Performance measurement in public and regulated organisations: evaluation and improvement of productive efficiency using data envelopment analysis.

Thesis submitted for the degree of
Doctor of Philosophy
at the University of Leicester

by

Michael Doble BA (CNAA) MA (Leicester)
Public Sector Economics Research Centre

December 1991

UMI Number: U041198

All rights reserved

INFORMATION TO ALL USERS
The quality of this reproduction is dependent upon the quality of the copy submitted.

In the unlikely event that the author did not send a complete manuscript
and there are missing pages, these will be noted. Also, if material had to be removed,
a note will indicate the deletion.

UMI U041198
Published by ProQuest LLC 2015. Copyright in the Dissertation held by the Author.
Microform Edition © ProQuest LLC.
All rights reserved. This work is protected against
unauthorized copying under Title 17, United States Code.

ProQuest LLC
789 East Eisenhower Parkway
P.O. Box 1346
Ann Arbor, MI 48106-1346

UNIVERSITY LIBRARY LEICESTER

7501076009

Acknowledgements.

A work such as this could not have been completed without the input and support of a large number of people. I would like to acknowledge those who have helped to make this thesis possible.

Special thanks are due to my partner Carrie who has always understood the difficult conditions under which I was working. It has not been easy for either of us as for a period of time we were apart, Carrie's study for a PGCE took her to a college some distance away. I would also like to thank my parents who are very proud that I have completed and realise what an achievement it is. They have always been there to support me if I have needed help.

Academically there are a number of people who have given of their interest and time. I am grateful to my supervisor Professor Peter Jackson for at the outset believing in my ability to conduct the Ph.D. research and his advice and encouragement during the frustrations of a long period of study. He has also usefully provided me with employment opportunities to sustain myself financially.

I would also like to acknowledge the contribution of Professor John Cubbin who provided me with some initial aid in my study of Data Envelopment Analysis and also generously gave me access to his computer software.

I am obliged to Mike Moir of the Department of Operational Research and Statistics at Post Office Counters Headquarters for giving me access to the dataset used in this thesis. Adrian Gaines and Mike Hellier also gave freely of their time in helping me to understand the data and much useful background in the workings of Counters.

I would also like to thank Tom Weyman-Jones of Loughborough University for working with me on the study of the Area Electricity Boards and providing many useful suggestions on improving this thesis. Also, his enthusiasm and help gave me the impetus I needed to complete my writing up.

There are innumerable colleagues who have been supportive and provided me with much useful discussion over the years. I would also like to thank the friendly and co-operative staff at Leicester University Computer Centre, whose efficient service has made my work much easier.

The diagrams and typing in this thesis were produced entirely by myself and I therefore accept full responsibility for any errors.

Abstract.

Data envelopment analysis (DEA) represents a new and innovative way of measuring performance in the public sector. First suggested by Farrel (1957), linear programming techniques have made it possible to measure technical efficiency through the estimation of non-parametric production frontiers. DEA measures technical efficiency, it does not need information on prices or costs and produces a single efficiency criterion using data purely on measured volumes of inputs and outputs, including qualitative ones. The type of DEA programme used in this thesis is one suggested by Banker (1984) allowing for variable returns to scale in the construction of the production frontier.

It is argued that DEA is superior to other attempts to measure the performance of public sector organisations. It has been traditional to use partial productivity ratios to measure efficiency and these have recently become popular in the public sector as 'performance indicators'. Other techniques used are regression analysis and the estimation of econometric production frontiers. It is demonstrated that DEA has a number of advantages over these methods.

This is not to say that DEA does not have its drawbacks or that there are not practical problems with its use. These have been explored through the conduct of two case studies, one of Post Office Counters and the other of the Area Electricity Boards. Themes explored during the course of the case studies are as follows. How useful is DEA as a performance measurement tool in the context of Post Office Counters? Can more information on efficiency be yielded by clustering the datasets into smaller groups which have a common factor and re-estimating the frontier. How robust is the DEA isoquant given that it is a frontier constituted on the basis of maximum observations and is susceptible to outliers, and can this be overcome? Also undertaken in this thesis is one of the first British studies using DEA in a dynamic context.

