the second half of the Nineties. The introduction of the new payment system was aimed at increasing producers' efficiency, and thus controlling public spending growth. The relevant role of private nonprofit hospitals in Western European countries represents an excellent opportunity to test theories that deem nonprofit organizations as more efficient than public (or private for-profit) providers when asymmetric information and uncertainty prevail in a market.

In order to evaluate the impact of different ownership structures on hospital efficiency, we estimate the (technical) efficiency of a sample of hospitals for the period 1995-2000, by adopting both parametric and non-parametric approaches. We use DEA with an output oriented model for the non-parametric approach. We also follow a parametric approach using both COLS and SF techniques, and estimate a translog output distance function, to accommodate multiple inputs and outputs. Our results show a convergence of mean efficiency scores between NFP and public hospitals, and seem to emphasize the private nature of NFP hospitals, supporting the hypothesis that public and private nonprofit hospitals differ in their response to the introduction of the new payment system, with the latter responding more promptly than the former to PPS introduction.

Contrary to expectations, we also observe a decline in technical efficiency, probably due to policies aimed at reducing hospitalization rates. In other words, the observed reduction can be explained as the sum of two countervailing effects: on the one hand, the introduction of the DRG-based payment system might have improved hospitals efficiency; on the other hand, the reshaping of the hospital industry and the process of de-hospitalization – by reducing the number of patients - might have more than offset the gains in (technical) efficiency. One possible extension of the paper would be to disentangle these two effects to offer an evaluation of the effect of the DRG payment system of the two types of hospitals. This is left for future research.

REFERENCES

1. Alchian A. A. and Demsetz H. Production, Information Costs, and Economic Organization, *Am Econ Rev* 1972; **62 (5)**: 777-795.
2. Hansmann, H. The role of nonprofit enterprise, *Yale Law J* 1980; **89**.
3. Weisbrod, B. *The nonprofit economy*, Harvard University Press: 1988.
4. Arrow K. Uncertainty and the welfare economics of medical care, *Am Econ Rev* 1963; **53**: 941-973.
5. Glaeser, E.L. and Shleifer, A. Not-for-profit entrepreneurs, *J Public Econ* 2001; **81 (1)**: 99-115.
6. Hansmann, H. *The ownership of enterprise.* Harvard University Press: Cambridge MA, 1996.
7. Ballou J. P. and Weisbrod B. A. Managerial rewards and the behavior of for-profit, governmental, and nonprofit organizations: evidence from the hospital industry, *J Public Econ* 2003; **87**: 1895-1920.
8. Fried, H. O., Lovell, C. A. K., and Schmidt, S. S. (eds.) *The measurement of productive efficiency.* Oxford University Press: New York, 1993.
9. Sloan F. Not-for-profit ownership and hospital behavior, in Culyer A. J. and Newhouse J. P. (eds.), *Handbook of Health Economics* 2000; 1: 1141-1174.
10. Marmor T., Schlesinger, M., and Smithey, R. Nonprofit organizations and health care, in Powell W. (ed.), *The nonprofit sector. A research handbook.* Yale University Press: 1987.
11. Wilson, G.W. and Jadlow, J.M. Competition, profit incentives, and technical efficiency in the provision of nuclear medicine services, *Bell J Econ* 1982; **13(2)**: 472-482.
12. Vitaliano, D.F. and Toren, M. Hospital cost and efficiency in a regime of stringent regulation, *East Econ J* 1996; **22(2)**: 161-175.
13. Zuckerman, S., Hadley, J., and Iezzoni, L. Measuring hospital efficiency with frontier cost functions, *J Health Econ* 1994; **13**: 255-280.
14. Puig-Junoy, J. Technical efficiency in the clinical management of critically ill patients, *Health Econ* 1998; **7**: 263-277.
15. Duggan M. G. Hospital ownership and public medical spending, *Q J Econ* 2000; **CXV(4)**: 1343-1373.
16. Salkever, David S. Regulation of prices and investment of hospitals in the U.S., in Culyer A. J. and Newhouse J. P. (eds.), *Handbook of Health Economics* 2000; **1**: 1489-1535.
17. Coulam R.F and Gaumer G.L. Medicare's prospective payment system: a critical appraisal, *Health Care Financ Rev* 1991; **13,** Supplement.
18. Dafny, L. S. How do hospitals respond to price changes?, *Am Econ Rev* 2005; **95(5)**: 1525-1547.
19. Silverman E. and Skinner J. Medicare upcoding and hospital ownership, *J Health Econ* 2004; **23**: 369-389.

