"Technical Efficiency of Hospitals in Ireland"

Brenda Gannon

Working Paper 18

Research Programme on "Health Services, Health Inequalities and Health and Social Gain"

This programme is supported by the Health Research Board, and is being carried out by researchers at the Economic and Social Research Institute (ESRI), University College Dublin and the University of Ulster.

• Economic and Social Research Institute, 4 Burlington Road, Dublin 4, Ireland.
Correspondence to:Brenda Gannon, Economic and Social Research Institute, 4 Burlington Road, Dublin 4, Ireland. Email: brenda.gannon @esri.ie

TECHNICAL EFFICIENCY OF HOSPITALS IN IRELAND*

Brenda Gannon

Economic and Social Research Institute,

November 2004

ABSTRACT

Similar to many other European countries, the funding system for Irish hospitals is partially based on casemix, whereby resources are redistributed annually to hospitals with greater efficiency. For this reason, accurate measurement of efficiency is essential, so in this paper, we use Data Envelopment Analysis and Stochastic Frontier Analysis to measure technical efficiency of acute public hospitals in Ireland between 1992 and 2000. Although previous research elsewhere has used DEA extensively, the more recent panel model approach to estimating the stochastic frontier has not been employed to the same extent. The results from both methodologies will provide estimates of average efficiency in the hospital sector in Ireland, and this is the first time an application of this type has been performed on input and output data from Irish hospitals. While our output measure is not adjusted for casemix, it is useful for obtaining initial estimates of technical efficiency in Irish hospitals, and we show that our results are in the same range as those obtained recently for hospitals in Northern Ireland. We then use a measure of output adjusted with Diagnostic Related Groups weights and determine whether or not efficiency estimates are different to those found when we use the simple measure of output. Internationally, the comparison of the results of the stochastic frontier model to the DEA efficiency scores contribute to the expanding literature of comparisons between DEA and SFA applications.

[WORK IN PROGRESS: NOT FOR QUOTATION WITHOUT PRIOR PERMISSION]

***This research forms part of the research programme "The Provision and Use of Health Services, Health Inequalities and Health and Social Gain", funded by the Health Research Board. Data on Diagnostic Related Groups was received from the Hospital Inpatient Enquiry system (HIPE).

1. Introduction

This paper is the start of a programme of research on hospital efficiency in Ireland and applies standard methods of efficiency analysis to a hospital setting. It aims to measure efficiency in acute hospitals in Ireland between 1992 and 2000. The limitations at this stage of the programme of research are firstly that our measure of output has not been corrected for differing procedures among hospitals. Once we obtain data on the type of procedures carried out by each hospital we can correct for this limitation. Secondly, the methods we apply are the basic methods introduced in the early eighties. Further refinements have been introduced since then, but for this first paper on hospital efficiency in Ireland we focus on the original methods. The results from this paper may be considered as initial estimates of efficiency, but these may change over the course of this programme of research.

Efficiency is the success of the hospital in using its resources to produce output – the degree to which the observed use of resources to produce output of a given quality matches the optimal use of resources to produce outputs of a given quality. This can be assessed in terms of technical efficiency or allocative efficiency. The latter type measures whether for any level of production, inputs are used in the proportion that minimises the cost of production, given input prices. Technical efficiency is concerned with the conversion of physical inputs such as labour services or raw materials into outputs. Cost efficiency refers to the combination of technical and allocative efficiency. We currently do not have access to cost data in hospitals so in this paper we focus on technical efficiency in acute hospitals in Ireland and use non–parametric and parametric techniques to measure performance of hospitals relative to best practice as opposed to the average practice.

Farell (1957) defined a simple measure of firm efficiency that could account for multiple inputs, stating that technical efficiency is the ability of a firm to obtain maximal output for a given set of inputs. His definition of technical efficiency led to the development of methods for estimating technical efficiencies in the context of a firm. Data Envelopment Analysis (DEA) is a non-parametric linear programming approach and was first introduced by Charnes, Cooper, and Rhodes in 1978 and further formalized by Banker, Charnes and Cooper in 1984. The technique was first used to study hospital production in 1986 (Banker, Conrad and Strauss) followed by

Grosskopf and Valdmanis in 1987. A number of more recent studies have employed DEA to measure hospital efficiency, Magnussen (1996), Hollingsworth and Parkin (1995), Ferrier and Valdmanis (1996), Parkin and Hollingsworth (1997) and Rosenman, Siddharthan and Ahern (1997). McKillop et al (1999) estimated the technical efficiency of all hospitals in Northern Ireland from 1986 to 1992. All acute hospitals were categorised into small, medium and large (based on total number of inpatients and outpatients). In Norway, Biorn, Hagen and Iversen (2002) measure technical efficiency of hospitals to test the hypothesis that hospital efficiency is expected to be greater with activity based funding of hospitals than with fixed budgets. They find that there was a large improvement in efficiency in the first year after the reform of the funding system.

An alternative method is the use of econometric models, and the development of the stochastic frontier model was first proposed by Aigner, Lovell and Schmidt (1977). Webster, Kennedy and Johnson (1998) use this approach and estimate a Cobb-Douglas production function and obtain the mean efficiency score for 301 hospitals between 1991 and 1995. They find that the efficiency scores under Stochastic Frontier Analysis (SFA) are lower than those using Data Envelopment Analysis.

The purpose of this paper is to estimate the variation in technical efficiency across hospitals in Ireland. Although, previous research in other countries has used DEA extensively, the more recent panel model approach to estimating the stochastic frontier, which we will describe later, has not been employed to the same extent. The results from both methodologies will provide estimates of average efficiency in the hospital sector in Ireland, and this is the first time an application of this type has been performed on input and output data from Irish hospitals. Internationally, the comparison of the results of the stochastic frontier model to the DEA efficiency scores will contribute to the growing literature on the comparison of DEA and SFA methodologies.

In Section 2, we provide a detailed account of the various types of hospitals in Ireland and define the measures of inputs and outputs. We then describe the Data Envelopment Analysis methodology and the Stochastic Frontier approach, and illustrate how they can usefully measure technical efficiency in Irish hospitals.

