INVESTIGACIONES ECONOMICAS. Volumen XX (3) - Septiembre, 1996, pp. 377-402

CHANGES IN THE EFFICIENCY OF SPANISH PUBLIC HOSPITALS AFTER THE INTRODUCTION OF PROGRAM-CONTRACTS*

Beatriz GONZALEZ LOPEZ-VALCARCEL
Patricia BARBER PEREZ
Universidad de Las Palmas de Gran Canaria

This paper studies the evolution of the technical efficiency in Spanish public hospitals during the period 1991-93 using Data Envelopment Analysis, and discusses the determinants of efficiency, as well as the relationship between technical efficiency and unit costs. A cost frontier model is also estimated in order to get an alternative measure of efficiency. Both procedures show significant efficiency gains after the introduction of program-contracts in 1992.

1. Introduction

In Spain, as in most OECD countries, there is a concerted effort to make the public health system more efficient. With new systems of control and a decentralization of responsibility, government authorities have attempted to overcome the lack of financial discipline and the chronic underfunding that has led to substantial budget deficits. In recent years, hospital costs have continued to go over budget, and Spain's central government, more or less regularly, has had to step in and cover the difference (most recently in 1989, 1992, and 1994). On these occasions, the central government, after considerable delay, allotted a lump-sum equivalent to the amount of cost overrun in *Insalud Gestión Directa* to each regional government with responsibility for health care[1]. The lack of financial discipline in hospital costs has major repercussions on government budgets, and has perverse effects on the administration of public hospitals. Furthermore, there is a general perception that technical inefficiency is widespread and that, even holding steady or increasing budgets very little, public hospitals could render more and better service.

Hospital cost inflation and the perception of technical inefficiency have led to certain changes in the relations between funder and hospital. *Insalud Gestión*

* This paper forms part of a group of studies coordinated by FEDEA and supported by the Bank of Spain. We thank the anonymous referees for useful comments.

[1] *Insalud Gestión Directa* is the public agency which regulates, finances and administrates the provision of health care in the Spanish regions without responsibility in health care. In 1993, it included 11 out of the 17 Spanish regions and amounted around 50% of the public expenditure.

Directa and the Health Authorities of Galicia and Andalucía have attempted to address both problems through «program-contracts» that provide an incipient delimitation of roles and risks between funder and providers, and replace the global retrospective budgets of the past with budgets calculated with fees progressively disagregated, refined, and based on a given unit of payment. These changes attempt to stimulate competition between internal units and between hospitals. As a result of the incentives built into the new funding systems, the authorities hope for increased efficiency, greater control of expenditure, and reactive synergies such as cooperation and specialization among neighbouring hospitals.

In the last decade experts in health economics and management have come to a consensus about ways to improve efficiency and control costs: divide the functions of finance, purchase and provision of health-care services among different public agencies or institutions; and distribute in a balanced way financial risk between the funding agency and the hospital, even if the hospital is public. In the earlier system, which covered de facto all the expenditures of the hospital with a budget that no one respected, the hospital assumed no risk. The consensus is that the old system of retrospective payment should gradually be replaced by a system in which payment is prospective. The budget of a particular hospital should cover only justifiable expenditures, and every hospital should commit itself to goals of quantity and quality for particular activities, these goals negotiated with the funder/purchaser. These commitments are called «program-contracts».

The program-contract of a particular hospital contains its budget, which is calculated by applying to production goals in physical units prices adjusted to estimated average costs for a sample of hospitals. For historical reasons or because of legal restrictions on management, some hospitals maintain in their budgets a subsidy in excess of the average costs, but every effort is made to reduce these deficits, and over time they should tend to disappear. At the time of this study there had been two program-contracts, that of 1992, which was tentative and experimental, and that of 1993.

Program-contracts measure hospital production in weighted care units, called UPAs *(Unidades Ponderadas Asistenciales)*, based on length of stay and type of service. For example, a one-day stay on an internal medicine ward equals one UPA, but a day in intensive care would be 5.8 UPAs; or an outpatient emergency visit would be 0.3 UPAs, and a first walk-in clinic visit would be 0.25 UPAs. Certain expensive high-tech surgical procedures, like transplants, and some other procedures have been removed from the UPA payment system and have a fee for service basis. If their fee is above or below the marginal cost, the hospital has incentives to increase or decrease the number of these procedures.

The main attraction of program-contracts is that they make health planning easier by permitting a link between hospital activities and the epidemiological objectives previously stated in the health agenda. Furthermore, they allow planners to even out the level of health care across regions. The most serious worry is whether the incentives in the program-contracts are sufficient to induce improvement in efficiency and financial discipline in public hospitals.

In Section 2, after some methodological considerations and background, we use Data Envelopment Analysis (DEA) to evaluate changes in technical efficiency of *Insalud Gestión Directa* hospitals over the years 1991-1993 (before, during, and after the introduction of the program-contracts). Furthermore we evaluate the share of inefficiency due to pure technical reasons –overusing inputs to produce given amounts of outputs– as well as the share due to scale problems –hospitals not having the optimal size–.

DEA has some limitations. The most serious is its sensitivity to outliers in the sample. In order to overcome this problem, we have designed and used different methods to validate DEA results, including statistical tests. We have computed and compared the scores of relative technical efficiency through some restricted DEA models, which try to homogenize the sample of hospitals by adding additional constraints to the initial model. In Section 2.4 we look for the determinants of DEA scores of efficiency with a censored regression model. The relationship between these scores and unit hospital costs is analyzed in Section 2.5. Finally, Section 3 specifies and estimates a cost frontier model with the panel sample of hospitals to get an alternative measure of efficiency –a mixture of technical and allocative efficiency– and analyze its evolution in these years.

The study applies both methods –data envelopment models and stochastic frontier models– to data provided by the *Dirección General de Aseguramiento* of Spain's Ministry of Health for 75 general hospitals run by *Insalud Gestión Directa*. We found an increase in efficiency after adoption of program-contracts, although we cannot attribute this increase necessarily to the change in the method of financing.

2. A data envelopment analysis of the evolution of Hospital Efficiency in Spain

2.1. Introduction

Technical efficiency consists in maximizing output with a given quantity of input, or minimizing input while maintaining a given level of output. One seeks to estimate a frontier function of production or costs and measure individual deviations from the frontier. The frontier of production expresses in physical units the maximum quantity of output obtainable from given amounts of inputs. The output effectively produced by the firm or *Decision Making Unit* (DMU) i may be below its frontier, and the difference is the absolute technical inefficiency.

In practice, the estimation of frontiers uses only sample information from the units whose efficiency we want to establish. As a result, what in fact one calculates is the «frontier of best practice», in Farrell's words, and not the frontier of absolute possibilities of production with existing technology. The most efficient unit in the sample automatically defines the reference level of efficiency as 100% efficient.

There are a variety of methods to estimate frontiers and measure inefficiencies. Data Envelopment Analysis (DEA) involves models with a deterministic and non-parametric frontier. They were developed by Farrell (1957) and later refined by Charnes, Cooper and Rhodes (1978). A modern introduction to DEA models is provided by Lovell (1994). The models are based on the solution of problems of mathematical programming to find the most efficient subset of production units and measure the relative efficiency of each unit. They are non-parametric models since they do not require a given functional form of the transformation function of inputs into outputs. The only hypothesis required is that the space of the technology of production is convex; that is, that a weighted sum of inputs and outputs of any two hospitals would result in a «compound hospital» with plausible technology. There are two main kinds of DEA models, those that define radial or non-radial measures of productive efficiency. The former may be input-oriented or output-oriented, and may or may not allow variable returns to scale. Our baseline model is radial and input-oriented. It was originally proposed by Charnes, Cooper and Rhodes (CCR model), and specifies that the DMU*i* under evaluation seeks the minimum feasible radial scaling of m inputs (vector x_i) to produce given amounts of k outputs (vector y_i):

$$Z_i \equiv Min_{z,\,\lambda \geq 0}\, Z \text{ subject to } ZX_i \geq X\lambda \text{ and } Y\lambda \geq y_i, \qquad [1]$$

where X is the $m \times n$ input matrix (n is the number of DMU in the sample) and Y is the $k \times n$ output matrix. The minimum value of z, Z_i, provides a radial measure of technical efficiency for the DMU*i*. It is judged to be inefficient if $z_i < 1$, and efficient if $z_i = 1$. For the inefficient DMUs, the $n \times 1$ intensity vector λ provides the weights of the technically efficient DMUs enveloping the DMU*i* and drawing its «best practice» frontier. Therefore model [1] is called *envelopment problem*. Its dual problem is the *multiplier problem*, which gives the optimal weights or normalized shadow prices for the inputs and outputs of the DMU*i* being evaluated.