20. Bordignon M., Mapelli V. and Turati G. Fiscal Federalism and National Health Service in the Italian System of Governments, in *Annual Report on Monitoring Italy.* ISAE: Roma, 2002.
21. France G., Taroni F. and Donatini A. The Italian health care system, *Health Econ* 2005; **14**: S187-S202.
22. ISTAT *Censimento 2001. Imprese, Istituzioni e Unità Locali.* Roma, 2004.
23. Cellini, R., Pignataro, G., and Rizzo, I. Competition ad efficiency in health care: An analysis of the Italian case, *International Tax Public Finance* 2000; **7**: 503-519.
24. Barbetta, G.P. and Turati, G. L'analisi dell'efficienza tecnica nel settore della sanità. Un'applicazione al caso della Lombardia, *Econ Pubblica*, 2001; **XXXI**(2).
25. Bordignon M. and Turati G. *Bailing out expectations and health expenditure in Italy*, CESifo Working Papers, n. 1026 2003.
26. Louis D. Z., Yuen E. J., Braga M., Cicchetti A., Rabinowitz C., Laine C., and Gonnella J. S. Impact of a DRG-based hospital financing system on quality and outcomes of care in Italy, *Health Serv Res* 1999; **34 (1-II)**: 405-415.
27. Cerbo M. and Langiano T. L'impatto a livello nazionale del sistema di remunerazione, in Falcitelli N. e Langiano T. (a cura di), *Politiche innovative nel SSN: i primi dieci anni dei DRG in Italia.* Il Mulino – Fondazione Smith Kline: Bologna, 2004.
28. Banker, R. D., A. Charnes, and W. W. Cooper Some Models for Estimating Technical and Scale Inefficiencies in Data Envelopment Analysis, *Manage Sci* 1984; **30 (9)**: 1078-1092.
29. Debreu, G. The Coefficient of Resource Utilization, *Econometrica*, 1951; **19 (3)**: 273-292.
30. Koopmans, T. C. An Analysis of Production As an Efficient Combination of Activities, in T. C. Koopmans (ed.), *Activity Analysis of Production and Allocation*, Cowles Commission for Research in Economics. New York: Wiley 1951.
31. Koopmans, T. C. *Three Essays on the State of Economic Science.* McGraw Hill: New York, 1957.
32. Farrell, M. J. The Measurement of Productive Efficiency, *J R Stat Soc* Ser A Gen 1957; **120 (3)**: 253-281.
33. Färe, R. *Fundamentals of production theory.* Springer-Verlag: Berlin, 1988.
34. Chambers, Robert G. *Applied Production Analysis: A Dual Approach*. Cambridge University Press: New York, 1988.
35. Shephard, Ronald W. *Cost and Production Functions.* Princeton University Press: Princeton, 1953.
36. Färe, R., Grosskopf, S. and Lovell, C.A.K. *Production Frontiers.* Cambridge University Press: New York, 1994.
37. Coelli, T. and Perelman, S. Technical efficiency of European railways: a distance function approach, *Appl Econ* 2001; **32(15)**: 1967-76.