Section 3 explains these methods in detail. This is followed in Section 4 with a summary of the results. At this stage of our research, we rely mostly on crude measures of hospital output in order to illustrate this technique and its power, but future work will incorporate more complex output measures. However, we have access to preliminary data of this type for 1995 and in Section 5 we illustrate how using this data can change the efficiency scores. In section 6, we provide a summary and conclusions from the research to date.

2. Data on Acute Irish Hospitals

This paper analyses technical efficiency in acute hospitals during the period 1992-2000, selected on the basis of most recently available comparable data. The main question is how many years to combine together - data constraints and hospital closures determined that the frontier be constructed from three years' observations. Anything less would not solve the degrees of freedom problem and on the other hand if some observations of a hospital are too widely dispersed over time, this may lead to outlier or output-mix problems. Data constraints and some closures of hospitals mean that we now split the sample into 3 categories, analysing hospitals from 1992 to 1994, 1995 to 1997 and 1998 to 2000. This provides us with a sample of 62 hospitals during the period 1992-1994, 61 from 1995 to 1997 and 58[1] from 1998 to 2000.

By looking at all hospitals together and comparing the efficiency of one to the production frontier of all hospitals combined, we are ignoring the fact that there is much variation in size in terms of bed capacity and types of procedures carried out. For this reason, we split the sample into groups based on the type of hospital. Before we specify the numbers in each group, it is important to know how the hospital sector is organised in Ireland and what distinguishes these different groups. This is illustrated in Appendix Figure 1. This is a conventional categorisation in Irish terms, reflecting ownership as well as complexity. There may be a more appropriate categorisation for the purposes of efficiency analysis, where voluntary hospitals are not treated as a homogenous group and we will investigate this later in our programme of research.

[1] There was no data available for one hospital in 2000, therefore was excluded from the analysis for 1998-2000.

Regional hospitals, owned and funded by the Health Boards, provide the most comprehensive range of services and most of them are teaching hospitals. County hospitals, owned by the Health Boards, have consultant-staffed units of general medicine, general surgery, obstetrics and gynaecology, and separate children's wards. Voluntary public hospitals, supported but not owned by the Health Boards, are general hospitals that often function as teaching hospitals and are located mainly in Dublin and other large centres of population.

The next two groups of hospitals are self-explanatory providing maternity and orthopaedic services respectively. Finally we group all other types of hospitals together noting that different services may be provided by each, i.e. cancer, long stay, and infectious diseases.

The variables we use consist of inputs to hospital production in the form of capital and labour, and outputs from production. In terms of capital we use the average number of beds in the year in each hospital and this data is available in the Health Statistics publications. Labour inputs are measured by the number of people employed in each hospital and is counted in December of each year. We use full time equivalent staff to measure labour input. Given the shift work nature of hospitals and prevalence of part-time employment this will give a more accurate indication of the amount of labour used to provide services than does a simple count of the number of staff employed or the overall cost of labour. The employment figures are obtained from the Department of Children and Health Personnel Census and are categorised by General Support Staff, Health and Social Care Professionals, Management/Administration, Medical/Dental and Nursing. Table A1 provides an example of the breakdown of employment categories into grades of employees.

The outputs of hospital production initially consist of in-patients and out-patients cases. To measure in-patients we look at the total number of discharges and deaths within a year. Discharges are used instead of length of stay, as there may be huge differences across hospitals in efficiency in terms of occupancy rates and duration of stay. Out-patients are counted as total yearly number of attendances at consultant-controlled out-patient clinics in each hospital. Previous authors for example (McKillop et al., 1999) and Gregan and Bruce (1997) have questioned reliance on the number of discharged patients per hospital in measuring efficiency. They argue that the complexity of the casemix measured for example by employing Diagnostic Related Groups (DRG) should be taken into account. In fact, the Department of Health currently uses a casemix model to determine an element of hospital funding. Using casemix adjusted cost per case provides a means of relating hospital costs to workload, so may be considered a measure of relative efficiency, (Wiley, 2001). We intend to extend our analysis using information on DRGs in the future. However in this paper, we use the number of discharged patients in order to illustrate how DEA and SFA may be used to measure technical efficiency in the hospital sector in Ireland.

The classifications of hospitals not only allow us to distinguish between different services provided but also compare hospitals of similar sizes. For example across all hospitals the average number of beds in each hospital in each year is 194. But the range in terms of beds varies substantially, from 31 to 667. When we classify the hospitals into different groups we find that regional hospitals are the largest by far in terms of bed numbers. The next highest group is ‘general’, but the minimum value is quite low at 55. The reason for such a low number is because we include paediatric hospitals into the general category, as there are only three such hospitals. When we break down the years we will see that after 1997 when the National Childrens Hospital became part of Tallaght hospital, this minimum value increases.

Before proceeding with the analysis it is interesting to look at the trends over time in outputs from acute hospitals. Figure 1 shows the changes between 1992 and 2000 in average Length of Stay (LOS in days) and total number of in-patients and number of out-patients (millions) for all hospitals. This shows that the average length of stay fell slightly to 6.4 days by 2000 from 6.7 in 1992. At the same time, the number of in-patients increased slightly from 510,608 to 551,834 in 2000. The biggest increase in outputs occurred in out-patients clinics, rising from 1.8 million in 1992 to over 2 million in 2000.

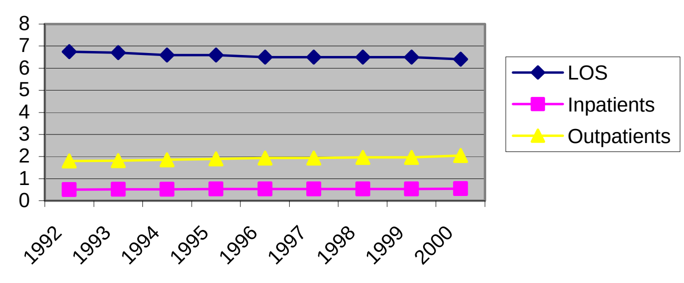


Figure 1:Outputs from Acute Hospitals 1992-2000

Inputs to hospital production also grew during this period, but Figure 2 shows that this was mainly due to increases in labour rather than capital (beds). The total number of beds for all hospitals remained fairly static and in fact decreased from 12,136 in 1992 to 11,891 by the year 2000. Employment on the other hand grew substantially, from approximately 27,500 to over 38,000.

