TESTING FOR VARIATION IN TECHNICAL EFFICIENCY OF HOSPITALS IN IRELAND[*]

Brenda Gannon

The Economic and Social Research Institute

June 2005

ABSTRACT

Funding in Irish hospitals is partially based on casemix, whereby resources are redistributed annually to hospitals with greater efficiency. 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 1995 and 2000. The results provide estimates of average technical efficiency in the hospital sector in Ireland for the first time, and highlight the variation in technical efficiency levels across hospitals. Internationally, the results contribute to the growing literature on the comparison of DEA and SFA methodologies.

1. Introduction

In the production of health care, hospitals should act efficiently in terms of using their inputs to obtain maximum output. In many countries, efficiency in one year affects the budget of hospitals in the following year. Funding in public hospitals in Ireland is partially based on casemix[1], whereby 20% of the annual budget of a hospital is determined by their relative efficiency in the previous year. In this paper, we measure technical efficiency by assessing the production of physical output given the physical inputs of labour and capital. This may be viewed as a complement to the casemix measure in the overall determination of hospital efficiency.

Farrell (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) using data from a sample of hospitals in the U.S., followed by Grosskopf and Valdmanis in 1987. A number of more recent studies have also 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). 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. In Northern Ireland, McKillop et al (1999) estimated the technical efficiency of all hospitals from 1986 to 1992. All acute hospitals were categorised into small, medium and large (based on total number of inpatients and outpatients). In the Republic though DEA has not yet been applied to data from hospitals.

An alternative approach to studying efficiency is based on the use of econometric models, in particular the development of the stochastic frontier model first proposed by Aigner, Lovell and Schmidt (1977). Webster, Kennedy and Johnson (1998) used this approach to estimate a

---

[1] Casemix categorises each hospitals caseload into discrete groups, and this allows the comparison of activity and costs between hospitals.

Cobb-Douglas production function and obtained the mean efficiency score for 301 hospitals in Australia between 1991 and 1995. They find that the efficiency scores under Stochastic Frontier Analysis (SFA) are lower than those using Data Envelopment Analysis. 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. Once again, this approach has not been applied to Republic of Ireland hospitals.

The purpose of this paper is to estimate the variation in technical efficiency across hospitals in Ireland, and to investigate the robustness of these two different methodologies. The results from both methodologies will provide estimates of relative efficiency in the hospital sector in Ireland, and is the first time an application of this type has been performed on input and output data from Irish hospitals. This allows us to capture the variation in efficiency in terms of producing treated cases[2] and pin-point where efficiency is lower compared to best practice. It does not however allow efficiency in Ireland to be compared with efficiency elsewhere. Internationally, the comparison of the results of the stochastic frontier model to the DEA efficiency scores contributes 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 in Section 3 the Data Envelopment Analysis methodology and the Stochastic Frontier approach, and illustrate how they can usefully measure technical efficiency in Irish hospitals. This is followed in Section 4 by empirical results, and conclusions are presented in Section 5.

2. Data on Acute Irish Hospitals

This paper analyses technical efficiency in acute public hospitals during the period 1995 to 2000, selected on the basis of most recently available comparable data. There are over 60 hospitals but 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

---

[2] McKillop et al (1999) discuss the issue of quality of output. Most research has relied on output quantity measures but do not control for quality of outcomes. We follow a similar approach.

in each group, it is important to know how the hospital sector is organised in Ireland and what distinguishes these different groups. We discuss the conventional categorisation in Irish terms, reflecting ownership as well as complexity.

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. 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 in future work. The remaining three groups are Maternity, Orthopaedic and Other (includes long term illness) but because there are only a few hospitals in each of these groups, in this paper we focus only on Regional, General and County hospitals.

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. For the purposes of our model, we combine these categories into two groups, Medical and Non-Medical.

The outputs of hospital production consist of in-patients, out-patients and day cases. To measure in-patients we look at the total number of discharges 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). The headcount of patients discharged 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 the Hospital In-Patient Enquiry (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 2005 and so far, we have access to the 1995 to 2000 data. Preliminary research using this data has shown that there is less variation in efficiency once we account for different procedures among hospitals, and highlights the need for using casemix adjusted data in analysing hospital efficiency (Gannon, 2005). In this paper, we use the DRG adjusted measure of inpatient and day cases. The HIPE system does not adjust the number of outpatients for casemix, so we use unadjusted outpatients data.