The case study of Counters was conducted using data that was derived from 1281 Crown Post Offices for a 13 week period from September-November 1989. The input used was labour (in hours) and the outputs were a quality variable (average waiting time) and the outputs were ten different categories of transactions. The results were examined for the whole of Counters and on a regional basis. In the study of the Area Electricity Boards a time-series approach was taken on a pooled data set for 12 AEBs from 1969-88. Three inputs and four outputs were used. Technical efficiency between AEBs is compared, the change in efficiency over time is examined and conclusions reached about the implications for regulation now they have been privatised.

Performance measurement in public and regulated organisations: evaluation and improvement of productive efficiency using data envelopment analysis.

CONTENTS.

Chapter 3.
Conducting a study of Post Office Counters technical efficiency.

Chapter 4.
The Use of DEA as a Performance Measure to Aid Management Control and its application to Post Office Counters.

Figures.

Tables.

Introduction.

The present government's commitment to improving the performance of the public sector has resulted in the continuing search for better ways of performance evaluation. Such performance can only be measured in terms of the objectives of an organisation and there are usually a range of these. This means that financial measures such as profitability which are popularly used to evaluate performance in the private sector cannot be used. Given that there can be trade-offs between objectives, the public sector organisations objective that is least in conflict with the others is that of technical efficiency and this will be the criterion used in this study. In addition, due to the absence of competition, market prices and costs may be lacking for some or all of the inputs and outputs. A measure of efficiency that concentrates on volumes of inputs and outputs is therefore desirable.

Popular techniques used to measure efficiency in the public sector have been regression analysis and performance indicators (PIs). The drawbacks to these methods are that regression analysis only measures efficiency in relation to the average efficiency attained rather than the maximum obtainable and PIs are inadequate because of the partial way in which they measure.

Dissatisfaction with these measures has resulted in the development of a large number of new methods for measuring efficiency which have in common the concept of the frontier. A production frontier is a technological relationship which describes the maximum output that an efficient firm can produce from any given combination of inputs. That is, efficient organisations are those operating on the production frontier, whilst inefficient organisations are those operating below it (this type of efficiency is technical efficiency). This approach is superior to measuring efficiency with an average production function because we need to relate actual output to the maximum output that is potentially achievable. It is better than ratios because it can accommodate many factors and produces a single result.

The foremost of these methods for estimating frontiers is a non-parametric linear-programming technique called Data Envelopment Analysis (DEA). The aim of this thesis is to evaluate DEA and examine its shortcomings and apply it to two organisations, a public sector enterprise and a regulated industry.

Chapter 1 surveys the old and new approaches to efficiency measurement, how they have been applied to the British public sector and their drawbacks. The rationale for choosing technical efficiency as the criterion for performance is explained. Then different concepts of efficiency are examined and the aspect of efficiency that we are interested in, that measured by DEA, which is technical efficiency, is put into context. DEA and its main features are then explained graphically.

Chapter 2 explains the linear-programming techniques used to estimate the DEA frontier. Chapters 1 and 2 constitute a survey of the literature on performance measurement and production frontier-models in that Chapter 1 examines methods of performance measurement, the DEA frontier model and other forms of frontier estimation and Chapter 2 looks specifically at approaches to DEA frontier estimation.

Chapter 3 introduces the background to the first case study of Post Office Counters and how it should be conducted. The structure and management of Counters is explained. Golany and Roll's systematic application procedure for DEA is set out and this is borne in mind whilst defining the information needed to conduct the analysis. To assist in this, studies of similar financial institutions are examined and data that Counters is currently producing.

Chapter 4 looks at performance measurement in the management context, that is, the ability of DEA to be used in a control system. As the purpose of a control system is to improve performance it is explained how DEA can do this. This is compared with the way DEA has begun to be used to improve performance at Counters in practice.

Chapter 5 specifies the final model for the Post Office Counters case study. The overall technical efficiency is examined and how performance can be improved by identifying inefficient DMUs and their peer groups and establishing resource targets to enable them to move to the frontier. The degree of scale efficiency is also established. To yield more information on the reasons for technical inefficiency the offices are also examined on a regional basis. Clustering can be used to glean more information by providing similar groups for comparison and establishing more accurate efficiency discrimination between units. This is done by creating groups with similar characteristics (in this case regions) and then re-estimating separate frontiers. The effect of clustering on efficiency ratings, peer groups and targets is examined and compared to the non-clustered regional results.