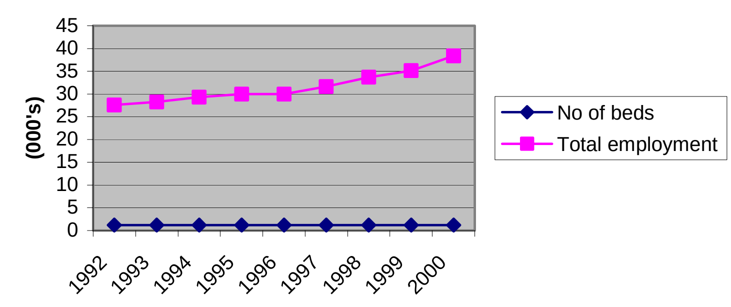
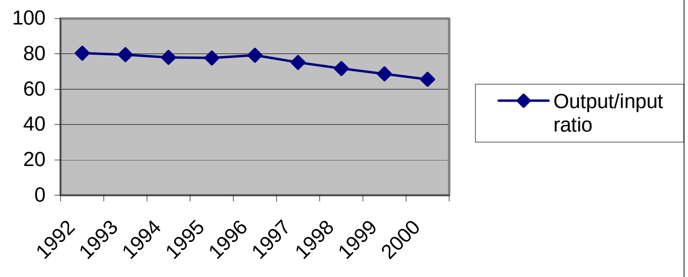


Figure 2:Inputs to Acute Hospitals 1992-2000

The simplest method of analysing hospital performance in terms of efficiency is to look at the rate of converting inputs into outputs for each hospital - Efficiency=output/input –and the higher this figure, the more efficient the hospital. In figure 3, we show the output/input ratios for all hospitals during 1992-2000. This shows how the efficiency ratio decreased from 80.4 to 65.6 by 2000.

Figure 3: Output/input ratio in Acute Hospitals 1992-2000



However, this figure does not distinguish hospitals in terms of size so is not very precise. Furthermore, inputs consist of total number of beds and employees and outputs consist of total number of in-patients and outpatients. There are usually a number of factors determining efficiency in hospitals, in terms of both inputs and outputs. For this reason, we need to employ more complicated measures of efficiency. In the next section we outline the methods we use to estimate technical efficiency.

3. Efficiency measurement using Data Envelopment Analysis and Stochastic Frontier Analysis

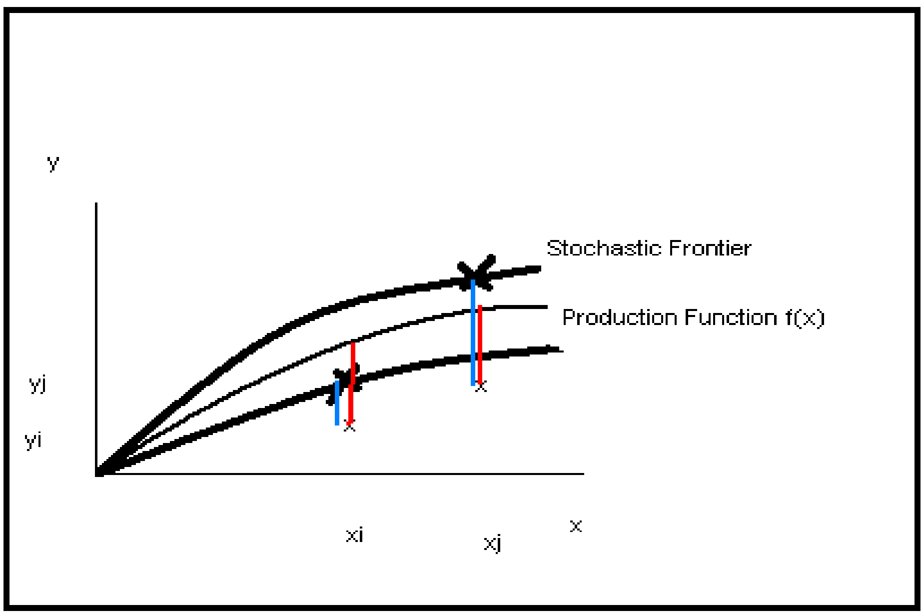
3.1 DEA v SFA:

'DEA and SFA employ quite distinct methodologies for frontier estimation and efficiency measurement, each with associated strengths and weaknesses, such that a trade off exists in selecting the correct approach', (Mortimer, 2002). Banker, Gadh and Gorr (1993) show that DEA is favoured when measurement error is an unlikely threat and where the assumptions of neoclassical production theory are questionable. SFA, on the other hand, deals with severe measurement error and where simple functional forms provide a close match to the properties of the underlying production technology. In estimating hospital efficiency, both measurement error and functional form are likely to cause problems. Therefore, what we require is an approach incorporating both. While both methods have been furthered advanced in recent years, there is no one standard of measuring efficiency. We concentrate in this paper on comparing the traditional methods of DEA and SFA.

The difference between DEA and SFA is described in Figure 4 below. Due to measurement error and other random factors affecting output, the stochastic frontier may differ from the best practice DEA frontier. For example, if the error is negative, the stochastic frontier will lie below the deterministic frontier. Using DEA we measure inefficiency as the distance from the estimated production function f(X) and the x produced by xi inputs, (measured by the red line). Using SFA, the estimated frontier lies below and the distance from it to x is shorter (measured by the blue line). In this case, DEA will result in a higher estimate of inefficiency. Deviations from the production frontier are due to inefficiency. With SFA however, deviations may also arise from a stochastic error. The distance between the DEA frontier and the SFA frontier represents this stochastic error. If the error is positive, the stochastic frontier will lie above the DEA best practice frontier, and DEA will result in lower estimates of inefficiency (higher efficiency).

The two frontiers are estimated by different methods, and in the next section we describe these in detail.

Figure 4: Comparison of DEA (estimated production frontier) and Stochastic Frontier



3.2 (a) Data Envelopment Analysis

Data Envelopment Analysis (DEA) is a linear programming technique which identifies best practice within a sample and measures efficiency based on differences between observed and best practice units, and is typically used to measure technical efficiency. DEA is used to estimate best practice in a sample by estimating a production frontier, and constructs a piece wise linear approximation of the efficient frontier (isoquant). A distribution of sample points is observed and a kinked line constructed around the outside of them, 'enveloping'.