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 in our sample the average number of beds (including daybeds) in each hospital in each year is 268. But the range in terms of beds is quite substantial, going from a minimum of 69 to a maximum of 712 beds. When we classify the hospitals into different groups we find that Regional and General hospitals are by far the largest in terms of bed numbers. In our final sample, we have data for 22 County hospitals. We only have data for 6 Regional and 5 General hospitals so because they are similar in size we combine them to facilitate our analysis. Our dataset includes 6 years of panel data, following the same hospitals in each year – providing 132 and 78 observations in each group.

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.

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 'enveloping' constructed around the outside of them.

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. Because we have data on many inputs and outputs for many hospitals, we need to solve a 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[3]:

Minimise E_n with respect to $w_1, \ldots w_N, E_n$

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

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

$w_j \geq 0 \quad j=1,\ldots,N$[4] $\hspace{6cm}$ [1]

In this model 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 n, these weights allow the most efficient way of producing hospital n's 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 model is solved 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 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 - this assumption of CRS is not 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

---

[3] The decision to use an input orientated model versus an output oriented one, should be based on whether or not the managers have much control over inputs. However, in many instances the choice or orientation has only a minor influence on the scores obtained (Coeilli and Perelman, 1996) so we follow Biorn et al (2002) and use an input orientated model.

[4] Constraint 1 requires that the weighted average of other hospitals must produce at least as much output as hospital n. Constraint 2 means 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.

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 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.

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 35 hospitals overall to analyse each year, 3 inputs and 3 outputs, this condition should be satisfied. In this paper we categorise the hospitals by type, so with the combined group of Regional and General hospitals, there are 13 hospitals but a total of 6 inputs and outputs. 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). Using pooled data may solve this degrees of freedom problem – i.e. we can combine data on the same hospitals for a number of years.

We could use panel data where each hospital is treated as a separate observation in each year and compare efficiency of one hospital to hospitals in all years. However, this 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)[5]. For example, we could choose a window width of 2 or 3 years, so the hospitals are compared to other hospitals within that time span only.

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[6]. However, Schmidt and Sickles (1984) observe that when panel data are available there is no 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

---

[5] Estimates are derived using *EMS* software.
[6] See Coeilli, Rao and Battesse (2002) for an overview of these methods.

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. 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, our first model is a panel model with fixed effects.

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

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

where y_{it} is the output for hospital i in time t and x_{it} is a vector of inputs, and u_i represents hospital specific fixed effects or time invariant technical efficiency and v_{it} is a normally distributed random error term and is uncorrelated with the explanatory variables. There is no distributional assumption on the u_i, distinguishing this model, so estimation is straightforward[7]. Returns to scale are estimated by the value of β.

We derive estimates of u_i by calculating $u_i = \hat{a} - \hat{a}_i$ where i=1,2,3....N. Technical efficiency is derived as:

$$TE = \exp(-u_i) \qquad [3]$$

To apply this model to our data, we need to specify an appropriate functional form for the production of hospitals output. In previous studies of hospital efficiency the parametric production function has been represented by a Cobb Douglas function, (e.g. Mortimer, 2001), representing unitary elasticity of substitution. While the Cobb Douglas form is easy to estimate, its main drawback is that assumes constant input elasticities and returns to scale for all hospitals. The Translog model does not impose these restrictions but is susceptible to degrees of freedom problems and multicollinearity. A number of studies estimated both the Cobb Douglas and Translog functional forms and then tested the null hypothesis that the Cobb-Douglas form is an adequate representation of the data. We follow a similar approach.

The model outlined above (following Schmidt and Sickles, 1984) assumes that the inefficiency effects are time invariant. In a panel with many years, this assumption may be questionable, so we also estimate a model that assumes time-varying inefficiency for

---

[7] Estimation is carried out using *Stata* Version 8.

comparison purposes. Cornwell, Schmidt and Sickles (1990) propose a model to replace a_i with a firm specific function of time, where $a_{it} = a_{i1} + a_{i2t}t + a_{i3}t^2$. This model allows for temporal variation of the firm effects, but the cost is that there may be too many parameters to estimate, (Coelli, Rao and Battese, 2002). Battese and Coelli (1992) recommended an alternative approach, a time varying model for technical inefficiency effects in the stochastic frontier production function for panel data. The technical inefficiency effects are defined by:

$$u_{it} = \{\exp[-\eta(t-T)]\}u_i, i = 1,2...,N; t = 1,2...,T \qquad [4]$$

where the u_is are assumed to be i.i.d. as the generalised truncated-normal random variable and η is an unknown parameter to be estimated. The drawback of this model, (compared to Cornwell, Schmidt and Sickles 1990) is that it does not account for situations in which some firms may be relatively inefficient initially but become more efficient in subsequent years. Nonetheless it provides us with an indication of the presence of any time variation in inefficiency. So, we apply this time varying model to the Cobb Douglas, Translog and production functions and then test the restrictions on each model to determine a preferable functional form for the production function. Using the Battese and Coelli (1992) model[8], we can also test the null hypothesis of no change in technical efficiency effects over time $H_0 : \eta = 0$. A generalised likelihood ratio test compares the log likelihoods of the model with time varying inefficiency versus the model with time invariant inefficiency.