Model [1] involves a technology with constant returns of scale. A modified version is the BCC (Banker, Charnes and Cooper) model, which allows for variable returns of scale. To measure *pure technical efficiency*, this model includes in [1] the constraint that the components in λ must add up to one. In this way it is guaranteed that the DEA model evaluates pure technical efficiency without including consideration of scale. Once the global technical efficiency (GTE) and the pure technical efficiency (PTE) have been determined through the CCR model and the BCC model respectively, we can calculate the *efficiency of scale* = GTE/PTE, which measures that portion of inefficiency due exclusively to problems of size.

We use input-oriented models because they fit better than output-oriented models the behaviour of public hospitals, whose output is exogenous to the extent that its level is determined by the demand in the area and fixed among the objectives negotiated in the program-contract. Hospitals try to produce these services with minimum levels of inputs, and we attempt to estimate the savings in input levels obtainable if all the hospitals were efficient.

Recent DEA models define non-radial measures of efficiency: some of the inputs can be reduced, but not all of them in the same proportion. We do not yet know any applications of these models to hospitals.

One of the principal benefits of DEA models is that they do not require the definition of a production function that describes the process of transformation of inputs into outputs. Nevertheless, the formation of the frontier is very sensitive to outliers, and hence the DMUs in the sample should be homogeneous and one should check for the presence of skewed observations and of outliers. Wilson (1993, 1995), Burgess and Wilson (1996) and Ley (1996) are useful references for detecting outliers and influentional observations in DEA.

The DEA models are flexible enough to accommodate easily solutions to their own limitations. Thus Banker and Morey (1986) added capital as fixed input, constraining the set of possibilities of production so as not to permit variation in capital. In some models, environmental variables are added that can affect the set of possibilities of production, making comparisons more difficult. These variables act as zero-priced, uncontrolled inputs.

The general problem is to find a way to control for the heterogeneity of DMUs. A common solution is to add constraints that limit the set of hospitals to those structurally comparable to the one being analyzed and to those that operate in a similar setting. Our DEA specification follows that approach.

2.2. Antecedents: The application of DEA to the study of hospital efficiency

Rosko (1990) summarizes the applications of DEA to measure the efficiency of different organizations providing health services. Table 1 lists those after 1989 and others not mentioned by Rosko referring to Spain. These studies demonstrate that the specification of the model strongly depends on the objectives of the analysis and the availability of information. Objectives may be more or less ambitious, from the fixing of a budget and planning for a region (Guey and Huang, 1990) to the measurement of cost savings that would be achieved by increasing technical efficiency to the level of the hospitals on the frontier (Sexton *et al.*, 1989).

The analyses differ as to the selection of the set of hospitals in the sample, the definition of hospital inputs and outputs, and to a lesser extent, the specification of additional constraints that increase the homogeneity of the sample. In general an attempt is made to see that the hospitals are structurally comparable and operate in similar environments. Ehreth (1994) groups hospitals according to size, teaching status, and region; Ozcan *et al.* (1992) and Ozcan and Luke (1993) group them by metropolitan areas. Other analyses limit the number of beds to a certain range or the sample to a certain city or specific region, where it can be understood to share an environment (Banker, Conrad and Strauss, 1986; Regnier *et al.*, 1993; Valdmanis, 1990).

For the definition of inputs and outputs, there are a great variety of options. A tripartite classification of resources (doctors, beds, operating rooms...), services (stays, admissions, operations...), and health, has led to the following typology of applications (Regnier *et al.*, 1993): a) resources produce services; b) services produce health; and c) resources produce health. The first formulation is by far the most prevalent, although it is also conceptually the least useful. The result of these analyses could systematically overestimate efficiency as far as the same measures of inputs and outputs are used to calculate ratios of productivity and to evaluate and control the achievement of the objectives. Some studies place constraints on the generic approach we have called type a), adjusting output by its quality or ensuring a given level of quality (Morey *et al.*, 1990). But most empirical studies measure quality negatively, as a worsening of health (the rate of iatrogenic infections, hospital mortality as a standardized rate, or the frequency of readmissions). There is barely any literature with an analysis of type c) (resources provide health), which we consider the most appropriate conceptually. The few attempts to analyze models of all three types for the same group of hospitals have produced conflicting results (Regnier *et al.*, 1993), although this may be due to methodological problems in the definition of variables.

The inputs in type a) and c) analyses are financial resources, physical resources, or a combination of the two, depending on the objectives of the analysis and the information available. The introduction of financial resources leads to a combined estimate of technical and allocative efficiency, thus sacrificing one of the potential advantages of the DEA models as compared to stochastic cost frontiers. Some variables, like the number of physicians in a medical training program (MIR), alternately act as inputs and outputs (Valdmanis, 1990).

The applications that have been published have generally analyzed cross-section information for a given year. Only Ehreth (1994) and more recently Burgess and Wilson (1996) analyze panel samples.

Authors validate their results by comparing them with those of alternative DEA models or with other methods like the traditional analysis of financial ratios (Ehreth, 1994). The correlation between the DEA measure of global efficiency of each hospital and the unit costs has also been used (Valdmanis, 1990; Nunamaker, 1983; Hogan *et al.*, 1987; Huang, 1990).

Some studies compute a DEA measure of efficiency and explain *a posteriori* its determinants using linear regression models, whether probit (Sexton *et al.*, 1989) or censured and truncated (Vitaliano and Toren, 1994).

For the Spanish health sector, Ley (1991) uses DEA to obtain individual measures of the productive efficiency of different hospitals. Prior and Solá (1993) use it to analyze the efficiency of the public and private hospitals of Catalonia.

Criticism of the application of DEA to hospitals points to various drawbacks. Its deterministic frontiers confuse statistical noise with inefficiency and impede the use of tests of significance and measures of goodness of fit. Since they are frontiers «of best practice», they are vulnerable to the presence of outliers[2]. Furthermore, since the analysis is non-parametric the number of hospitals and the number of variables condition the frequency of efficient units. Other things being equal, the larger the sample, the fewer hospitals will show up as efficient, and a greater number of variables will lead to more hospitals appearing efficient. For this reason one cannot make direct comparisons between the results of two DEA that correspond to different groups of units or that use different variables.

It may also be that the hypothesis of the convexity of the function that transforms inputs to outputs is unrealistic (Rosko, 1990). In any case, this hypothesis is required as well in conventional econometric models, forcing as well more rigid functional forms. And if necessary, a data envelopment model can be relaxed by adding constraints or formulating a problem with quadratic programming.

A more serious criticism refers to the choice and definition of inputs and outputs and how to control among hospitals for quality of care, for severity of the illnesses of the patients, and for the supplementary services of a hotel variety (Newhouse, 1994). One must refine these measures, which many times when aggregated are too crude. Nevertheless the possibilities of disagregation are limited because there must be a balance between the number of hospitals and variables in the model[3]. As the population size is finite (*Insalud Gestión Directa*, for instance, has 89 hospitals) and the hospitals being compared must be homogeneous, the only solution is to limit the disagregation of factors and products within reason.

2.3. *Empirical Application: Results*

Our objective is to evaluate with data envelopment models the evolution of technical efficiency in the general hospitals of *Insalud Gestión Directa* from 1991 to 1993.

By means of cluster analysis, Insalud has classified its 89 hospitals in 5 homogenous groups, taking into account number of beds, size of medical staff, number of medical specialities, high technology facilities etc.[4] This classification is widely accepted and robust. It ranges from the smallest hospitals in group 1 to the largest in group 4. Group 5 is comprised of atypical hospitals.

Our analysis moves from the general to the particular as we resolve three problems (P1, P2 and P3) sequentially to measure the global technical efficiency of each hospital every year. The three analyses have the same objective function as in [1]; only their constraints vary.