Chapter 6 examines a problem with the non-statistical deterministic approach of DEA is that it does not take into account 'statistical noise' (things like measurement error and outliers). This chapter looks at ways of identifying and overcoming this problem.

Chapter 7 applies a new way of using DEA, that of pooling the observations from different time periods, rather than doing a cross-section, to measure technical efficiency over time. The analysis is conducted using Area Electricity Board data from 1969-88 and the techniques success ascertained. It is also contrasted with another dynamic DEA method, which is 'window analysis'. Conclusions are drawn about the efficiency of the Area Electricity Boards over the last twenty years and the factors that have affected their performance. The possibility is examined of using DEA as a regulatory tool now that the industry has been privatised.

Chapter 1.
Measurement of performance in the public sector.

1.1 Introduction.

1.1.1 How should public sector performance be measured?

The purpose of this section is to show how technical efficiency measurement should be the starting point in evaluating the performance of public sector organisations. There are two major considerations for this being the case in the public sector and these will be explained in turn.

Public sector organisations must participate in the implementation of the governments objectives in the same way as other components of the public sector. Consequently, their performance is to be evaluated by the extent by which they do in fact achieve these objectives. A private firm's operations could also be appraised according to the same criterion. But, its primary objective is to maximise rates of return on capital to pay shareholders. In the public sector context there is no analogue for profit-seeking behaviour or adequate feed-back system for learning about the quality of decisions. As a result of this the problem of evaluating performance of public sector organisations and the development of insights to guide performance improvement has been much more difficult.

Measuring the performance of public firms thus amounts to appraising the extent to which they in fact achieve the goals which are given to them. This is known as the 'Performance approach' as expounded by Tulkens (1986). He says that evaluating performance involves proceeding through four stages: a. stating the objectives attributed to public organisations, b. providing a justification for them, c. translating the objectives in terms of indicators that are observable and measurable and d. deriving methods by which comparison can be made between the observed values of the indicators and those values that correspond to a complete fulfilment of the corresponding objective.

There are a range of conceivable economic objectives traditionally assigned to publicly owned firms (Rees [1984], Bos [1986]). These objectives, which are more generally those of public policy can be summarised under four different headings: ***efficiency, equity, financial balance*** *and* ***macroeconomic***.

Efficiency is related to the way resources are used and goods are allocated. It can be divided into two subcomponents reflecting the physical efficiency of the input-output production transformation (productive or technical efficiency) and the price efficiency of optimal resource allocation (allocative efficiency). It can be argued that this distinction is somehow artificial as production decisions are joint, but it is usually assumed that the effects of these decisions can be measured separately empirically.

The three other objectives will now be explained briefly. Equity implies concern for the effects of employment and pricing policies on the distribution of real income. Financial balance could be included in macroeconomic objectives, it is often cited separately because a number of public enterprises have expensive ongoing deficits, and it is a major issue whether these are due to mismanagement or simply to unavoidable huge fixed costs. Finally, the government addresses its macroeconomic objectives with all available instruments, including public production, to solve problems like unemployment, inflation, the trade balance and growth of the economy.

Having reached some agreement on the list of possible objectives, the next step involves developing for each of them a specific indicator to establish how close they come to achieving them. A performance measure would then be obtained from a multi-criterion combination of these indicators. Unfortunately such a programme is unrealistic as a whole simply because some parts of it do raise logical and methodological difficulties that are unresolved so far (an example of this approach is provided by Marchand, Pestieau and Tulkens [1984,a]). All that can be provided is a set of indicators which essentially deal with the stated allocative objectives, the other objectives achievement being measured in an ad hoc way and then used to qualify the results.

