# Measuring and evaluating the efficiency of NHS Trusts in England: a non-parametric and econometric analysis

Sofia Silva
University of York

With financial support from the Foundation for Science and Technology (Portugal)

# 1. Background to this research and assumptions

- According to the basic principles of a NHS, the quality of services is expected to be homogeneous across the country;

- After having adjusted for all those variables that are exogenous to the system (such as severity, age and gender casemix), a rational resource allocation should be enough to minimise undesirable variation of care;

- *Ceteris paribus*, the same cost per case should lead to the same quality of care.

**However**, significant variations in productivity and quality of care exist across England.

It is often argued by providers that they are less productive because they are providing better care. **But** is this true?

Until recently, it was not possible to analyse this question, because data for casemix-adjusted costs and indicators for clinical outcomes simply were not available.

# 2. Objectives

- To develop a league table of NHS Trusts according to their quality-productivity ratio by using Data Envelopment Analysis;

- To analyse the possible reasons for the variation of this ratio across the country by applying a censored model (tobit).

# 3. Data Envelopment Analysis

## 3.1 Unit of analysis

The unit of analysis is the English NHS Trust. In this study, we analyse 79 Trusts, for which all the necessary information is available, and of good quality.

## 3.2 Data

We use data for the period 1997-98.

### 3.2.1 Output variables

The primary source of clinical indicators is the Performance Assessment Framework.

- Rates of survival within 30 days of surgery after an emergency admission per 100,000 hospital stays;

- Rates of survival within 30 days of surgery after a non-emergency admission, per 100,000 hospital stays;

- Rates of survival in hospital within 30 days of emergency admission with a hip fracture (neck of femur), for patients aged 65 and over, per 100,000 hospital stays;

- Rates of survival in hospital within 30 days of emergency admission with a heart attack (myocardial infarction), for patients aged 50 and over, per 100,000 stays;

- Rates of non-readmission to hospital within 28 days of discharge from hospital, per 100,000 admissions;

- Rates of discharge to usual place of residence within 56 days of emergency admission with a stroke, for patients aged 50 and over. The data are expressed as percent of continuous inpatient stays of stroke patients;

- Rates of discharge to usual place of residence within 28 days of emergency admission with a hip fracture (neck of femur), for patients aged 65 and over. The data are expressed as percent of continuous inpatient stays of hip fracture patients.

### 3.2.2 Input variable

We use the Department of Health's Casemix-adjusted Unit Cost Index (CCI).

This is calculated as the ratio of actual costs (total costs for inpatients, outpatients and A&E) to casemix-weighted activity, where costs per HRG and per outpatient-specialty first attendance are used as weights.

This index is equivalent to a ratio of actual to expected costs, and is therefore a measure of productivity.

It has been standardised so that it has a national average of 1.

### 3.2.3 Model

Both the input and the outputs are presented as ratios, and so it is not possible to determine the scale at which each Trust is operating. Therefore, we must assume the presence of variable returns to scale.

We use an input-oriented model. Note that we will **not** be measuring the technical efficiency of the NHS Trusts, because our input is already, to some extent, a measure of efficiency.

The scores will indicate the extent to which the productivity of a given Trust can be increased (by reducing its CCI) without reducing its quality of care.

### 3.2.4 Results

31 NHS Trusts (39.2%) obtained scores of 1.

An additional 6 were given scores between 0.95 and 1.

The average score for the non-frontier Trusts is 0.82.

The overall average score is 0.898. This implies that the productivity of the 79 Trusts could have been increased by 11.4% (1/0.898 – 1), without affecting in any way the quality of care being provided.

Clearly, it is not acceptable that such a large proportion of the resources available for the provision of health care is being wasted, when it could have been put to good use somewhere else in the NHS.

We must now determine the reasons that lead to these variations in the quality-productivity ratio across the NHS.

# 4. Econometric analysis of the DEA scores

Many factors may affect the productivity and quality of a NHS Trust, as well as the relation between the two.

These include Trust structural and process characteristics, patient characteristics, and patient satisfaction.

Data are for the period of 1997-98, unless stated otherwise. There is a variety of sources: the Hospital Episode Statistics database, the Performance Assessment Framework, the DoH webpage, and the Sunday Times Good Hospital Guide.

## 4.1 Structural characteristics

### 4.1.1 Size

We use the number of admissions per year (variable *ADMISSIO*) and the number of admissions squared (*ADMISS2*). Data was only available for 1998-99.

We don't expect these variables to be significant, because we assumed VRS in our DEA model, which should have eliminated the effect of size in the scores.

### 4.1.2 Location

A dummy variable (*CITIES*) indicates whether the Trust is situated in one of the three major cities in Britain (London, Birmingham and Manchester).

### 4.1.3 Labour intensity and composition

We use the number of doctors per 100 beds (*DOCTORS*) and the number of nurses per 100 beds (*NURSES*), measured as whole time equivalents. The data are for 1998-99.

### 4.1.4 Teaching

The dummy variable *TEACHING* tells us whether the Trust has been classified as Acute Teaching.

### 4.1.5 Training and Research

We use the amount of income received for education, training and research (*INCOME*) as a proxi. The indicator is presented in £000 per 100 beds.