38. Jondrow J., Lovell C.A.K., Materov I.S., and Schmidt P. On the estimation of technical inefficiency in the stochastic frontier production function model, *J Econom* 1982; **19**: 233-238.
39. Barbetta, G. P., Turati G. and Zago A. M. The impact of DRG payment system on Italian hospitals, *Department of Economics Working Paper*, Università di Verona, n. 17-2004.
40. Banker, R. D. Hypothesis Tests Using Data Envelopment Analysis, *J Productivity Anal* 1996; **7 (2-3)**: 139-159.
41. Linna, M. Measuring hospital cost efficiency with panel data models, *Health Econ* 1998; **7**: 415-427.
42. Greene W. H. *Econometric Analysis.* Prentice Hall: 1998.
43. Atkinson S. E. and Primont D. Stochastic estimation of firm technology, inefficiency, and productivity growth using shadow cost and distance functions, *J Econom* 2002; **108**: 203-225.
44. Ray S. Measuring scale efficiency from the translog multi-input, multi-output distance function, *Department of Economics Working Paper Series*, n. 25, University of Connecticut 2003.

Table 1. Descriptive statistics
(Mean values; standard deviation in parentheses)

	NFP	*PUB*	*All*
Inputs			
Beds	402 (410)	277 (299)	282 (305)
Beds for DH	24 (34)	27 (38)	27 (38)
Physicians	146 (170)	107 (132)	108 (134)
Nurses	354 (416)	273 (319)	276 (324)
Teaching staff	3 (5)	1 (3)	1 (3)
Other personnel	33 (62)	24 (51)	25 (51)
Outputs			
Inpatient days	117726 (119941)	75355 (87419)	77030 (89290)
Discharged patients	15599 (14890)	10429 (10855)	10634 (11086)
DH treatments	10215 (17530)	9460 (18826)	9484 (18785)
Emergency room treatments	24609 (23001)	39452 (143475)	38901 (140885)
Nr. Obs.	*126*	*3060*	*3186*

Table 2. ALOS by years and type of hospitals

	1995	*1996*	*1997*	*1998*	*1999*	*2000*	*Before DRG*	*After DRG*
All hospitals	7.80 (1.81)	7.41 (1.81)	7.10 (1.75)	7.02 (1.77)	7.37 (2.25)	7.42 (2.50)	7.44 (1.81)	7.27 (2.20)
NFP	8.80 (1.53)	8.08 (1.21)	7.61 (1.31)	7.39 (1.57)	7.71 (1.78)	7.46 (1.61)	8.16 (1.42)	7.52 (1.64)
Public	7.76 (1.81)	7.39 (1.82)	7.08 (1.77)	7.00 (1.77)	7.35 (2.27)	7.41 (2.53)	7.41 (1.82)	7.26 (2.22)

Mean values; standard deviation in parentheses

Table 4. Efficiency scores by year: output distance function (all hospitals)

Methodology	*1995*	*1996*	*1997*	*1998*	*1999*	*2000*	*Before DRG*	*After DRG*	**Test (§)** *MW*	*KS*
DEA	0.64	0.66	0.68	0.69	0.62	0.63	0.66	0.65	2.41	0.06
	(0.18)	(0.16)	(0.17)	(0.17)	(0.17)	(0.17)	(0.17)	(0.17)	[0.02]	[0.01]
COLS	0.61	0.61	0.61	0.61	0.61	0.61	0.61	0.61	0.43	1.22
	(0.09)	(0.07)	(0.08)	(0.08)	(0.09)	(0.09)	(0.08)	(0.09)	[0.68]	[0.10]
SF	0.80	0.83	0.82	0.82	0.74	0.72	0.82	0.75	3.64	2.49
	(0.22)	(0.15)	(0.17)	(0.21)	(0.29)	(0.32)	(0.18)	(0.28)	[0.00]	[0.00]

Mean values; standard deviation in parentheses.
(§) Mann-Whitney test for equality of means and Kolmogorov-Smirnov test for equality of distributions before and after DRG introduction; p-values in parentheses.