[2] We wish to emphasize, however, that this vulnerability to outliers is absolutely asymmetric, as it applies to hospitals that turn out to be efficient. Their effect, hence, would at worst introduce a bias toward underestimating the level of efficiency in comparison to the rest of the hospitals, never the reverse. A good symptom that this bias does not occur is if a high number of hospitals constitute the envelopment frontier of each hospital that turns out to be inefficient.

[3] Otherwise, one runs the risk of finding that all the hospitals to be efficient, except for those completely dominated (that is those that with more of every input produce less of every output).

[4] For technical details see Insalud: Program Contract 1993.

Table 1
Applications of Data Envelopment Analysis to the measure of hospital efficiency

	Number of Hospitals	Inputs	Outputs	Specific objetives, main results and observations
Ehreth (1994)	Not given. Separate analysis by groups	Various alternative specifications are tested. In one of them: – Fixed assets – Number of full-time equivalent employees	In the last of the specifications: – Medicare discharges – Medicaid discharges – Other discharges (All discharges adjusted for case-mix) – Outpatient consultations as discharge equivalents	– Evaluate the effect of the payment on financial performance and technical efficiency – Calculate the correlations among the different DEA measures of efficiency and measures of financial performance – Analyses on hospitals grouped by size, teaching status, and region
Ley (1991)	139 Spanish general hospitals with 70 to 700 beds	– Doctors including MIR – Staff with a technical degree – Other personnel – Purchases (in millions of pesetas) of sanitary supplies – Beds (as a proxy for capital)	– Patient-days in hospital – Discharges because of recovery (from medicine, surgery, obstetrics, pediatrics, and intensive care) – Patient-days on other wards (psychiatry, tuberculosis, long-term) – All emergency ward cases – Operations – Newborns	– Result: private hospitals are on average more efficient, but vary more than the public hospitals – No significant differences in technical efficiency between teaching and non-teaching hospitals
Morey *et al.* (1992)	300 United States hospitals with 20 to 918 beds	Current expenses in six categories (does not include doctors)	– Discharges – Births – Abnormal newborns – Patient-days in rehabilitation – Patient-days in sub-acute stays – Emergency ward cases – Outpatient visits – Ambulatory surgical operations – Production of medical education, in dollars	– Analysis of the quality-cost relation (elasticity) – Use of the DEA model to identify the minimal costs for a given level of quality – Quality measured by an index of real deaths over predicted deaths in the hospital

Table 1
(Continuation)

	Number of Hospitals	Inputs	Outputs	Specific objetives, main results and observations
Ozcan *et al.* (1992) Ozcan and Luke (1993)	All urban hospitals in the United States not run by the federal government (3000), by metropolitan area (82)	– Number of diagnostic and specialized services (complexity of services) – Beds – Full-time personnel – Current expenses	– Discharges adjusted by case-mix – Outpatient visits – Number of persons in full-time medical training	– Compare technical efficiency for public, nonprofit, and for-profit hospitals, controlling for region, size, and affiliattion – Results: Public hospitals are in general more efficient, but less efficient in their use of the labor factor
Prior and Solá (1993)	169 hospitals in Catalonia (1989)	– Full-time doctors and others professionals – Full-time nursing personnel with technical degrees – Other full-time personnel – Beds – Purchases in pesetas, as a proxy for consumption of material in physical units	– Patient-days in medicine, surgery, obstetrics and gynecology, pediatrics, and other services – Discharges in these five categories – Outpatient visits, including emergency ward – Other activities (test, rehabilitation, chemotherapy...)	– Measure global technical efficiency and scale – Compare public and private hospitals – Results: no significant differences in technical efficiency between public and private hospitals
Regnier *et al.* (1993)	76 Quebec hospitals with more than 100 beds for acute cases and more than 50 beds for medicine and surgery	Tests three kinds of models 1. Resources 2. Resources 3. Services	 → Health → Services → Health	– Measure relative efficiency by comparing the results of different models – Results: different models produce very different results
Valdmanis (1990)	41 Michigan hospitals (8 public, 33 private) with more than 200 beds	– Doctors – Residents – Non-medical staff in full-time equivalents – Capital assets	– Residents – Patient-days in hospital – Patient-days on intensive care – Operations – Outpatient visits and emergency ward	– Compare the efficiency of public and private hospitals – Determine the factors that affect efficiency – Residents are considered both inputs and outputs by the model – Results: Public hospitals are more efficient; private hospitals use more complex technology

Table 1
(Continuation)

	Number of Hospitals	Inputs	Outputs	Specific objetives, main results and observations
Burguess, J. F. and Wilson, P. W. (1993)	137 non-psychiatric hospitals in theVeteran Administration System. Panel data	– Number of acute-care hospitals beds – Number of long-term hospitals beds – Clinical labor (full-time equivalents) – Non-clinical labor (full-time equivalent) – Physician hours	– Inpatient days – Number of inpatient discharges – Outpatient visits. – Ambulatory surgical procedures – Inpatient surgical procedures	– Statistical test for outliers
Yueh Guey and L. Huang (1990)	213 Florida hospitals	– Beds in acute and intensive care wards – Full-time personnel – Case-mix index (to control for severity of cases) – Service-mix index (to control for the variety of services in the hospital) – Capital assets as a proxy for technology)	– Outpatient visits – Patient days, adjusted – Admissions, adjusted	Compare average costs of efficient and inefficient hospitals – Purpose: to automate the process of budget revision in Florida's hospitals

Of the 89 hospitals in the Insalud public network we have excluded the 14 group 5 hospitals that are in some way exceptional, usually because of a particular specialization. For each of the remaining 75 hospitals we consider data for 1991, 1992, and 1993, a total sample size of N=225 observations. We calculated for each of the observations global technical efficiency, pure technical efficiency, and efficiency of scale including the kind of returns to scale. We used 3 inputs and 11 outputs and solved 675 problems of linear programming [1] with 225 variables and 14 constraints each.

Our sample contains n_{gt} observations of hospitals of group g during year t (g=1,2,3,4; t=1,2,3), so that the total panel contains 225 decision-making units (75 hospitals in 4 structural groups for 3 years). The first analysis (P1) is unconstrained, permitting intergroup and interperiod comparisons. The solution for these 225 P1 problems provides the unconstrained intensity vector $\lambda_i^{(1)}$. Hence their reference units compare each hospital with the «best» hospitals in any of the three years. These are the most absolute frontiers. The second analysis (P2) compares each hospital i in a given year with the hospitals of its group in any of the three years. It is thus constrained, allowing interperiod but not intergroup comparisons. For this reason, we added constraints so that only the hospitals in hospital i's group are included. The P3 problems permit intergroup comparisons of each hospital in a given year with all the others, but only for that year.

The DEA models P1 and P2, comparing the same DMUs in different years, presuppose that the production frontier does not vary over the period. Otherwise one would be generating an index that was a meaningless hybrid, reflecting technological changes as well as efficiency. Three years, we think, is a period short enough for this hypothesis to be reasonably applied.

Hence the solution of the three analyses provides for each hospital and year three values $z_i^{(p)}$ of global technical efficiency (p=1,2,3). By definition, $z_i^{(1)} \leq z_i^{(2)}$ and $z_i^{(1)} \leq z_i^{(3)}$, but the relations among the optimal values of efficiency that result from the P2 and P3 analyses are not known *a priori*. For the hospitals that show up as inefficient, the models give three different intensity vectors $\lambda_i^{(p)}$ size 225, but because of the constraints, the vectors can coincide only in the values of the n_g elements of their group and year. The inter-year and inter-hospital comparisons, discussed above in reference to the P1 analysis, are qualified and validated by the results of the other analyses.

We present the inputs and outputs used in the analysis with their descriptive statistics in Table 2. Definition and measurement of what it is that hospitals produce continues to be a matter for debate. The inputs of this study seek to reflect, on the one hand, the structure of the hospital capital, and on the other, labour. To quantify the outputs we used the same measures of hospital production as in the program-contracts. These measures distinguish care requiring hospitalization from care on an out-patient basis. In addition to patient-days in different services, ambulatory surgical procedures, operations with hospitalization, and hospital admissions, we have included two novel outputs: UPAMIX is an index of ambulatory visits and emergency activities. It weighs (using the weights defined by the UPA's) all emergency ward treatment that does not include hospitalization (equivalent to 0.3 UPA's), as well as the number of first and successive outpatient visits (with values of 0,25 and 0,15 UPA's respectively); TECNO represents the weighted sum, according to fees, of high-tech activity (kidney, heart, and liver transplants; kidney and multi-organ removals; hemodynamic diagnosis and therapy; dialysis). Both UPAMIX and TECNO use the weights adopted in the program-contracts to measure hospital activity. These weights were determined in an earlier Insalud study that estimated relative consumption of resources and the costs of different diagnostic and therapeutic services in five pilot hospitals.