4. Results

4.1 Description of Data

Before proceeding with the analysis it is interesting to discuss the trends over time in production of treated cases in acute hospitals. The number of in-patients increased slightly from 510,608 to 551,834 in 2000. The biggest proportional increase in outputs occurred in daycases, rising from 155,000 in 1992 to over 300,000 in 2000. Outpatients also increased significantly during this time. Inputs to hospital production also grew during this period, but 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

---

[8] All estimates are derived in *Frontier*.

to over 38,000. An additional input is the number of day beds, and the number available almost doubled between 1992 and 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=outputs/inputs – and the higher this figure, the more efficient the hospital. We calculate this ratio by simply adding all the inputs together and all outputs together, but this involves combining numbers employed with number of beds, and outputs are made up of outpatients, inpatients and day cases. Clearly, this is not a satisfactory way of measuring efficiency, so this ratio is only a descriptive measure of efficiency, and not very precise. In figure 1, we show the output/input ratios for all hospitals during 1992-2000 using data from Health Statistics publications. This shows how the efficiency ratio decreased from 63.5 in 1996 to 57.3 by 2000. We should bear in mind though that this ratio has been calculated for all hospitals, regardless of size and casemix of each hospital. The DEA and SFA measures of efficiency will take both of these factors into account.

Figure 1:Output/input ratio in Acute Hospitals 1992-2000
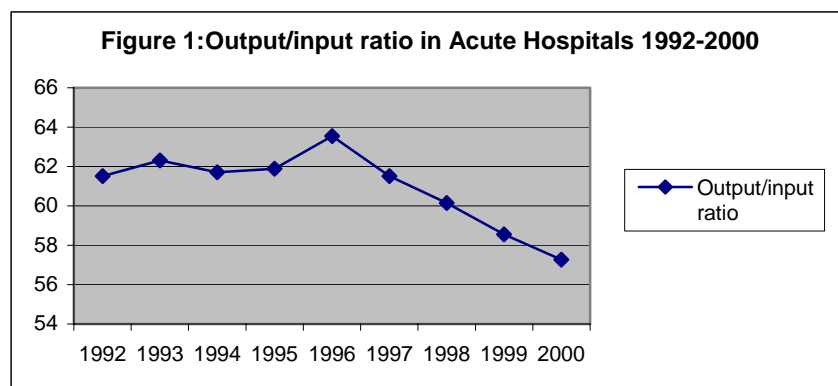
In the graph above, inputs consist of total number of beds and employees and outputs consist of total number of in-patients, outpatients and daycases. There are usually a number of factors determining efficiency in hospitals, in terms of both inputs and outputs. Because we do not want to treat all inputs as equal and all outputs as equal, we need to employ more complicated measures of efficiency and we present the results of the DEA analysis in the next section.

4.2 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 Table 1, we present results from a window analysis in each sub-period, using a window of two years and assuming variable returns to scale. This shows that for county hospitals the mean DEA efficiency score is 0.96 in 1995-1996 decreasing to 0.94 by 1999-2000. This suggests that there is variation in efficiency among county hospitals and by 1999-2000 on average the inefficient hospitals could have reduced their inputs by 6% and still achieved the same output in that year. In Regional and General hospitals there was less variation in efficiency over this period, but they could still reduce inputs by about 3 per cent to obtain the same output.

Table 1: DEA Mean Efficiency Scores (Window panel width=2 years, assuming Variable Returns to Scale)

County		Regional /General	
1995-1996	0.96	1995-1997	0.96
1996-1997	0.97	1996-1998	0.96
1997-1998	0.95	1997-1999	0.97
1998-1999	0.94	1998-2000	0.97
1999-2000	0.94		
	22		
N hospitals	44	N hospitals	13
N in window		N observations in window	52

The results in Table 1 show two important aspects of efficiency in hospitals between 1995-2000. Firstly, they indicate inefficiency across hospitals – the scores given indicate the mean efficiency within each group but they do not show that there is substantial variation within the group. For example our efficiency scores for each hospital (not reported here) show that during the years 1999-2000 there is a range of 0.16 within county hospitals, with the lowest scoring hospital obtaining a relative efficiency score of 0.84. This variation is substantial in Regional and General hospitals also, with the lowest relative efficiency score estimated at 0.81.