Table 5. Efficiency scores by year and ownership type: output distance function (DEA estimates)

									Test (*)	
Type	***1995***	***1996***	***1997***	***1998***	***1999***	***2000***	***before DRG***	***after DRG***	**MW**	**KS**
NFP	0.73	0.74	0.76	0.78	0.67	0.65	0.74	0.70	1.61	0.21
	(0.15)	(0.15)	(0.16)	(0.17)	(0.16)	(0.15)	(0.15)	(0.17)	[0.11]	[0.10]
PUB	0.64	0.66	0.67	0.69	0.62	0.63	0.66	0.65	2.12	0.05
	(0.18)	(0.16)	(0.17)	(0.17)	(0.17)	(0.17)	(0.17)	(0.17)	[0.03]	[0.02]
MW (§)	2.36	2.3	2.05	2.14	1.13	0.66	3.85	2.12	[redacted]	[redacted]
	[0.02]	[0.02]	[0.04]	[0.03]	[0.26]	[0.51]	[0.00[	[0.03]	[redacted]	[redacted]
KS (§)	0.31	0.31	0.24	0.25	0.16	0.16	0.24	0.15	[redacted]	[redacted]
	[0.03]	[0.04]	[0.14]	[0.10]	[0.56]	[0.61]	[0.00]	[0.08]	[redacted]	[redacted]

Mean values; standard deviation in parentheses.
(§) Mann-Whitney test for equality of means and Kolmogorov-Smirnov test for equality of distributions between different types of hospitals; p-values in parentheses.
(*) Mann-Whitney test for equality of means and Kolmogorov-Smirnov test for equality of distributions before and after DRG introduction; p-values in parentheses.

Table 6. Efficiency scores by year and ownership type: output distance function (COLS and SF estimates)

										Test (*)	
Methodology	*Type*	***1995***	***1996***	***1997***	***1998***	***1999***	***2000***	***before DRG***	***after DRG***	**MW**	**KS**
SF	NFP	0.93	0.93	0.92	0.91	0.62	0.46	0.92	0.60	4.14	2.17
		(0.03)	(0.03)	(0.03)	(0.05)	(0.41)	(0.46)	(0.03)	(0.42)	[0.00]	[0.00]
	PUB	0.79	0.83	0.82	0.81	0.74	0.72	0.82	0.76	3.03	2.29
		(0.21)	(0.15)	(0.18)	(0.21)	(0.28)	(0.31)	(0.18)	(0.27)	[0.00]	[0.00]
	MW (§)	3.57	3.04	2.77	2.21	0.49	1.73	5.41	0.85	[redacted]	[redacted]
		[0.00]	[0.00]	[0.00]	[0.03]	[0.63]	[0.08]	[0.00]	[0.40]	[redacted]	[redacted]
	KS (§)	1.85	1.80	1.53	1.54	0.77	1.51	2.71	1.51	[redacted]	[redacted]
		[0.00]	[0.00]	[0.02]	[0.02]	[0.59]	[0.02]	[0.00]	[0.02]	[redacted]	[redacted]
COLS	NFP	0.70	0.68	0.68	0.65	0.61	0.63	0.68	0.63	2.89	1.45
		(0.07)	(0.08)	(0.07)	(0.07)	(0.05)	(0.06)	(0.07)	(0.06)	[0.00]	[0.03]
	PUB	0.61	0.60	0.61	0.61	0.61	0.61	0.61	0.61	0.16	1.20
		(0.09)	(0.07)	(0.08)	(0.09)	(0.09)	(0.09)	(0.08)	(0.09)	[0.87]	[0.11]
	MW (§)	2.85	2.40	2.30	1.54	0.34	1.14	4.42	1.58	[redacted]	[redacted]
		[0.00]	[0.02]	[0.02]	[0.12]	[0.73]	[0.25]	[0.00]	[0.11]	[redacted]	[redacted]
	KS (§)	1.46	1.34	1.50	0.98	0.72	0.93	2.25	1.28	[redacted]	[redacted]
		[0.03]	[0.05]	[0.02]	[0.30]	[0.67]	[0.35]	[0.00]	[0.08]	[redacted]	[redacted]

Mean values; standard deviation in parentheses.
(§) Mann-Whitney test for equality of means and Kolmogorov-Smirnov test for equality of distributions between different types of hospitals; p-values in parentheses.
(*) Mann-Whitney test for equality of means and Kolmogorov-Smirnov test for equality of distributions before and after DRG introduction; p-values in parentheses.