Table 2
Descriptive statistics of the measures of inputs and outputs

		Mean	Standard Deviation
INPUTS	DOCTORS	214,85	214,73
	REST OF STAFF	1.091,64	1.136,08
	BEDS	400,93	375,36
OUTPUTS	MEDICAL INPATIENT-DAYS	46.485,92	46.753,00
	SURGICAL INPATIENT-DAYS	49.856,00	52.845,00
	INTENSIVE CARE INPATIENT-DAYS	3.165,75	4.409,50
	OBSTETRIC INPATIENT-DAYS	7.795,43	8.604,90
	NEWBORN INPATIENT-DAYS	2.233,00	3.568,90
	PEDIATRIC INPATIENT-DAYS	5.722,50	7.189,80
	AMBULATORY SURGICAL PROCEDURES	1.989,02	2.265,00
	OPERATIONS WITH HOSPITALIZATION	3.688,78	3.781,55
	UPAMIX	55.395,15	58.761,74
	ADMISSIONS	12.432,95	11.196,23
	TECNO	783,30	1.763,28

The unconstrained analysis indicates that the average level of relative global efficiency is high: 0.92 in 1991, 0.94 in 1992 and 0.95 in 1993, with less variance in 1993 than in 1991. The number of hospitals with maximum relative efficiency rises; from 22 in 1991 to 28 in 1992 and 37 in 1993. This implies that the proportion of inefficient hospitals has fallen from 70% in 1991 to 50% in 1993. Twenty-one hospitals moved from a situation of relative inefficiency in 1991 to one of relative efficiency in 1993, while the converse applied to only six hospitals.

Detailed results of the unconstrained analysis (P1) are given in Tables 3 and 4. Table 3 contains, for each group and year, the mean and minimal scores of global, pure and scale efficiency. Both the global and the pure efficiency are in average higher in 1993 than in 1991. This result occurs also for each group. In Table 4 one can see that the median score of efficiency has increased from 1991 to 1993, and the standard deviation has decreased.

Table 3
Unconstrained Data Envelopment Analysis

	GLOBAL EFFICIENCY						PURE EFFICIENCY						SCALE EFFICIENCY					
	EFI91		EFI92		EFI93		EFIP91		EFIP92		EFIP93		EFIESC91		EFIESC92		EFIESC93	
	MEAN	MIN	MEAN	MIN	MEAN	MIN	MEAN	MIN	MEAN	MIN	MEAN	MIN	MEAN	MIN	MEAN	MIN	MEAN	MIN
G1	0.88	0.68	0.92	0.73	0.93	0.71	0.95	0.79	0.96	0.82	0.98	0.85	0.93	0.70	0.96	0.83	0.95	0.72
G2	0.93	0.68	0.93	0.66	0.95	0.74	0.94	0.68	0.94	0.66	0.95	0.74	1.00	0.97	0.99	0.91	1.00	0.96
G3	0.98	0.93	0.99	0.91	0.99	0.95	0.97	0.80	0.99	0.91	1.00	0.97	1.00	0.98	1.00	0.98	0.99	0.96
G4	0.92	0.85	0.95	0.84	0.98	0.85	0.98	0.91	1.00	0.99	1.00	0.94	0.95	0.85	0.95	0.84	0.98	0.89
TOT	0.92	0.68	0.94	0.66	0.95	0.71	0.95	0.68	0.96	0.66	0.97	0.74	0.96	0.70	0.97	0.83	0.98	0.72

Table 4
Unconstrained Data Envelopment Analysis

	GLOBAL EFFICIENCY			PURE EFFICIENCY			SCALE EFFICIENCY		
	EFI91	EFI92	EFI93	EFIP91	EFIP92	EFIP93	EFIESC91	EFIESC92	EFIESC93
Standard deviation	0.086	0.075	0.070	0.065	0.060	0.048	0.062	0.044	0.047
Rank	0.316	0.338	0.285	0.313	0.336	0.252	0.296	0.161	0.273
Median	0.937	0.968	1.000	0.976	1.000	1.000	0.995	0.997	1.000

Table 5 shows that the number of hospitals in each group with inefficiency of scale in 1993. Hospitals with increasing returns to scale in 1993 were small district ones (13 in group 1 and 9 in group 2), while those with decreasing returns to scale were mainly large hospitals (group 4).

Table 5
Efficiency of scale for 1993

	OPTIMAL SIZE	DECREASING RETURNS TO SCALE	INCREASING RETURNS TO SCALE
GROUP 1	11	0	13
GROUP 2	16	1	9
GROUP 3	8	2	0
GROUP 4	7	8	0
TOTAL	42	11	22

A majority of hospitals in group 3 (those with more than 1.700 medical staff, 6 to 8 complex specialities, and which, if they do transplants at all, transplant one organ only) belong to the frontier of efficiency in 1993, in terms of global, pure, and also scale efficiency.

To find out whether there were significant differences between levels of global efficiency in the three years under consideration, we applied a Friedman non-parametric test for *k* related samples, in which the null hypothesis was:

H_0= the probability distribution of global efficiency is the same for the years 1991, 1992, and 1993.

The resulting Chi-square of 17.26 led us to reject the null hypothesis of equal distributions and conclude that the levels of relative global efficiency were significantly higher in 1993 than in previous years.

We performed this same test on the measures of hospital efficiency obtained from the restricted analysis P2 (interperiod comparisons). The null hypothesis was rejected, and an increase of productivity was anoted for the period 1991-1993. As expected since the sample were more homogeneous in P2 and P3, the scores for efficiency increased. The resulting relationship is $z_i^{(1)} \leq z_i^{(3)} \leq z_i^{(2)}$, that is, when each hospital is compared only with those of its group during the sample period P2, the relative indices of efficiency are higher than when all the hospitals are compared for a single year (P3), and in both cases higher than the unconstrained model (P1).

Table 6
Average level of relative global efficiency

	Unconstrained model (P1)	Constrained model (P2)	Constrained model (P3)
1991	0.92	0.93	0.94
1992	0.94	0.96	0.96
1993	0.95	0.97	0.98

2.4. *Determinants of efficiency*

Are there features of hospitals that explain differences in technical efficiency, or groups of hospitals that stand out from the others? The following Tobit model of censored regression attempts to answer this question.

The sample includes the same 75 general hospitals of *Insalud Gestión Directa* submitted to DEA analysis for the years 1991-1993. The dependent variable, with values not greater than 1, is the DEA score for pure technical efficiency resulting from P1. We did not work with the global technical efficiency to avoid comparisons between hospitals with different returns to scale. The independent variables include the dummy ones of the group, except that of group 1 which serves as the reference group. Another dummy variable is 1 for the hospitals of Madrid. Furthermore there are three continuous independent variables. MIRP is the percentage MIR doctors on the staff of doctors. We included this variable because the DEA analysis did not consider teaching an output of a hospital, which meant that if the MIRP coefficient was negative and significant it might indicate an error in the specification of the outputs of the DEA analysis. CONCP is the amount of expenditure in subcontracts (consortia) with other institutions over the total current expenditures of the hospital. TFREQ is the rate of hospital use (hospital admissions per thousand inhabitants of the population of its designated hinterland). Finally, we have added the variable YEAR (=1,2,3), to estimate the mean annual variation in technical efficiency. If its coefficient is significantly positive, the model finds a systematic improvement in technical efficiency after the introduction of program-contracts. This turned out to be the case. The results in Table 7 of the estimation are conclusive. Other things being equal, there was an annual mean increase in efficiency of 3.5%. In regard to the other independent variables, the model indicates that the (larger hospitals) in groups 3 and 4 were significantly more efficient than the (smaller) ones in group 1, and that those of Madrid were more efficient than those elsewhere. The hospitals that subcontract out more medical services also, according to the model, do better. Hospitals with higher levels of use, on the other hand, are associated with lower technical efficiency. The proportion of MIR on the staff does not significantly affect the efficiency of hospitals either way.