## 4.2 Process characteristics

### 4.2.1 Waiting lists

*WAITLIST* shows the number of patients in the waiting list per 100 beds, for each NHS Trust.

*MORE12MO* indicates de percentage of waiting list patients that have been waiting for more than 12 months.

*LESS13WE* shows the percentage of outpatients that have been seen within 13 weeks of referral.

*LESS30MN* indicates the percentage of outpatients that have been seen within 30 minutes of their appointment time.

### 4.2.2 Daycase rate

The casemix adjusted daycase rate (*DAYCASE*) is obtained by subtracting the actual percentage of spells that were treated as day cases in a Trust from the percentage that would have been day cases, had the national average daycase rate been achieved for each HRG. The lower the indicator, the higher the daycase rate of a Trust.

### 4.2.3 Length of stay

*AVLENGT* indicates the average length of stay per spell.

### 4.2.4 Cancellations

*NON_READ* shows the percentage of patients from the waiting list that have not been readmitted within a month after a last minute cancellation of their operation by the hospital, for non-medical reasons.

### 4.2.5 Dealing with complaints

*DEALTWIT* indicates the percentage of complaints dealt with within the 4 weeks performance target for local resolution, excluding those that require independent review.

## 4.3 Patient characteristics and patient satisfaction

### 4.3.1 Age

*MEAN_AGE* is the average age of the patients attended in each Trust.

Given that the DEA outputs have been adjusted for age, this variable should not be significant.

### 4.3.2 Complaints

*COMPLAIN* shows the number of complaints per 1000 patients. The data are for 1998-99.

### 4.3.3 Level of trust in the provider

*TRUST* is the percentage of survivors of heart operations that say they trust their doctor.

# 5. Applying a censored regression model to the DEA scores

A large percentage of the Trusts in our sample have their scores clustered at 1, which is the highest score DEA allows.

In cases like this, where there is a non-negligible proportion of observations reaching the upper value limit (and where we have, therefore, a limited dependent variable), OLS cannot be applied; the expected errors will not equal zero, and so standard regression will provide a biased estimate.

There seems to be substantial literature with distributions similar to the DEA scores where a tobit model has been applied.

For computational reasons, it is desirable that a tobit model assumes a censoring point at zero. Therefore, we must transform the DEA scores with the formula:

Tobit score = 1/DEA score – 1

From now on, our dependent variable will be this score.

The tobit regression model takes the form:

$y_i = ß'x_i + u_i$, if $y_i^* > 0$
$y_i = 0$, otherwise

where $y_i$ is the observed variable, transformed from the original one, $y_i^*$. In other words, we will have

$y_i = y_i^*$, if $y_i^* > 0$
$y_i = 0$, otherwise

## 5.1 Results

- The size variables are not statistically significant, as we expected;

- *CITIES* is highly significant. It seems that Trusts outside the three main cities tend to achieve a better quality-productivity ratio, contrary to what is often defended in the literature. This is probably due to the level of deprivation that exists in big cities, which has negative effects on the effectiveness of care;

- *DOCTORS* is also significant, showing a negative association with the tobit scores, leading us to believe that, either there is a lack of doctors in the English Trusts, or the ones that are working there do not produce as much good quality care as they could have;

- *INCOME* is significant, even if only at the 10% level. This is probably due to the fact that research and training activities require time and resources from the Trust's staff.

- *MORE12MO* has a negative association with the tobit scores; the longer people wait to be treated, the higher that Trust's quality-productivity will be. Perhaps by making people waiting longer, the hospital manages to treat the right number of patients each period, according to its capacity, which leads to better and more efficient care;

- *DAYCASE* shows a positive association with the tobit scores, as expected. Note that, the higher this indicator, the lower the day case rate of the Trust. By treating more patients in ambulatory, it is possible to increase the turn-over of beds and other resources (both physical and human).

- *NON_READ* shows a positive association with the tobit scores, as would have been expected from a variable that can be said to represent pure mismanagement of a Trust.

- *COMPLAIN* shows a positive association with the tobit scores, as expected.

- Not surprisingly, *DEALTWIT* shows the opposite association with the score. Even though sorting out the complaints involves a cost, it allows members of staff to get to know which aspects of care were not satisfactory to patients, and therefore improve their performance. The patient also feels he has been heard, and patient satisfaction often has a positive influence on his health.

The remaining variables are not statistically significant.

In order to test for misspecification in the model, we conducted a RESET test. There was no evidence of misspecification.

# 6. Weaknesses of this analysis

The sample only includes 79 of the 389 NHS Trusts listed in 1997-98, as all the necessary information was not available for the remaining. Assuming that these 79 are the ones with better management, which allowed them to provide complete and accurate data, we may expect that, by including the remaining Trusts, the variations in the performance would have been ever larger.

The data for some of the explanatory variables used in the tobit model did not refer to the period for which the DEA was applied. However, it did refer to periods immediately before or after, and we do not have reasons to believe that important changes have occurred.

Finally, the act that some variabels were not significant does not necessarily inficate that the aspects of care for which they are proxies are not relevant.