There are several advantages in using DEA, in particular that the efficiency measure is related to best practice, not average practice. Being a non-parametric approach, it also does not require the use of a pre-specified functional form for technology nor distributional assumptions about error terms. Furthermore, data on costs is often difficult to obtain but the DEA method does not require cost minimisation.

In simple terms, based on constant returns to scale, if we have two inputs, beds and employees (Capital and Labour), we construct the number of inputs per treated case and plot the combinations of each input. A kinked frontier is constructed by enveloping all the input points. Hospitals on the frontier are efficient and the technical efficiency score may be constructed from the ratio of total inputs in best practice to total observed inputs. If this score were for example 0.66, then the hospital would have to reduce inputs by one third, while maintaining current output to attain the highest level of performance observed in the sample. If a hospital is looking for a role model, they would examine the performance and inputs of the more efficient best practice hospitals that are similar to itself – these are know as Peers in DEA. Sometimes, a hospital may be required to reduce inputs in different proportions but still maintain the same level of output. This special case is known as input Slack.

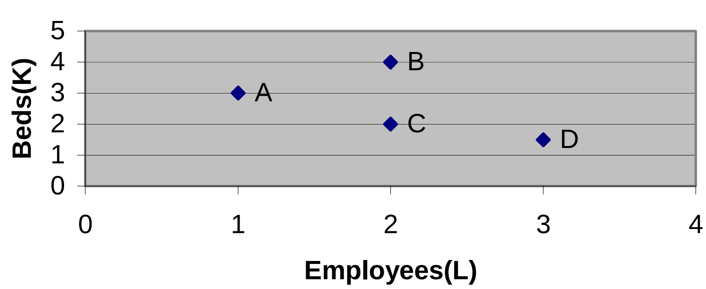
In Table 1 we provide data to illustrate a simple example of DEA. Employees and beds are inputs and in-patients discharged are the output. To facilitate comparisons input levels are converted to those needed by each hospital to produce one case of output.

Table 1: Example of inputs and outputs required to construct frontier

Hospital	Employees (Labour)	Beds(Capital)	In-patients Discharged(Output)	Labour per output	Capital per output
A	200	600	200	1	3
B	600	1200	300	2	4
C	200	200	100	2	2
D	600	300	200	3	1.5

Figure 5 then demonstrates the simplest case of Data Envelopment Analysis: if we plot each of these inputs on to a graph we obtain the efficiency frontier:

Figure 5: Efficiency points of 4 hospitals



The hospitals with less inputs are more efficient, and kinked frontier can be drawn from the lower points of A, C and D. This frontier 'envelopes' all the points and approximates a smooth isoquant. Any hospital on this frontier represents best practice and is efficient, but any hospital above will be inefficient – it can reduce input use and still maintain the output level compared with the performance of the best practice hospitals. In the case of hospital B, it could reduce its inputs by one third and reach the frontier at point between points A and C. The optimal use of total inputs is now 4 (2.66 beds and 1.33 employees). The observed use is 6 (4 beds and 2 employees). The efficiency score is measured as the ratio of optimal use to observed use, and is therefore 4/6, i.e. 0.67. In other words this hospital (B) should reduce its inputs by one third before it can reach the efficient frontier.

To get to this point, the inefficient hospital must reduce its inputs in equal proportions. If hospital 2 wishes to compare itself to other hospital, it should look to hospitals at points 1 and 3, as these are efficient hospitals that are closest to it in terms of output. These are known as 'peers'. Even if a hospital is 100 per cent efficient, but there are not many peers in the group, there may be scope for them to increase efficiency. The efficiency results can only be compared to those within the group – if we have another group of hospitals, these will appear on a different frontier and hence may only be compared to peer hospitals within their group.

This simple model is based on constant returns to scale (CRS), implying that the size of a hospital is not relevant when assessing efficiency. However it is likely that the size of the hospital will influence its ability to produce services more efficiently, and this assumption of CRS is no longer valid. A variable returns to scale frontier allows best practice level of outputs to inputs to vary with size of hospitals. The scale efficiency can then be determined by comparing technical efficiency scores of each hospital under CRS and variable returns to scale.

However, this example will not be possible when we extend the number of inputs and outputs involved and increase the number of hospitals. For this reason we need to solve a linear programming problem. When there are many inputs, and many hospitals we use a non-parametric approach to measuring efficiency using linear programming. Efficiency Management System (EMS) software may be used to produce efficiency scores, peer hospitals, and slacks. The standard approach is to use at least three times the number of hospitals as inputs/outputs. Other software includes DEAP, similar to EMS but also providing information on the optimal inputs required to reach technical efficiency. We use EMS software, as this allows us to directly perform window analysis, (a panel data technique we describe later in the paper). Alternatively Excel Solver may be used to solve the linear programming problem.

This problem is as follows: Starting with the most simplest form we assume constant returns to scale, and the objective for each hospital is to minimise inputs for a given level of output[2]:

[2] The decision to use an input orientated model should be based on whether or not the managers have much control over inputs or not. However, in many instances the choice or orientation has only a minor

Minimise E_n with respect to $w_1, \ldots w_{N,}, E_n$

subject to $\sum_{j=1}^{N} w_j y_{ij} - y_{in} \geq 0$ i=1,…,I

$$\sum_{j=1}^{N} w_j x_{kj} - E_n x_{kn} \leq 0 \quad k=1,\ldots,K$$

$$w_j \geq 0 \quad j=1,\ldots,N^3 \qquad [1]$$

There are N hospitals producing i different outputs y_{in} for i=1,…,I using K different inputs x_{kn} for k=1,…,K .

The w_j are weights applied across the N hospitals. When the linear program is being solved for hospital A, these weights allow the most efficient way of producing hospital no. 1 outputs to be determined. The efficiency score for this hospital is E_1^* i.e. the smallest number E which satisfies the constraints.

The linear program solves for the convex combination of the N data points (frontier) that can produce at least the observation n output and use at most E_n^* times the observation n combination of inputs. To get efficiency scores for each hospital, the problem must be solved N times. The efficiency score should be minimised and this is achieved by varying the weights and the efficiency score itself. The weights are used to construct the hypothetical best practice for the hospital on the frontier.