While it is common in the specialized literature to explain DEA measures of efficiency by censored Tobit models (Vitaliano y Toren, 1994; Chilingerian, 1995), strictly speaking the DEA score does not fit the theory of sample censoring that gives rise to these models: the accumulation of sample observations at the highest level of efficiency is not the result of imperfect observation, as Tobit originally supposed, but rather is intrinsic to the model (Chilingerian, 1995). The hypothesis that the latent efficiency is normally distributed is implausible; in this case the Tobit maximum likelihood estimators based on a normal distribution are not consistent (Amemiya, 1984). This hypothesis of a normal distribution has been tested and rejected for our case using the Pagan and Vella (1989) test for normality. There are thee possible ways to estimate that overcome this problem: 1) use a robust method; 2) remove the censoring problem directly, dropping the ith hospital from its own reference set, i.e. from the constraint set (Burgess and Wilson, 1996); and 3) estimate the model assuming different distributions of probability of disturbance, and check whether the differences in estimates are substantial. We adopted the latter strategy and recalculated the estimates assuming alternative distributions of the error term: normal, Weibull, logistic and exponential. The coefficients we estimated have the same sign and are quite similar except in the exponential case.

Table 7
Tobit model of censored regression with normal disturbance
Panel of *Insalud Gestión Directa* hospitals, 1991-93
Dependent Variable: DEA pure technical efficiency

Variable	Coefficient	Standard error
Constant	0.989**	0.032
GROUP2	- 0.026	0.019
GROUP3	0.083**	0.035
GROUP4	0.096**	0.042
MADRID	0.067**	0.030
MIRP	- 0.093	0.079
CONCP	0.389**	0.133
YEAR	0.035**	0.009
TFREQX10^3	- 0.075**	0.031
$\sigma^{(1)}$	0.090**	0.007

** Significant at 5%
(1) Estimated standard deviation of the disturbace

2.5. *Tecnical efficiency and unit costs*

In this section we estimate a model of random effects with a period effect for the panel of *Insalud Gestión Directa* hospitals. This model explains the current expenditures per UPA. These expenditures are measured as deviations from the mean expenditure per UPA for the set of hospitals. The explanatory variable the model focuses on is the DEA measure of global technical efficiency. The more efficient hospitals use less input to produce a given amount of output, so they should show up with a significant negative coefficient. We have chosen the measure of global technical efficiency because the structural problems that derive from a hospital not working at the optimal scale of operation could foreseeable affect costs. In addition to efficiency, we have included in the list of explanatory variables, to control for the issues discussed in Section 2.4, the dummy variables hospital GROUP, MADRID, MIRP, and CONCP, defined above. The model is:

$$Y_{it} = \alpha_1 + \sum_{J=2}^{4} \alpha_j \, GROUPj_{it} + \alpha_5 \, MADRID_{it} + \alpha_6 \, MIRP_{it} + \alpha_7 \, CONCP_{it} + \alpha_8 \, EFI_{it} + e_i + u_t + w_{it}, \qquad [2]$$

with $e_i \sim iid\ (0, \sigma_e^2)$, $u_t \sim iid\ (0, \sigma_u^2)$, $w_{it} \sim iid\ (0, \sigma_w^2)$, *and* $E\,(e_i.u_t) = E\,(e_i.w_{it}) = E\,(u_t.w_{it}) = 0.$

It also assumes that the three components of the random noise are independent of the explanatory variables. This hypothesis has been tested previously. The model, estimated by generalized least squares, generated as expected a highly significant coefficient of efficiency (Table 8). Other results worth noting include the lower cost of producing a UPA in the middle-sized hospitals (groups 2 an 3) as opposed to the smallest ones (group 1), and the high unit costs of the hospitals in Madrid. Neither the presence of MIR doctors on the staff nor subcontracting affected significantly the cost per UPA.

Table 8

Random effect model
Panel of Insalud Gestión Directa hospitals, 1991-1993
Dependent variable: Expenditure per UPA (as a deviation from the annual mean)

Variable	Coefficient	Standard error
Costant	24.219**	3.188
GROUP2	-6.061**	1.275
GROUP3	-3.397*	1.881
GROUP4	-2.201	2.189
MADRID	3.895**	1.608
MIRP	-5.150	3.631
CONCP	-2.024	4.673
EFI	-22.455**	3.324
σ_e	6.532	
σ_u	17.295	
σ_w	1.785	

* Significant at 10%
** Significant at 5%

3. A cost frontiers analysis of the evolution of hospital efficiency in Spain

3.1. Introduction: Methodology of stochastic parametric frontier models

Stochastic parametric frontier models overcome the limitations of data envelopment models at the price of greater rigidity in functional form and the requirement of additional hypothesis about the probability distributions of the stochastic elements of the model.

All stochastic parametric frontier models have a «compound error» with two components. One of them, v_i, is part of the frontier itself, which hence is stochastic, representing errors of measurement, the omission of variables, and the presence of unpredictable and uncontrollable exogenous events that affect production. The second component of error, u_i, is asymmetric ($u_i \geq 0$ in the model of cost frontiers, and $u_i \leq 0$ in the model of production frontiers), measures inefficiency, and is equivalent to the u_i in deterministic frontier models. Aigner, Lovell and Schmidt (1977) proposed methods for estimating the frontier by maximum likelihood and by corrected least squares, assuming the following alternative distributions of the asymmetric term for technical inefficiency u_i: semi-normal, exponential, and truncated normal. Nevertheless their method did not estimate individual technical efficiency consistently, although it did measure the sample average of the u_i's Jondrow, Lovell and Schmidt (1982) discovered a method for unbiased although inconsistent estimation of u_i conditional on the total error $v_i + u_i$. Schmidt and Lovell (1979) for the first time estimated separately technical and allocative efficiency for a Cobb-Douglas production function with a single output.

The main limitation of these cross-sectional models is that their results depend on the probability distribution assumed for u_i. This limitation has been overcome by Kalirajan and Hand (1994). For a single output and a Cobb-Douglas function, they estimated a frontier assuming a general distribution of u_i. Technical efficiency varies between hospitals but allocative efficiency is assumed to be «persistent», that is, to vary among inputs but not among hospitals.

Stochastic frontier models for panels of data are generally more flexible in the specification of u_i. Some of them allow dynamic effects. Schmidt and Sickles (1984) formulated a production model that does not require the assumption of a given probability distribution for the errors u_i and v_i and that allows the measurement of individual efficiency (u_i). The only (quite restrictive) constraining hypothesis is that the level of inefficiency of each hospital remains constant over the period of observation.

For the model of fixed effects, u_i is assumed to be fixed for each hospital so that these n inefficiencies form part, together with the α, of the set of parameters to be estimated. But this model presents difficulties for estimation when some of its explanatory variables are invariant over time. In our case, there are various structural aspects of hospitals that do not vary over our three-year period, and these prevent us from using this model.

In the random effects model it is assumed that the inefficiency u_i of a hospital is independent of the inputs and other explanatory variables, and that furthermore it is independent of the transient component of error v_{it}. The coefficients of this model can be estimated by generalized least squares or by maximum likelihood if some particular probability distribution of the two components of error is specified.

Lichtenberg (1988), with a Cobb-Douglas production function for a single output, suggested a method for estimation of technical efficiency, that allows for temporal variation in efficiency, but only for homogeneous groups of hospitals.

The novel idea of Cornwell, Schmidt and Sickles (1990) to permit technical inefficiency to vary over time and space is quite suggestive. They generalize the model of Schmidt and Sickles (1984) parametrizing the inefficiency u_{it} as a free distribution de probability function of time. Their model is a Cobb-Douglas production function with one output for a panel with heterogeneity in the intercepts and the other coefficients. In our application we also introduced explicitly time into the model of cost frontier.