The second point is that efficiency seems to get lower over time for county hospitals and increases over time for regional and general hospitals. These are average scores, but on investigation of the data there were many positive and negative changes in efficiency in

several hospitals. It could be that is purely due to efficiency change or this could be due to technological change. Decomposition of productivity change into these two components is possible using DEA and SFA methods. However, the focus of this paper is on exploring the various models to derive average efficiency scores, and we will return to the role of technological change in future research.

Even though the theoretical side of DEA[9] 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 address these difficulties in efficiency measurement, so in the next section we provide estimates of technical efficiency using Stochastic Frontier Analysis.

4.3   Stochastic Frontier Analysis

Following Sickles and Schmidt (1984), we estimate a Cobb Douglas production function [based on equation 2] and derive inefficiency scores based on the value of the intercept for each hospital following equation 3. This model assumes time invariant inefficiency but as we have seen already (using DEA estimates) efficiency may change over time so we relax this assumption later on in the paper. The output measure is a combination of DRG adjusted inpatients and daycases, and unadjusted outpatient numbers. Compared to the DEA approach, this output measure is less satisfactory. We are now combining all outputs into one output to facilitate the SFA analysis, but this means we are treating inpatients as the same as outpatients or daycases. However, our data has been adjusted for casemix, so in that sense our output measure has been adjusted for differences across inpatients and outpatients[10]. For the Sickles and Schmidt model, the SFA results are presented in Table 2 below. The average efficiency scores are now 0.63 for county and 0.60 for regional/general hospitals between 1995-2000. 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. Factors that may have contributed to this difference

---

[9] See Simar and Wilson (2000).

[10] We could attach weights to each output based on relative costs of inpatients versus outpatients versus daycases. However, we do not yet have cost data so alternatively we could estimate a multi input-output distance function – this is an important issue and will be addressed in future work on this project.

could include technological change, environmental factors such as location or ownership or simply random error within the data. The first two reasons will be explored in future work, allowing us to further analyse pure efficiency change compared to technological change.

Table 2: Mean Efficiency Scores derived from Cobb Douglas Fixed Effects Model 1995-2000

	County	Regional and General
1995-2000	0.63	0.60
N hospitals	22	13
N observations	132	78

The model from which estimates of efficiency were derived show that an increase in medical staff and number beds has a significant effect on productivity in Regional and General hospitals. The corresponding results for County hospitals were insignificant. The results from our model allow us to estimate the level of returns to scale, and as found in hospitals in Northern Ireland (McKillop et al, 1999) the larger (Regional and General) hospitals perform with increasing returns to scale but the smaller ones (County hospitals) show decreasing returns to scale.

As mentioned earlier, specification of the functional form in SFA is important. To relax the assumption of unitary elasticity of substitution within the Cobb Douglas function, we also estimated a Translog function and tested the hypothesis that the cross product coefficients in are equal to zero. We could not reject this hypothesis (see table A1 for results), and concluded that the Cobb- Douglas is a better specification of the production function under the fixed effects model (where no distribution is assumed for inefficiency effects and we assume time invariant inefficiency).

As discussed in section 3, the Schmidt and Sickles (1984) may not be appropriate if inefficiency effects are time varying. The DEA estimates of efficiency suggest that this may apply to our data, so we now present results from the Battese and Coelli (1992) model[11]. This

---

[11] We tried to estimate the Cornwell, Schmidt and Sickles (1990) model but with only 132 observations and many parameters, the model could not be solved.

differs to the model presented so far in that it specifies a distribution for the inefficiency effects and assumes time varying inefficiency. We tested several hypotheses regarding the specification of the production function (see results of the Likelihood Ratio tests in Table A1), and concluded from these tests that the Translog specification is a more appropriate specification for the county hospitals in our sample. Due to the low number of observations in regional/general sample, it is possible that the Translog model suffers from multicollinearity, and we were unable to calculate an appropriate likelihood ratio test statistic. We therefore are unable to conclude that either specification is most appropriate in the case of regional/general hospitals, so we report results from both the Translog and Cobb Douglas model. For comparison purposes, we do the same for county hospitals.