This method assumes constant returns to scale, but there may be diseconomies of scale or economies of scale. If we include variable returns to scale in the model, then there is an additional constraint that the weights should equal to 1. This has the effect of pulling a tighter frontier to envelope the data. The measure of scale efficiency is then derived by the ratio of constant returns to variable returns efficiency scores. If this ratio is less than 1, then the hospital is too small or too large relative to its

---

influence on the scores obtained (Coeilli and Perelman, 1996) so we follow Biorn et al (2002) and use an input orientated model.

[3] Constraint 1 is that the weighted average of other hospitals must produce at least as much output as hospital n. Constraint 2 is that the weighted average of other hospitals should not use any more of input than does hospital n. Constraint 3 limits the weights to be zero or positive.

optimum size. The final scores on technical efficiency indicate how efficient a hospital is relative to best practice. For example, a score of one equals total efficiency, and a score of less than one represents inefficiency. The values on slacks in each hospital show how much of each input can be reduced in order to reach the same level of output.

3.2 (b) Window Analysis

Generally the combined number of inputs and outputs should be 3 times greater than the number of hospitals under analysis. Given that we have around 60 hospitals to analyse each year, 6 inputs and 1 output, this condition should be satisfied. In this paper however we categorise the hospitals by type, we will be left in some cases with only 4 or 5 hospitals to measure efficiency in any year. In this scenario, DEA would classify a number of hospitals as technically inefficient just because the total number of outputs and inputs is too large relative to the number of hospitals, (Nunamaker, 1995). This degrees of freedom problem may be solved by using pooled data – i.e. we can combine data on the same hospitals for a number of years. Furthermore, using panel data where each hospital is treated as a separate observation in each year may be unrealistic as there may be technological improvements over time, leading to higher efficiency scores in later years. In this case we should use a window approach rather than a panel analysis. Within DEA analysis this technique is known as 'window analysis' and was introduced by Charnes et al. (1985). For example, we could choose a window width of 4 years, so the hospitals are compared to other hospitals within that time span only.

The frontier is constructed from the total number of hospitals in each year. For example if there are 6 hospitals with data over 3 years, we now have 18 observations. If we choose a 2 year window width, each hospital is given an efficiency score by running DEA for each observation relative to the frontier constructed from 12 observations. The average of these two scores is the efficiency measure for that hospital during that three year period.

3.2 (c) Stochastic Frontier Analysis

While DEA does not separate out the effects of a stochastic error term, SFA disentangles the two sources of error, due to (1) inefficiency and (2) random noise. The Stochastic Frontier may be estimated by Maximum Likelihood, or OLS in a cross section[4]. However, Schmidt and Sickles (1984) observed that when panel data are available there is not need to specify a particular distribution for inefficiency effects, and measures of technical efficiency are obtained relative to the most efficient firm/hospital. Panel data estimators are preferable as they are less likely to yield biased estimates of the parameters of inputs due to omitted variables and require fewer distributional assumptions about the deterministic error. Skinner (1994) also suggests using panel data methods to estimate a fixed effect for each hospital. The results may then be used to make conservative assessments of how each hospital differs from the best practice hospital. We can also infer statistical significance using this method. For these reasons, we estimate a panel model with fixed effects.

In general, following Schmidt and Sickles (1984) the model is specified as:

$$y_{it} = a + x_{it}\beta + v_{it} - u_{it} \qquad [2]$$

where y_{it} is the output for hospital i in time t and x_{it} is a vector of inputs, and u_i represents technical efficiency and v_{it} is the random error term.

In previous studies of hospital efficiency the parametric production function has been represented by a Cobb Douglas function. (Mortimer, 2001 etc). We follow this approach, convert all inputs and outputs to logs and estimate:

$$\ln y_{it} = a_i + \ln x_{it}\beta + v_{it} \qquad [3]$$

We derive estimates of u_i by calculating $u_i = \hat{a} - \hat{a}_i$ where i=1,2,3….N.

[4] See Coeilli, Rao and Battesse (2002) for an overview of these methods.

4. Results

4.1 Data Envelopment Analysis

In comparing mean efficiency scores from two studies we can only look at the dispersion of efficiencies within each sample – i.e. the scores cannot tell us anything about the efficiency of one sample relative to another. In Section 2 we noted that there were 6 inputs - number of beds and five types of labour. In this analysis on efficiency from 1992 to 2000, we cannot use the various categories of labour as inputs, due to the degrees of freedom problem when using sub-periods of 3 years. For this reason, we combine labour inputs into medical and non-medical. In Table 2, we present results from a window analysis in each sub-period, using a window of two years and assuming variable returns to scale[5].

Table 2: DEA Mean Efficiency Scores (Window panel width=2 years, assuming Variable Returns to Scale)

	Regional	County	General	Maternity	Orthopaedic	Other
1992-1993	0.95	0.91	0.85	0.92	0.87	0.89
1993-1994	0.94	0.91	0.84	0.92	0.75	0.91
N	6	25	12	6	6	6
1995-1996	0.89	0.90	0.86	0.98	0.90	0.97
1996-1997	0.97	0.91	0.87	0.98	0.90	0.97
N	6	25	12	5	5	6
1998-1999	0.96	0.89	0.92	0.95	0.82	0.97
1999-2000	0.98	0.88	0.92	0.98	0.82	0.97
N	6	25	12	4	5	6

The results presented in table 2, are based on an output variable – total inpatients and outpatients. No allowance is made for the fact that there are different resources used for these two types of patients and that the cost of an outpatient is much lower compared to that of an inpatient. Incorporating this into the DEA analysis is complicated as we do not know exactly what relative cost to attach to each type of patients. We begin by repeating the analysis above and use two output variables – inpatients and outpatients. We then include weights in the model. In terms of the

[5] Due to hospital closures and amalgamations between 1992 and 2000, we split the timeframe into 3 sub-periods in order to look at the same hospitals in each sub-period. Given the different sizes of hospitals we also split the groups based on the type of services provided, i.e. regional etc.

model presented in section 3.2(a) this is achieved by including a restriction into the linear programming model. For example, if we assume that one inpatient represents between 9 and 10 outpatients, we add in two restrictions, (1) $q_1 - 9q_2 \geq 0$ and (2) $-q + 10q_{21} \geq 0$, where q represents the output weight. Given that we do not know the exact weight to attach to inpatients or outpatients, we do several variations of the DEA analysis by attaching varying weights: [1] cost of inpatients is greater than outpatients, but we do not specify by how much, [2] one inpatient equals five outpatients and [3] one inpatient equals ten outpatients. This analysis is for illustrating how different weights affect the mean efficiency scores, and we concentrate on the efficiency of regional, general and county hospitals for the years 1998-2000. The results for regional hospitals are presented in Table 3.