Table 7. Second stage analysis (Tobit models)

Dep. Var.: EFF scores		*SF*			*COLS - w/o FE*			*COLS - FE*	
DRG	-0.04***			-0.02***			-0.03***		
	(0.010)			(0.007)			(0.007)		
DRG*PUB		-0.03***	0.004		-0.017**	-0.0004		-0.02***	-0.015
		(0.010)	(0.02)		(0.007)	(0.014)		(0.007)	(0.014)
DRG*NFP		-0.214***	-0.167***		-0.144***	-0.124***		-0.15***	-0.14***
		(0.041)	(0.046)		(0.028)	(0.032)		(0.028)	(0.032)
PREXP	0.04***	0.04***	0.04***	0.025***	0.026***	0.026***	0.02***	0.02***	0.02***
	(0.005)	(0.005)	(0.005)	(0.004)	(0.004)	(0.004)	(0.003)	(0.003)	(0.003)
PREXP*DRG			-0.004**			-0.001			-0.001
			(0.002)			(0.001)			(0.001)
Nr. obs.	2572	2572	2572	2572	2572	2572	2572	2572	2572
Log-L.	-569.20	-559.30	-556.89	279.96	289.96	290.94	307.22	317.55	317.91
LR test (a)		19.80	24.63		20.01	21.97		20.66	21.38

MLE. Asymptotic t-ratios in parentheses. Regional fixed effects included in all regressions. Lev. of significance: *** 1%, ** 5%, * 10%.

(a) Test equality of DRG*PUB and DRG*NFP. Critical values Chi-sq (1).

Table 8. Decomposition of DEA efficiency scores

	No. obs.	*Overall technical efficiency*	*Scale Efficiency*	*Pure technical efficiency*	*No. obs. (%) with IRS*
Whole period (1995-2000)					
All	*3186*	*0.568*	*0.881*	*0.655*	*494 (15.5%)*
Public	3060	0.566	0.882	0.652	486 (15.9%)
NFP	126	0.603	0.855	0.721	8 (6.3%)
Before DRG (1995-1997)					
All	*1593*	*0.573*	*0.882*	*0.662*	*231 (14.5%)*
Public	1530	0.571	0.883	0.658	226 (14.8%)
NFP	63	0.617	0.847	0.742	5 (7.9%)
After DRG (1998-2000)					
All	*1593*	*0.563*	*0.881*	*0.648*	*263 (16.5%)*
Public	1530	0.562	0.882	0.646	260 (17%)
NFP	63	0.589	0.863	0.699	3 (4.8%)
By dimension					
Small	2904	0.572	0.906	0.638	494 (17%)
Medium	258	0.524	0.644	0.817	0 (0%)
Large	24	0.489	0.526	0.928	0 (0%)

Table 9. Scale efficiency estimates from SF model

	Mean	*Std. Dev.*	*Min.*	*Max.*
All hospitals	1.002	0.003	1.000	1.026
By dimension				
Small hospitals (1st quartile)	1.003	0.003	1.000	1.026
Medium hospitals (2nd-3rd quartile)	1.001	0.002	1.000	1.012
Large hospitals (4th quartile)	1.000	0.0007	1.000	1.002
By ownership types				
NFP	1.004	0.006	1.000	1.023
Public	1.002	0.003	1.000	1.026
By years				
Before DRG	1.003	0.003	1.000	1.021
After DRG	1.003	0.004	1.000	1.026