The above objectives can rarely be fulfilled together without conflict. As they conflict there is a need to prioritise them, choosing efficiency on the one hand and equity, financial balance and stabilisation on the other. This implies that a lower achievement of one objective can be justified by concern for the other objectives. For instance, inefficient operations could be justified by the search for redistribution or full employment. The only objective that does not conflict with any other and so is compatible with all of the others is that of technical efficiency. In other words, producing little or employing too many factors, compared to what is technically feasible, cannot be justified in the name of any objective. As Parris, Pestieau and Saynor (1986) state, 'Technical efficiency is a very attractive concept of performance. A firm, private or public, which is charged with technical efficiency, cannot invoke non-allocative objectives to justify itself. A performing enterprise should always be technically efficient; there is no trade-off between technical efficiency and any other objectives assigned by the state to the firm'.

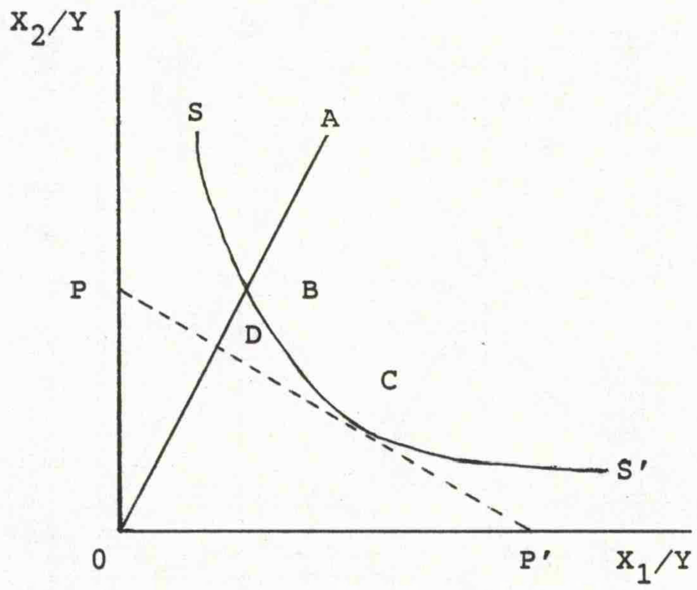
Another major consideration is the problem of measurement in the public sector. Only information about the volume of inputs and outputs is needed to calculate technical efficiency. In the absence of competition market prices are not available to compute, aggregate and compare the economic value of outputs and costs may also be lacking for inputs. Physical data is often more readily available and more reliable, so it can be concluded that 'measuring technical efficiency can be viewed as the first and unavoidable stage in the agenda of a performance study' (Pestieau [1987]). Public sector performance will therefore be concentrated on from the sole point of view of technical efficiency as it is the least contentious criterion. The concept of efficiency now needs to be examined in more detail and this will be done in the rest of section 1. Section 2 examines the commonly used performance measures in the public sector and section 3 surveys methods of estimation of frontier production functions including data envelopment analysis. The concluding section will show that DEA fulfils the criteria for a good performance measure and overcomes most of the difficulties that other measures have.

1.1.2 The concepts of allocative and technical efficiency.

The concept of the efficiency of a productive unit can be decomposed into three requirements. The first requirement of efficiency is that the maximum possible amount is produced with the resources used, or to put it another way, it must be impossible to reduce the volume of any input without reducing the volume of output. This is what is known as ***technical efficiency.*** The second requirement is that the cost of any given level of output is minimised by combining inputs in such a manner that one input cannot be substituted for another without raising the total cost. This is known as ***allocative efficiency.*** The third requirement is that the mix of outputs of different goods and services produced from the given resources maximises the benefits to consumers. This means that it is impossible to produce more of one good at the expense of another without reducing the total value of output to consumers (***pareto efficiency***). In the simple theoretical world of perfect competition, efficiency is achieved by competition amongst cost-minimising firms for the custom of rational, well-informed consumers who seek to achieve the greatest value for their money. In the production of government services this is not the case and so the aim is to develop management and information systems which come as close as possible to maximising efficiency in the above sense.

The efficiency concepts given can be illustrated as follows. Consider for the sake of simplicity a firm employing two factors of production X_1 and X_2, producing output Y and that the firms production function is $y=f(X_1,X_2)$. Assuming that there are constant returns to scale (or linear homogeneity) so that it may be written $1=f(X_1/Y,X_2/Y)$ all the relevant information can be characterised by the ***unit isoquant*** SS′ (see Figure 1.1.1).