The first column repeats the efficiency score presented in Table 2. In column 2, we use two output variables and in the following columns we attach different weights to outpatients. The results show that by allowing for outpatients to cost less than inpatients, the mean efficiency scores are no different to when we include no weights. However, the less costly an outpatient is relative to an inpatient, the lower the efficiency score between 1998 and 1999. Overall though, we do not see much change in the results compared to when no weights were used. We should bear in mind that the sample of regional hospitals is small, and in DEA we cannot infer statistical significance from the efficiency of any hospital.

Table 3: DEA efficiency scores for Regional Hospitals

	DEA Inpatients and outpatients combined – no weights	**DEA Inpatients and outpatients separate – no weights**	**DEA Inpatients and Outpatients – weights [1,-1]**	**DEA Inpatients and Outpatients –weights [-4,5]**	**DEA Inpatients and Outpatients –weights [-9,10]**
1998-1999	0.96	0.96	0.96	0.96	0.95
1999-2000	0.98	0.99	0.99	0.98	0.98

In the next table, we continue to report DEA results and focus on general hospitals. There appears to be less variation in efficiency scores, once we separate the output variable into inpatient and outpatients. Using weights, but not attaching a particular value to each output, makes no difference to the average efficiency score. When we introduce more specific weights, the efficiency scores are lower. This suggests that when counting inpatients and outpatients as equal outputs, on average general hospitals are highly efficient at 0.99. Even when we control for weights, the average efficiency score is still quite high at 0.97 in 1998-1999 and 0.98 in 1999-2000.

Table 4: DEA Efficiency Scores for General Hospitals

	DEA Inpatients and outpatients combined – no weights	DEA Inpatients and outpatients separate – no weights	DEA Inpatients and Outpatients – weights [1,-1]	DEA Inpatients and Outpatients –weights [-4,5]	DEA Inpatients and Outpatients –weights [-9,10]
1998-1999	0.92	0.99	0.99	0.96	0.97
1999-2000	0.92	0.99	0.99	0.97	0.98

We turn now to county hospitals and find that the mean efficiency score is quite similar regardless of what weight is applied. Even when we weight inpatients as 9-10 outpatients, the efficiency score is almost the same as when we apply weights but attach no particular value. It may be that we need to define a different weight for outputs in county hospitals, depending on what procedures are employed. Data on DRG's would be particularly valuable for this analysis of county hospitals.

Table 5: DEA Efficiency Scores for County Hospitals

	DEA Inpatients and outpatients combined – no weights	DEA Inpatients and outpatients separate – no weights	DEA Inpatients and Outpatients – weights [1,-1]	DEA Inpatients and Outpatients –weights [-4,5]	DEA Inpatients and Outpatients –weights [-9,10]
1998-1999	0.89	0.90	0.89	0.90	0.89
1999-2000	0.88	0.91	0.90	0.90	0.89

Even though the theoretical side of DEA[6] has developed rapidly over the last few years allowing us to infer statistical significance on efficiency scores, there is no such approach available yet for small samples. For this reason, we cannot infer statistical significance on any measure of efficiency from a hospital. Furthermore, deviations from the best practice frontier are attributed to inefficiency and ignore the possibility of random measurement error. As mentioned earlier, a parametric approach can help to eliminate this difficulty in efficiency measurement. We now provide estimates of technical efficiency, using Stochastic Frontier Analysis.

4.2 Stochastic Frontier Analysis

The SFA results for regional hospitals are presented in Table 6 below. In the first column, we see efficiency scores derived from the fixed effects model when we combine inpatients and outpatients with no weights attached. When we apply weights to the output variable, we see that the efficiency scores are now slightly higher at 0.94 and 0.95 for 1998-1999 and 1999-2000 respectively. This score is much the same, when we apply a larger weight to outpatients. The importance of these results is that they are lower than those found using DEA. This suggests that the SFA results have removed any random noise that had been included in the DEA efficiency scores.

Table 6: SFA Efficiency Scores for Regional Hospitals

	Fixed Effects Inpatients and Outpatients combined	Fixed Effects Inpatients and outpatients – weights [Outpatient=0.1 Inpatient]	Fixed effects Inpatients and outpatients combined-weights [outpatient=.0.2 inpatient]
1998-1999	0.91	0.94	0.93
1999-2000	0.93	0.95	0.95

Looking at Table 7, we find a similar result for general hospitals, where now the efficiency scores are much lower. The importance of applying weights to the outputs can be seen in the final two columns. There we see that the efficiency score between

[6] See Simar and Wilson (2000).

1999-2000 is much lower at 0.84, compared to 0.89 when we do not apply weights. Similar to the case of regional hospitals, there is no real difference in results when we increase the weight of an outpatient, relative to an inpatient.

Table 7: SFA Efficiency Scores for General Hospitals

	Fixed Effects Inpatients and Outpatients combined	Fixed Effects Inpatients and outpatients – weights [Outpatient=0.1 Inpatient]	Fixed effects Inpatients and outpatients combined-weights [outpatient=.0.2 inpatient]
1998-1999	0.87	0.86	0.87
1999-2000	0.89	0.84	0.84

Finally, we look at the SFA results for county hospitals in Table 8 and see a similar pattern as before, i.e. the scores are lower once we apply weights to hospital output. In comparison though, the results for 1999-2000 are much lower when we increase the relative weight of an outpatient. The difference is that in 1999-2000, the efficiency score is much lower at 0.74, when we allow one outpatient to equal 0.2 of an inpatient. Again, this may reflect the type of procedures carried out in county hospitals, and we cannot determine exactly the difference between an inpatient and outpatient without information on DRGs. Data on DRG values for inpatients is available from HIPE (see section 5), but there is no such information available for outpatients in each hospital.