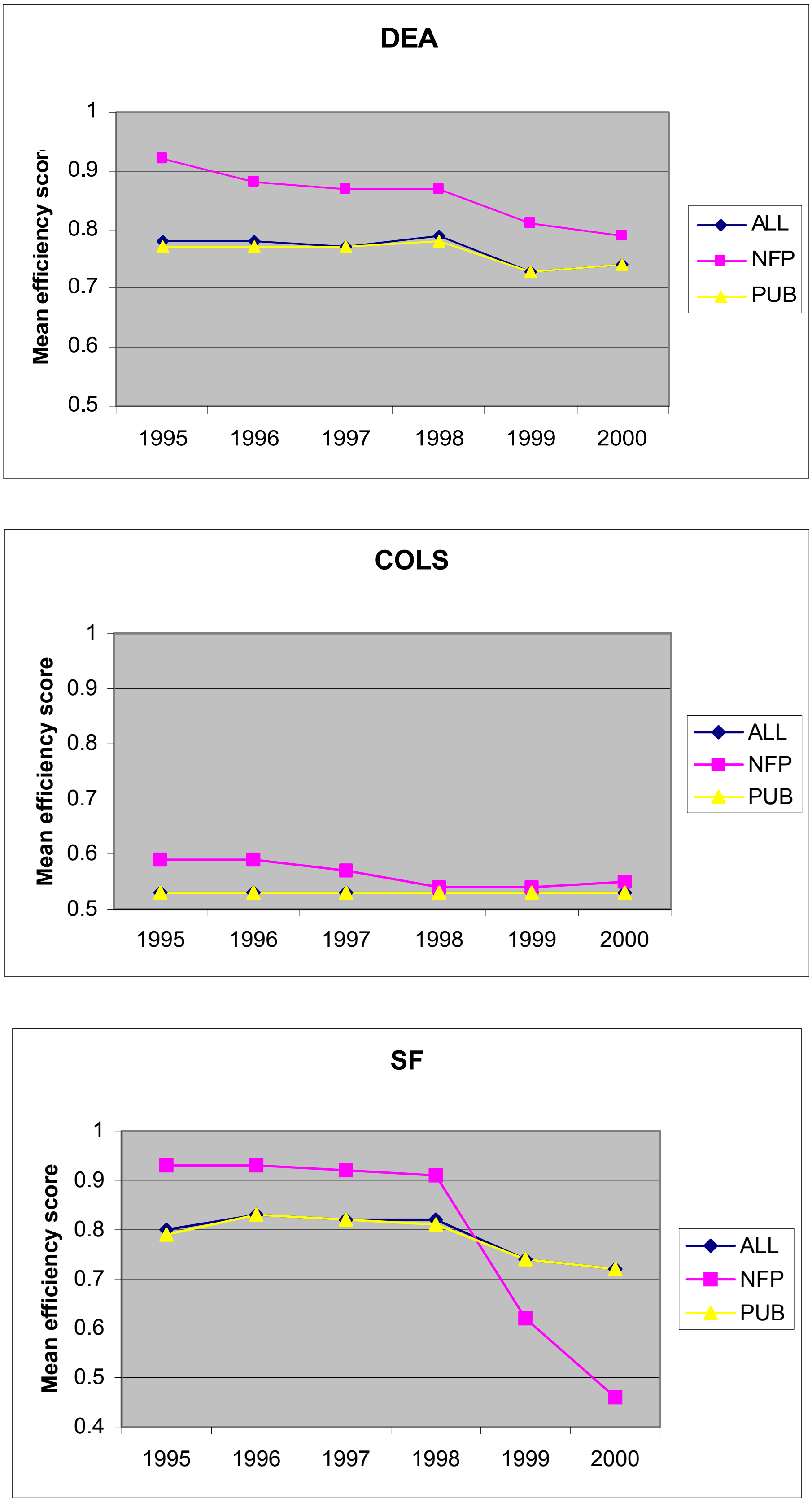

Figure 1. Mean efficiency scores