3.2. Empirical application: Results

We estimated a cost frontier model for the years 1991-1993 for a panel of 73 *Insalud Gestión Directa* general hospitals, after excluding the atypical hospitals of group 5, and two other hospitals for lack of information. The functional form is quite flexible, a non-homothetic translog. The logarithm of the total costs of hospital i in year t is the dependent variable, defined excluding investment costs. For independent variables we introduced in the equation linear and nonlinear functions of the logarithms of outputs in physical units, the prices of the inputs, and other structural features that could affect the cost frontier of a hospital. We considered three outputs: hospitalization, ambulatory service, and high-technology service. We measured the amount of hospitalization by patient-days (ESTANC). Ambulatory service includes first and successive outpatient visits, emergency ward treatment not leading to hospitalization, and ambulatory surgical procedures. The measure for high technology (TECNO) was the same we used in the DEA. These three outputs weigh the activities with the weights defined in the UPA, already mentioned. We considered two inputs, labour and material, which includes maintenance of equipment. The input prices are, for labour the average salaries in the hospital (p_1) and for materials the current expenditure in goods and services per bed (p_2). The other independent variables are average patient stay, teaching status (a dummy variable equal to 1 for hospitals with MIR), the logarithm of the number of beds, and a time trend *(t)*. This last variable is to estimate the systematic variations of the cost frontier over the three years. If its coefficient is negative and significant, the model indicates that average (technical plus allocative) efficiency improved after the introduction of the program-contract.

With the outputs as $Y_r (r=1, 2, 3)$, the prices of the inputs as p_j $(j=1, 2)$, and the other structural independent variables as Z_h $(h=1, 2, 3)$, the model for hospital i in year t is as follows:

$$LnCT_{it} = \alpha_0 + \sum_{r=1}^{3} \alpha_r \ln Y_{r,it} + \sum_{j=1}^{2} \beta_j \ln p_{j,it}$$
$$+ \frac{1}{2}\sum_{r=1}^{3}\sum_{j=1}^{3} \gamma_{rj} \ln Y_{r,it} \ln Y_{j,it} + \frac{1}{2}\sum_{j=1}^{2}\sum_{k=1}^{2} \delta_{jk} \ln p_{j,it} \ln p_{k,it} \quad [3]$$
$$+\sum_{r=1}^{3}\sum_{j=1}^{2} \eta_{rj} \ln Y_{r,it} \ln p_{j,it} + \theta_0 t + \theta_1 Z_{1,it} + \ldots\ldots + \theta_h Z_{h,it} + v_{it} + u_{it}$$

with $v_{it} \sim N(0, \sigma_v^2)$, $u_{it} \geq 0$, and $|u_{it}| \sim N(0, \sigma_u^2)$. Since it is a frontier model, its disturbance is composed of v_{it}, which is random noise due to errors in measurement and factors of good or bad luck not included explicitly, and u_{it} a random asymmetrically positive term representing inefficiency. A semi-normal distribution is assumed for u_{it}, and independence between u_{it} and v_{it} and between them and each of the explanatory variables. The term u_{it} registers the aggregate of technical and allocative inefficiency. Hence the inefficiencies estimated by our model are not comparable with the measures from the previous DEA, which only registered technical inefficiency.

As for the implicit hypotheses, we wish to emphasize that the fact of specifying a model of random effects implies an assumption of independence between the individual inefficiency of each hospital and the independent variables.

We impose the following constraints of symmetry:

$$Y_{rj} = Y_{jr} \ (r, j = 1, 2, 3); \ \delta_{jk} = \delta_{kj} \ (j, k = 1, 2) \qquad [4]$$

Furthermore, one must make some additional constraints to ensure that the cost function is linearly homogeneous in the prices of inputs. In our case, these additional constraints are:

$$\beta_1 + \beta_2 = 1$$

$$\delta_{11} + \delta_{12} = \delta_{21} + \delta_{22} = 0$$

$$\eta_{11} + \eta_{12} = \eta_{21} + \eta_{22} = \eta_{31} + \eta_{32} = 0 \qquad [5]$$

By substituting [5] into [3] we can obtain an equation in which the dependent variable is the logarithm of total cost divided by the price of input 1 (p_1). Hence this constrained model has 15 independent variables, in addition to the Z variables. The model is estimated by maximum likelihood.

Table 9 contains the results of three alternative specifications. The first model (M1) is the more general and includes the logarithm of the number of beds as the influential structural element, apart from the activity performed, in the cost frontier. The second model (M2) excludes beds but is otherwise identical. For illustrative purposes, a third model (M3) is based on the Cobb-Douglas specification.

A battery of Wald sequential tests provides the following results: One cannot reject the hypothesis of a homothetic function ($\eta_{12} + \eta_{22} + \eta_{32} = 0$) for any level of significance smaller than 0.14 ($x^2_1 = 0.297$). Nor can one reject the hypothesis of returns to scale being constant in the model M1[5] ($x^2_5 = 5.09$). Nevertheless, the Cobb-Douglas equation, which constrains the translog model by imposing null values on all the coefficients of interaction between outputs (γ) and prices (η) is clearly rejected at any level of significance ($x^2_9 = 21.197$)[5].

Although there is a high degree of multicollinearity, as in any direct estimate of a translog cost function, it is unlikely that this problem will affect individual predictions or the residuals of the regression. Indeed, the correlation between the residuals in the different models is very high[6].

[5] This hypothesis can be tested imposing a level 1 homogeneity in the M1 or M2 cost functions, which implies the following set of homogeneous linear restrictions: $\alpha_1 + \alpha_2 + \alpha_3 = 1$; $\gamma_{11} + \gamma_{12} + \gamma_{13} = \gamma_{21} + \gamma_{22} + \gamma_{23} = \gamma_{31} + \gamma_{32} + \gamma_{33} = 0$; $\eta_{12} + \eta_{22} + \eta_{32} = 0$.

[6] We have also estimated, although we do not report it here, a version of the model with two outputs and various other specifications that include the rate of hospital use (admissions for 1000 inhabitants in the reference population), indices of bed occupation, and percentage of MIR in place of the teaching dummy. The residuals of the alternative translog models have correlations always higher than 0.7.

The results are also robust in relation to alternative hypotheses about the probabilistic distribution of u, whether truncated or exponential.

The u_{it} residuals have the correct kind of asymmetry. Their variance in all of the models that we estimated was significantly higher than the variance of the noise v_{it}, indicating that the variation between hospitals in levels of efficiency is 5 to 7 times greater, depending on the model, than the random variations of cost frontier (v_{it}).

Table 9

Estimation of Cost Frontier Functions
Panel of *Insalud Gestión Directa* hospitals, 1991-93
Dependent Variable: LN(TOTAL COST(CT)/p_1)

VARIABLE	MODEL 1 (Translog)		MODEL 2 (Translog)		MODEL 3 (Cobb-Douglas)	
CONSTANT	7.137	(4.765)	10.984	(5.523)	-5.146**	(0.452)
LN(ESTANC)	-1.999*	(1.025)	-3.180*	(1.576)	0.508**	(0.047)
LN(AMBU)	0.842	(1.116)	1.534	(1.515)	0.359**	(0.050)
LN(TECNO)	-0.090	(0.208)	-0.072	(0.216)	0.020**	(0.007)
$(LN(ESTANC))^2$	0.143	(0.096)	0.177	(0.150)		
LN(ESTANC).LN(AMBU)	-0.197	(0.171)	-0.210	(0.240)		
LN(ESTANC).LN(TECNO)	0.024	(0.023)	0.030	(0.024)		
$(LN(AMBU))^2$	0.126	(0.091)	0.145	(0.107)		
LN(AMBU).LN(TECNO)	-0.029	(0.021)	-0.037	(0.024)		
$(LN(TECNO))^2$	0.001	(0.003)	0.000	(0.005)		
$LN(p_2/p_1)$					0.297**	(0.050)
$(LN(p_2/p_1))^2$	0.033	(0.066)	0.005	(0.069)		
$LN(ESTANC).LN(p_2/p_1)$	0.075	(0.087)	0.201	(0.214)		
$LN(AMBU).LN(p_2/p_1)$	-0.099	(0.158)	-0.199	(0.254)		
$LN(TECNO).LN(p_2/p_1)$	1.160	(1.615)	1.188	(1.966)		
YEAR	-0.025**	(0.007)	-0.036**	(0.008)	-0.038**	(0.006)
ESTANC AVERAGE (EA)	0.007	(0.012)	0.006	(0.015)	-0.006	(0.106)
TEACHING STATUS	0.026	(0.020)	0.045	(0.022)	0.031*	(0.018)
LN(BEDS)	0.724**	(0.105)				
σ^2_u/σ^2_v	8.724**	(3.286)	12.748**	(4.490)	12.733**	(4.018)
$Var(u)/\sigma^2_v$	4.979					
σ^2_v	0.003**	(0.000)	0.003**	(0.000)	0.004**	(0.000)
σ^2_u	0.023		0.040		0.050	
Var(u)[1]	0.013		0.023		0.028	
Ln(L)	235.876		202.201		178.883	

The standard errors are in parentheses.
* Indicates a coefficient significant at 10% and ** at 5%.
(1) If u is semi-normal: $|u| \sim N(0, \sigma^2_u)$, the variance of u is $Var(u) = (\pi/2 - 1)\,\sigma^2_u$.