Table 8: SFA Efficiency Scores for County Hospitals

	Fixed Effects Inpatients and Outpatients combined	Fixed Effects Inpatients and outpatients – weights [Outpatient=0.1 Inpatient]	Fixed effects Inpatients and outpatients combined-weights [outpatient=0.2 inpatient]
1998-1999	0.78	0.78	0.78
1999-2000	0.82	0.79	0.74

4.3 Summary

The results from the DEA and SFA approaches to measuring the production frontier and efficiency, suggest that efficiency of hospitals may vary depending on whether or not a parametric approach is employed, and also on the relative weights attached to an outpatient. Focusing on 1999-2000, for regional hospitals, the mean efficiency score ranges from 0.92 to 0.99. The scores for county hospitals range between 0.74 and 0.91 and for general hospitals are between 0.84 to 0.99. In all cases, the DEA efficiency scores are higher, indicating that inefficiency (deviation from the best practice frontier) is lower than inefficiency measured by SFA. This means that the stochastic frontier must lie above the DEA best practice frontier. The differences between the two efficiency measures calculated from DEA and SFA is quite high for county hospitals (0.15 efficiency points) and for general hospitals (0.13 efficiency points).

The results from this analysis suggest that the types of hospitals with the highest level of efficiency are regional. On average, regional hospitals between 1999 and 2000 could use inputs to increase efficiency by only 1-8 per cent. County hospitals could increase efficiency by around 9-26 percent, between 1999 and 2000. Likewise the inputs of general hospitals could be used more efficiently (by 1-16 per cent) between 1999 and 2000.

Focusing on the DEA efficiency scores, these results are in the same range as those obtained by McKillop et al (1999) for hospitals in Northern Ireland. For example, larger hospitals showed an average score of 0.93 for the years 1989-1992, assuming constant returns to scale, and 0.99 with variable returns to scale. These results are in the same region as our efficency scores for regional hospitals. Likewise in Finland, Linna and Hakkinen (   ) found that the average level of technical efficiency for all hospitals was 0.95 with variable returns to scale and 0.91 when assuming constant returns to scale. However, we cannot compare these scores to the results shown in Table 2, as they do not specify the size of the hospitals.

In terms of DEA v SFA, other research has shown differences of up to 0.11 efficiency points at low levels of measurement error and up to 0.40 efficiency points with high levels of measurement error, (Banker, Gadh and Gorr, 1993). Yu (1998) found a difference of 0.16 efficiency points, even with high levels of measurement error. Both of these studies used simulation methods to test the difference between DEA and SFA. Mortimer (2001) emphasised the need for real-world comparisons to determine the relative precision (and policy value) of DEA and SFA. In this context, the 'real-world' results from this paper contribute to the expanding literature of comparisons between DEA and SFA applications.

5. Output adjusted for DRG

The results presented so far are based on the head count of individuals discharged from hospitals. As we mentioned earlier in section 2, this may not represent the true output from hospitals if some are performing more costly or time consuming procedures relative to their peer hospitals. Using output data adjusted for these types of procedures means we can now control for this possibility. In each hospital, a discharge is recorded and given a DRG coding. There are over 500 different types of DRG so we weight the number of discharges in each DRG by the relative value (RV) of that DRG. The weighted numbers are then aggregated into a single output measure. This data is available to us from HIPE - a system that collates the records of each patient in each hospital on an annual basis. Many public hospitals participated on an annual basis between 1995 and 2003. So far, we have access to the 1995 data only, but we can use this data to illustrate how the efficiency scores change as a result of using the new adjusted output measure compared to the output variable we have used so far in this paper.

Given that we need sufficient hospitals to overcome the degrees of freedom problem, we focus in this example on three different groups – all hospitals, general hospitals and county hospitals. In order to get a sample that is comparable using the two different output measures, we had to exclude some hospitals. Firstly, some hospitals did not return records for all inpatients discharged, particularly the maternity and long stay hospitals. By including these hospitals, we may distort the efficiency result, so we exclude any hospitals where coverage is less than 80 per cent. Our final samples have 53 hospitals in total, 12 general and 24 county hospitals.

Table 9 shows that when we use the unadjusted output measure, the efficiency scores are lower in all three groups (using the DEA method).

Table 9: Comparison of Technical Efficiency using Regular and Adjusted Output Measures (Variable Returns to Scale), 1995.

	All	County	General
Regular	0.83	0.91	0.93
Adjusted	0.87	0.92	0.96
N	53	24	12

The efficiency scores from the adjusted data support the hypothesis that some hospitals may be rated with lower efficiency levels if output is not adjusted for the value of each DRG. In Table 10, we show how the number of hospitals near the efficient frontier increases after we adjust the output data. For example, in county hospitals we now see 13 efficient (100%) hospitals relative to best practice in that group, compared to 11 when we use the regular output measure.

Table 10: Number of efficient hospitals, 1995

	All		County		General	
	Regular	Adjusted	Regular	Adjusted	Regular	Adjusted
100%	10	17	11	13	8	7
90-99	13	9	4	3	1	4
80-89	13	14	5	4	1	1
<80	15	11	3	4	2	0

This illustrative comparison of output measures supports the view that hospital output in DEA should reflect the type of procedure carried out in each hospital. We use data from 1995 in this example, but aim to use data from 1995-2000 in future work.

5. Conclusions

The analysis of efficiency in hospitals can make a major contribution to improving health services. This paper shows how one can measure technical efficiency relative to best practice using Data Envelopment Analysis and Stochastic Frontier Analysis. Using a crude measure of output, it demonstrates how this may be applied in the context of the hospital sector in Ireland. Compared with the most efficient hospitals within their categories, our results show that regional hospitals are highly efficient while county and general hospitals are less efficient. It is not possible to compare across categories and our analysis provides an average of best practice performance, where the focus is on spread within categories only. In future research, by analysing total factor productivity, we will be able to measure changes in efficiency over time, and decompose this change into technological change and pure efficiency change.

In future work, we also intend to incorporate more complex measures of output based on categorisation of patients in terms of casemix. In section 5 we illustrated the importance of using an output measure adjusted for casemix, so we will aim to include these output measures for all years from 1995-2000.

Our comparison between DEA and SFA methods, shows that there may be high levels of random error in the data, leading to lower efficiency scores under the SFA method. Using other methods (ML) and other functional forms of the production function to estimate the parametric distance functions we could also check the robustness of these findings to the choice of methodology for efficiency scores.