We now present results for time varying models (where the inefficiency effect is distributed as truncated normal), for the Cobb Douglas specifications. In Table 4, once again we show that there is substantial variation within both County and Regional and General hospitals. The average efficiency score of County hospitals is much the same as when we used the fixed effects model, assuming that the Translog is the best specification in both models. This result suggests there may not be time variation in inefficiency in county hospitals – it raises the question whether or not the inefficiency levels remained the same throughout 1995-2000, so we test for this in the final part of this section. Although, we could not establish an appropriate specification for the regional/general hospitals, it is interesting to see that both models give the same average efficiency score. The most important aspect of these results is that efficiency is higher overall compared to the previous time invariant model. In contrast to the county hospitals, this suggests there may be time variation of inefficiency in the regional/general hospitals. So to answer the question of whether or not inefficiency is time varying, in the final part of this section, we tested the hypothesis that $\eta = 0$ and our results are set out in Table A3.

Table 4: Mean Efficiency Scores derived from Battese and Coelli (1992) Model for hospitals 1995-2000

	County (Translog and time varying)	County (Cobb Douglas and time varying)	Regional and General (Cobb Douglas and time varying)	Regional and General (Translog and Cobb Douglas)
1995-2000	0.58	0.81	0.76	0.76
N hospitals	22	22	13	13
N observations	132	132	78	78

In Regional and General hospitals, we find that there are time varying inefficiency effects, and hospitals have become more efficient between the years 1995 and 2000, although the variation across hospitals is substantial in all years. When we restricted the model to exclude η in the County hospitals, we find a different result – we cannot reject $H_0 : \eta = 0$. These results support the results found so far, and suggest that the average efficiency over the 1995-2000 period for county hospitals is approximately 0.63 and for regional/general hospitals is about 0.76.

In this section we have demonstrated that in using the SFA models there is much more variation in efficiency within each group of hospitals, compared to the estimates of efficiency obtained using the DEA model. Overall, we find higher estimates of inefficiency within the SFA models, indicating that the DEA frontier lies below the SFA frontier. Using the SFA models, we determined whether or not efficiency has changed over time, and results from Battese and Coelli (1992) model propose that this applies in the case of regional/general hospitals only.

4.4   Discussion of Results

The results from the DEA and SFA approaches to measuring the production frontier and efficiency suggest that measured efficiency of hospitals may vary depending on whether or not a parametric approach is employed. 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. The advantage of SFA though is that we can disentangle any random error from the inefficiency effect.

The results from this analysis suggest that the types of hospitals with the lowest level of variation in inefficiency are regional and general. 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 (1998) 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 1, 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. Conclusions

The analysis of efficiency in hospitals can make a major contribution to improving health services. The ultimate aim is to (1) identify poorly performing hospitals, (2) to understand why and (3) to address the underlying causes. In this paper we concentrate on the first issue, and show how one can measure technical efficiency relative to best practice using Data Envelopment Analysis and Stochastic Frontier Analysis. 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.

Our results contribute to the international comparison of DEA and SFA in estimating technical efficiency. Our comparison between DEA and SFA methods shows that there are lower efficiency scores under the SFA method, suggesting that DEA efficiency measures are not controlling for other factors such as the type of production process or other environmental factors that not included in the model. A topic for further research is to determine how environmental factors (e.g. location or factors beyond the control of managers) may influence DEA efficiency scores.

The DEA results in section 4 cannot be used to assess trends in pure efficiency, as there may have been some technological changes over time. In future research, by using DEA models to analyse 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. Technological change can also be analysed within the SFA models, so we will apply both methods to the data to assess levels of technical efficiency change and technological change.

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

Cornwell, C., Schmidt, P. and R. Sickles. Production Frontiers with Cross-sectional and Time-series variation in Efficiency levels. Journal of Econometrics, 1990, 46:185-200.

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

Gannon, B. Technical Efficiency of Hospitals in Ireland. February 2005, Research Programme on "Health Services, Health Inequalities and Health and Social Gain", Working Paper .

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.

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.

Appendix A

Table A1   Tests for functional form in the Schmidt and Sickles Model

		F statistic	Reject H_0 ?	Preferred
County				
Cobb Douglas v Translog	$H_0:\beta_{ij}=0$	$F_{6,101} = 1.04$	No	Cobb Douglas
Regional/General Cobb Douglas v Translog	$H_0:\beta_{ij}=0$	$F_{6,101} = 1.81$	No	Cobb Douglas

Table A2   Tests for functional form in the Battese and Coelli (1992) Model

		$\chi^2(df)$	Reject H_0 ?	Preferred
County				
Cobb Douglas v Translog	$H_0:\beta_{ij}=0$	46.6 (6)	Yes	Translog

Table A3   Tests for $H_0 : \eta = 0$ in the Battese and Coelli (1992) Model

	$\chi^2(df)$	Reject H_0 ?	Preferred
Translog			
County	$\chi^2(1) = 0.22$	No	Time invariant inefficiency
Regional/General	$\chi^2(1) = 4.76$	Yes	Time varying inefficiency