The coefficient for YEAR is negative in all the models tested, indicating a systematic decrease in the cost frontier between 1991 and 1993. Like the data envelopment analysis, these stochastic frontier models find a general improvement of efficiency after the implementation of program-contracts.

Assuming a semi-normal distribution of u_{it} an unbiased estimates of the inefficiency for the year t of hospital i, given the compound error of the hospital, has been calculated in the following way (Jondrow *et al.*, 1982):

$$E(u_{it}/v_{it} + u_{it}) = \frac{\sqrt{\sigma^2_v + \sigma^2_u}\,\lambda}{1+\lambda^2}\left[\frac{\phi\,((u+v)\,\lambda/\sigma)}{\Phi\,((u+v)\,\lambda/\sigma)} - \frac{(u+v)\,\lambda}{\sigma}\right], \quad [6]$$

$$\sigma = \sqrt{\sigma^2_v + \sigma^2_u} \text{ and}$$

where $\lambda = \sigma_u/\sigma_v$, ϕ (.) and Φ (.) are the density and distribution functions of a standard normal distribution.

Using [6], we could calculate the excess of expected costs, as a percentage, over the frontier for each hospital and year, since u_{it} is the difference between the logarithm of the real cost and the minimal estimated cost that corresponds to the frontier, and hence:

$$e^{u_{it}} - 1 = \frac{CT_{it} - CT^*_{it}}{CT^*_{it}} \quad [7]$$

The results on the percentage of the total cost that could be saved if all sources of inefficiency were eliminated, indicate an average potential savings of around 13.5% of expenditures over a frontier that has been reducing costs for three years. We emphasize that the heterogeneity of the inefficient behaviour has diminished. The standard deviation of estimated savings for the 73 hospitals goes down each year (0.14, 0.11, and 0.10 respectively), and the most inefficient hospitals have improved substantially. For example, the most inefficient spent 62% above the cost frontier in 1991, but only 46% in 1993.

The program-contracts appear to have imposed, according to these results, a certain discipline, especially on the hospitals farthest away from the cost frontier, which has reduced the disparities among hospitals.

Although as indicated above the measures of inefficiency of the translog models of cost frontiers are not directly comparable with those of our earlier data envelopment analysis, we made a Wilcox test of rank equality among the inefficiencies measured by both models, for each year. For no year can one reject the equality of rankings obtained at standard levels of significance.

4. Summary and conclusions

The main advantages of DEA for the determination of technical efficiency in productive units are that it is non-parametric and can analyze many outputs. But there are limitations both in general and in the application to hospitals. That is why, in addition to determining and analyzing DEA measures of technical efficiency for hospitals, we applied this analysis with additional constraints, and estimated some econometric models of cost frontiers.

We found a significant increase in the levels of relative efficiency in the general hospitals of *Insalud Gestión Directa* between 1991 and 1993, following a change in the model of financing reflected in the program-contracts. The percentage of inefficient hospitals declined from 70% in 1991 to 50% in 1993. Twenty-one of the 75 hospitals we analyzed had reached the efficiency frontier in 1993 from a position of inefficiency in 1991, while over the same period only six hospital lapsed into inefficiency from the frontier. Hospitals with increasing returns to scale in 1993 were small district ones (13 in group 1 and 9 in group 2), while those with decreasing returns to scale in the same year were mainly large hospitals (group 4). In 1993 the majority of hospitals in group 3 were at the frontier of efficiency.

Since what one determines from a DEA model is a frontier of efficiency that is *relative*, the results may be very sensitive to heterogeneity in the sample. To reduce this sensitivity we added new restrictions to the initial model that limited the comparison of each hospital unit to others that were similar. As expected, the scores for technical efficiency came out higher than those obtained from the unconstrained analysis, and the levels of efficiency in 1993 are significantly higher than those of the previous years.

Subsequently we looked for features that could explain the differences in technical efficiency using a Tobit model of censored regression. In this way we identified as significant the variables size of hospital, subcontracting, and rate of hospital use. This model confirmed the systematic improvement of pure technical efficiency in the period 1991-1993. A random-effects model for the panel of hospitals shows that the DEA scores for technical global efficiency contribute significantly to explain the cost of producing a UPA, which supports the results of the data envelopment analysis.

Finally, for the same panel of hospitals we evaluated technical and allocative efficiency together with a translog stochastic costs frontier model of costs. The results also showed some improvements of efficiency after the introduction of the program-contracts.

Referencias

Afriat, S. N. (1972): «Efficiency estimation of a production function», *International Economic Review* 13, pp. 568-598.

Aigner, D. J. and Chu, S. F. (1968): «On Estimating the industry production function», *American Economic Review* 58, pp. 826-838.

Aigner, D. J.; Lovell, C. A. and Schmidt P. J. (1977): «Formulation and estimation of stochastic frontier production function models», *Journal of Econometrics* 6, pp. 21-37.

Ali, A. I., Lerme, C. S. and Seiford, L. M. (1995): «Components of efficiency evaluation in data envelopment analysis», *European Journal Operational Research* 80, pp. 462-473.

Amemiya, T. (1988): «Una revisión de los modelos Tobit». *Cuadernos Económicos de ICE* 39, pp. 73-128.

Arellano, M. and Bover, O. (1990): «La econometría de datos de panel», *Investigaciones Económicas* 14, pp. 3-45.

Banker, R. D. (1984): «Estimating most productive scale size using data envelopment analysis», *European Journal of Operational Research* 17, pp. 35-44.

Banker, R. D.; Charnes, A. and Cooper, W. W. (1984): «Some models for estimating of technical and scale inefficiencies in data envelopment analysis», *Management Science* 30, pp. 1078-1092.

Banker, R. D.; Conrad, R. and Strauss, R. (1986): «An application of DEA to the empirical investigation of a hospital production function». *Management Science.*

Banker, R. D. and Morey, R. (1986): «Efficiency analisys for exogenously fixed inputs and outputs», *Operations Research* 34, pp. 513-521.

Banker, R. D. and Thrall, R. M. (1992): «Estimation of returns to scale using data envelopment analysis», *European Journal of Operational Research* 62, pp. 74-84.

Bauer, P. W. (1990): «Recent developments in the econometric estimation of frontier», *Journal of Econometrics* 46, pp. 39-56.

Bestard, J.; Sevilla, F.; Correlia, M. and Elola, J. (1993): «La unidad ponderada asistencial (UPA): nueva herramienta para la presupuestación hospitalaria», *Gaceta Sanitaria* 7, pp. 263-273.

Boussofiane, A.; Dyson, R. G. and Thanassoulis, E. (1991): «Applied data envelopment analysis», *European Journal of Operational Research* 52, pp. 1-15.

Burguess, J. F. and Wilson, P. W. (1993): «Technical efficiency in veterans administration hospitals», en Fried, Lovell y Schmidt (eds.): *The Measurement of Productive Efficiency*, Oxford University Press, Oxford

Burguess, J. F. and Wilson, P. W. (1996): «Variation in inefficiency among US hospitals» (mimeo).

Chang, K. P. and Yeah-Yuh, G. (1991): «Linear production functions and the data envelopment analysis», *European Journal of Operational Research* 52, pp. 215-223.

Charnes, A.; Cooper, W. W. and Rhodes, E. (1978): «Measuring efficiency of decision making units», *European Journal of Operational Research* 2, pp. 429-444.

Charnes, A.; Cooper, W. W.; Seiford, L. and Stutz, J. (1983): «Invariant multiplicative efficiency and piecewise Cobb-Douglas envelopments», *Operational Research Letters* 2, pp. 101-103.

Christensen, L. P.; Jorgenson, D. W. and Lau, L. J. (1973): «Trascendental logarithmic production function frontiers», *Review of Economics and Statistics* 55, pp. 29-45.