References

Aigner, D.J., Lovell, C.A. and P. Schmidt. Formulation and Estimation of Stochastic Frontier Production Function Models. *Journal of Econometrics*, 1977:6:21-27.

Banker, R.D., A. Charnes and W.W. Cooper. Some models for Estimating Technical and Scale Efficiencies in Data Envelopment Analysis. *Management Science*, 1984;30:1078-1092.

Banker, R., Gadh, V. and W. Gorr. A Monte Carlo comparison of two production frontier estimation methods:Corrected ordinary least squares and data envelopment analysis. *European Journal of Operational Research*, 1993:67:332-343.

Biorn, E., Hagen, T, Iversen, T., and J. Magnussen. A panel data analysis of DEA efficiency scores 1992-2000. *Health Economics Research Programme, University of Oslo*, Working Paper 2002:8.

Charnes, A., W.W Cooper and E. Rhodes. Measuring the Efficiency of Decision Making Units. *European Journal of Operational Research*, 1978; 2:429-444.

Coeilli, T. and S. Peralman . A comparison of Parametric and Non-parametric Distance Functions: With Application to European Railways. *CREPP Discussion* Paper no. 96/11, University of Liege, Liege.

Coeilli, T., D. Rao and G. Battese. An Introduction to Efficiency and Productivity Analysis., 1998. Kluwer Academic Publishers.

Farrell, M.J. The Measurement of Productive Efficiency. *Journal of the Royal Statistical Society, Series A*, 1957; CXX, Part 3:253-290.

Grosskopf, S. and V. Valdmanis. Measuring Hospital Performance: A non-parametric approach. *Journal of Health Economics*, 1987; 6:89-107.

Health Statistics 1992-2000, Department of Health and Children.

Hollingsworth, B. and D. Parkin. The Efficiency of Scottish Acute Hospitals – an application of data envelopment analysis. *IMA Journal of Mathematics Applied in Medicine and Biology*, 1995; 12(3-4):161-173.

McKillop, D., J. Glass, C. Kerr and G. McCallion. Efficiency in Northern Ireland Hospitals: A Non-parametric Analysis. *The Economic and Social Review*, 1999; 30(2):175-196.

Mortimer, D. A Systematic Review of Direct DEA vs SFA/DFA Comparisons. *Centre for Health and Evaluation, Australia*, 2002, Working Paper 136.

Mortimer, D. Methods for the Measurement of Hospital Efficiency: A comparison of Frontier estimation techniques in a sample of Victorian public hospitals. 2001, Department of Economics, Monash University:Unpublished Masters of Economics thesis.

Parkin, D. and B. Hollingsworth. Measuring Productivity Efficiency of Acute Hospitals in Scotland 1991-1994:Validity Issues in Data Envelopment Analysis. *Applied Economics*, 1997; 29:1425-1433.

Rosenman, R., K. Siddharthan and M. Ahern, Output Efficiency of Health Maintenance Organisations in Florida. *Health Economics*, 1997;295-302.

Schmidt, P. and R. Sickles. Production Frontiers with Panel Data. *Journal of Business and Economic Statistics*, 1984:4(2):367-374.

Skinner, J. What do stochastic frontier cost functions tell us about inefficiency? *Journal of Health Economics*, 1994, 13:323-328.

Simar, Leopold and Paul W. Wilson (2000) Statistical Inference in Nonparametric Frontier Models. *The State of the Art in Journal of Productivity Analysis*. Vol. 13, S. 49 – 78.

Webster, R., Kennedy, S. and L. Johnson. Comparing Techniques for Measuring the Efficiency and Productivity of Australian Private Hospitals. *Australian Bureau of Statistics Working Paper No. 98/3,* 1998.

Wiley, M. Case Mix in Ireland: Budgeting Basis for Acute Hospital Services in F. France, I. Mertens, M. Closon and J. Hofdijk (Eds.), *Case Mix: Global Views, Local Actions*, 2001.

Table A1: Employment types in Hospitals by Grade

General Support Staff	Health and Social Care Professionals	Management/ Admininstration	Medical/Dental	Nursing
Catering/Cleaner/Assistant	Chiropodist	Administration Unclassified	Consultant/Anaesthetist	Assistant Director of Nursing (obsolete)
Catering Officer, Assistant	E.C.G. Technician, Senior	Clerical Officer	Consultant/Dermatologist	Clinical Nurse Manager 2
Catering Officer Grade I	Medical Laboratory Technician, Senior	Fianance Officer	Consultant/General Physician	Clinical Nurse Manager 2 - Home (obsolete)
Chaplain (obsolete)	Occupational Therapist	Grade III (obsolete)	Consultant/General Surgeon	Director of Nursing (obsolete)
Chef I	Pharmaceutical Technician	Grade II (obsolete)	Consultant/Geriatrician	Nursing Unclassified
Chef II	Pharmacist	Grade IV	Consultant/Neurologist	Principal Nurse Tutor
Cook (obsolete)	Phlebotomist	Grade V	Consultant/Obstetrician & Gynaecologist	Staff Nurse - General
Cook, Trainee	Physiotherapist	Grade VI	Consultant/Ophthalmic Surgeon	Student Nurse III (obsolete)
Diningroom Supervisor	Physiotherapist-in-Charge (Grade I) (obsolete)	Secretary/Manager	Consultant/Orthopaedic Surgeon	Student Nurse II (obsolete)
Electrician	Physiotherapist, Senior	Supplies Officer Grade B	Consultant/Otolaryngologist	Student Nurse I (obsolete)
Linen Room/Laundry Assistant	Radiographer	Telephonist	Consultant/Radiologist	Tutor Principal III (obsolete)
Maintenance Craftsman/Technician	Radiographer, Senior		Consultant/Rheumatologist	
Maintenance Officer/Supervisor	Radiographer, Supt. III (obsolete)		House Officer (obsolete)	
Nursing Auxiliary/Orderly (obsolete)	Radiographer, Supt. I (obsolete)		Intern	
Painter	Social Worker, Head Medical (obsolete)		Medical & Dental Unclassified	
Plumber	Social Worker, Medical		Psychiatrist	
Porter	Social Worker, Senior (obsolete)		Registrar	
Seamstress	Speech and Language Therapist			

FIGURE A1: STRUCTURE OF HOSPITAL GROUPS