Cornwell, C.; Schmith, P. and Sickles, R. (1990): «Production frontiers with cross-sectional and time-series variation in efficiency levels», *Journal of Econometrics* 46, pp. 185-200.

Dusansky, R. and Wilson, P. W. (1994): «Thechnical efficiency in the descentralized care of the developmentally disabled», *Review of Economics and Statistics*, pp. 340-345.

Ehreth, J. L. (1994): «The development and evaluation of hospital performance measures for policy analysis», *Medical Care* 32, pp. 568-587.

Färe, R. and Grosskopf, Sh. (1994): «Estimation of returns to scale using data envelopment analysis: a comment», *European Journal of Operational Research* 79, pp. 379-382.

Farrell, M. J. (1957): «The measurement of productive efficiency», *Journal of the Royal Statistical Society A* 120, pp. 253-281.

Ferrier, G. and Lovell, K. (1990): «Measuring cost efficiency in banking: econometric and linear programming evidence», *Journal of Econometrics* 46, pp. 229-245.

Greene, W. H. (1980): «On the estimation of a flexible frontier production model», *Journal of Econometrics* 13, pp. 101-117.

Grosskopf, S.; Margaritis, D. and Valdmanis, V. (1995): «Estimating output substitutability of hospital services: a distance function approach», *European Journal of Operational Research* 80, pp. 575-587.

Hadley, J. and Zuckerman, S. (1994): «The role of efficiency measurement in hospital rate setting», *Journal of Health Economics*, pp. 335-340.

Huang, Y. L. (1990): «An application of data envelopment analysis: measuring the relative performance of Florida general hospitals», *Journal of Medical Systems* 14, pp. 191-196.

Jondrow, J.; Lovell, C. A.; Materov, I. S. and Schmidt, P. (1982): «On the estimation of technical inefficiency in the stochastic frontier production function model», *Journal of Econometrics* 19, pp. 233-238.

Kalirajan, K. P. and Shand, R. T. (1989): «A generalized measure of technical efficiency», *Applied Economics* 21, pp. 25-34.

Kao, C. (1994): «Efficiency improvement in data envelopment analysis», *European Journal of Operational Research* 73, pp. 487-494.

Kooreman, P. (1994): «Data envelopment analysis and parametric frontier estimation: complementary tools», *Journal of Health Economiscs*, pp. 345-346.

Kumbhakar, S. (1990): «Production frontiers, panel data, and time-varying technical inefficiency», *Journal of Econometrics* 46, pp. 201-211.

Ley, E. (1991): «Eficiencia productiva: un estudio aplicado al sector hospitalario», *Investigaciones Económicas* 15, pp. 71-88.

Ley, E. (1996): «Data screening and data envelopment analysis», in *Computational Economics and Finance: Modeling and Analysis with Mathematica*, Hal Varian (ed.), Springer-Verlag, Santa Clara, CA.

López Casasnovas, G. and Wagstaff, A. R. (1992): «Indicadores de eficiencia para la gestión pública: una revisión de método», *Cuadernos de Economía* 19, pp. 55-83.

López Casasnovas, G. and Wagstaff, A. R. (1992): «Eficiencia y competitividad en los servicios públicos: algunas consideraciones relativas a la asistencia sanitaria», in *Eficiencia y Competitividad: Políticas de oferta en el proceso de creación del Mercado Único Europeo*, V Simposio de Moneda y Crédito, Madrid.

Lovell, C. A. K. (1994): «Linear programming approaches to the measurement and analisys of productive efficiency», *TOP* 2, pp. 175-223.

Morey, R. C.; Fine, D. J. and Loree, S. W. (1990): «Comparing the allocative efficiencies of hospitals», *Omega* 18, pp. 71-83.

Newhouse, J. (1994): «Frontier estimation: how useful a tool for health economics?», *Journal of Health Economics*, pp. 317-322.

Nunamaker, T. (1983): «Measuring routine nursing service efficiency: a comparison of cost per patient day and data envelopment analysis models», *Health Services Research* 18, p. 183.

Olson, J. A.; Schmidt, P. and Waldman, D. M. (1980): «A Monte Carlo study of estimators of stochastic frontier production functions», *Journal of Econometrics* 13, pp. 67-82.

Ozcan, Y. A.; Luke, R. D. and Haksever, C. (1992): «Ownership and organizational performance», *Medical Care*, pp. 781-794.

Ozcan, Y. A. and Luke, R. D. (1993): «A national study of the efficiency of hospitals in urban markets», *Health Services Research*, pp. 719-739.

Prior Jiménez, D. and Solà i Tey, M. (1993): «L'eficiència dels hospitals de Catalunya. Comparació entre els hospitals públics i els privats», *Generalitat de Catalunya, Barcelona.*

Richmond, J. (1974): «Estimating the efficiency of production», *International Economic Review* 15, pp. 415-421.

Rosko, M. D. (1990). «Measuring techical efficiency in health care organizations», *Journal of Medical Systems* 14, p. 307.

Schmidt, P. (1976): «On the statistical estimation of parametric frontier production functions», *Review of Economics and Statistics* 58, pp. 238-239.

Schmidt, P. (1988): «Estimation of a fixed-effect Cobb-Douglas system using panel data», *Journal of Econometrics* 37, pp. 361-380.

Schmidt, P. and Lovell, C. (1979): «Estimating technical and allocative inefficiency relative to stocahastic production and cost frontiers», *Journal of Econometrics* 9, pp. 343-366.

Schmidt, P. and Sickles, R. (1984): «Production frontiers and panel data», *Journal of Business and Economic Statistics* 2, pp. 367-374.

Schmidt, P. and Lin, T. (1984): «Simple tests of alternative specifications in stochastic frontier models», *Journal of Econometrics* 24, pp. 349-361.

Sevilla, F.; Corella, I.; Bestard, J. J. and Elola, F. J. (1993): «Presupuestación prospectiva en los hospitales del Insalud», *Presupuesto y Gasto Público* 10, pp. 143-148.

Sexton, T. R.; Leiken, S. *et al.* (1989): «The impact of prospective reimbursement on nursing home care», *Medical Care* 27, p. 154.

Skinner, J. (1994): «What do stochastic frontier cost functions tell us about inefficiency?», *Journal of Health Economics*, pp. 323-328.

Sueyoshi, T. (1994): «Stochastic frontier production analysis: measuring performance of public telecommunications in 24 OECD countries», *European Journal of Operational Research* 74, pp. 466-478.

Valdamis, V. G. (1990): «Ownership and thecnical efficiency of hospitals», *Medical Care*, pp. 552-561.

Vitaliano, D. F. and Toren, M. (1994): «Frontier analysis: a reply to skinner, dor and newhouse», *Journal of Health Economics*, pp. 341-343.

Vitaliano, D. F. and Toren, M. (1994): «Cost and efficiency in nursing homes: a sthochastic frontier approach», *Journal of Health Economics*, pp. 281-300.

Wei, Q. L; Sun, B. and Xiao, Z. J. (1995): «Measuring technical progress with data envelopment analysis», *European Journal of Operational Research* 80, pp. 691-702.

Wagstaff, A. (1989): «Estimating efficiency in the hospital sector: A comparison of three statistical cost frontier models», *Applied Economics*, pp. 659-672.

Wagstaff, A. (1992): «Quantitative thecniques for investigating health service efficiency», in Seminario de Economia y Salud del Sistema Sanitario Español,Universidad Internacional Menéndez Pelayo, September 1992,

Wilson, P. W. (1993): «Detecting outliers in deterministic nonparametric frontier models with multiple outputs», *Journal of Business and Economic Statistics* 11, 3, pp. 319-323

Wilson, P. W. (1995): «Detecting influential observations in data envelopment analysis», *Journal of Productivity Analysis* 6, pp. 27-45

Resumen

En este trabajo, se analiza la evolución de la eficiencia técnica de los hospitales públicos españoles durante el período 1991-93 mediante el análisis envolvente de datos, y se investigan los determinantes de esa eficiencia y su relación con los costes unitarios. Además, se estima un modelo de frontera de costes con objeto de obtener una medida alternativa de eficiencia. Ambos procedimientos muestran mejoras significativas de eficiencia tras la introducción de los contratos-programa en 1992.

Recepción de original julio de 1995
Versión final, septiembre de 1996