Figure 1.1.1
Technical and allocative efficiency.



The isoquant SS′ represents the various combinations of the two factors that a perfectly efficient firm might use to produce unit output and it represents the frontier technology.

Now the point B represents an efficient firm using the two factors in the same ratio as the firm A. B produces the same amount of output as A using only a fraction OB/OA as much of each factor. The measure of technical efficiency of A then is OB/OA, the ratio of input combinations that is actually achieved to the ratio of input combinations that can be achieved. The technical efficiency measure of a perfect firm would be OB/OB=1. Suppose now that the line PP′

shows the available budget, all points along PP' have the same cost. This is a cost function so that C = c(p,Y) where C = total cost and p is a vector of input prices. Its slope reflects the relative prices of X1 and X2. Although all points along SS' are technically efficient, only C is allocatively efficient. C represents the minimum cost combination of inputs. Point B costs too much because the units of output could be produced at a cost of 0D by substituting X_1 for X_2. The allocative efficiency (sometimes called price efficiency) score for A is 0D/0B, as B moves closer to C the score rises. Total efficiency is the product of allocative and technical efficiency, 0D/0A.

The practical estimation of allocative efficiency is difficult in the public sector context. It requires data on the prices all inputs and it needs to take into account the possible movements over time in both prices and the responses of producers to such changes. It is not relevant to this study and is not examined further.

1.2 Commonly used public sector efficiency measurement techniques.

1.2.1 Performance indicators.

The most common measure of public sector efficiency are the hopefully named ***performance indicators*** (PIs). They are expressed as ***ratios*** and as will be seen should be thought of as ***partial productivity measures***. It will be explained precisely what they are, where and how in the public sector they are being used. It will then be discussed how good a measure of performance they are.

It is necessary to differentiate between performance measures and indicators. The distinction exists because the precision with which things can be measured is different. If efficiency can be measured precisely and unambiguously then it is usual to refer to performance measures. If it is not possible to obtain a precise measure it is usual to refer to performance indicators. PIs therefore give an indication rather than a clear measure of performance.

The type of PIs we are interested in, in this context, are those that are indicators of technical efficiency. That is, measures that relate the volume of inputs to produce a given output or outputs. This is expressed as a ratio, dividing inputs (the numerator) by outputs (the denominator). It is assumed that other factors of production and technology do not change in the period studied.

Performance indicators became very popular in the 1980s, they were embraced at the highest levels of government and promulgated throughout the public sector. There was a thrust towards managerial reform throughout government, central to which was the wider use of PIs. The shift of interest from the traditional focus on inputs to outputs was partly because the government needed to find out what departments were actually doing and partly because of the overriding concern of the Thatcher Administration to control public expenditure. Hence the attraction of a system that emphasized outputs rather then defining all improvements in terms of inputs. The government was also anxious to improve managerial competence. The ascendency of the 'three Es', economy, efficiency and effectiveness, was intended to increase central government control over service delivery as much as it was to cut costs.

The build-up of interest resulted in the Financial Management Initiative (FMI) announced in the 1982 White Paper on 'Efficiency and Effectiveness in the Civil Service' (Cmnd 8616). It emphasized that managers at all levels in government should have a 'clear view of their objectives, and assess, and wherever possible measure, outputs or performance in relation to these objectives', (Prime Minister and Chancellor of the Exchequer, 1983). The most obvious manifestation of the interest in PIs was their proliferation in the Public Expenditure White Paper, multiplying from 500 in 1985, to over 2,300 in 1989.

The quest to develop PIs has maintained momentum in different sections of government. Within local government, the establishment of the Audit Commission as an entity in 1982 has institutionalised the process of seeking improved value for money through quantitative assessments of performance. They produce evaluative reports on different function of local government. The approach is usually a comparative one, conclusions drawn about individual cases are based on measures of relative efficiency. Their prescriptive approach raises awareness and encourages measurement.

In central government the development of PIs has been given a boost by the implementation of the 'Next Steps' initiative to break up three-quarters of the civil service into separate executive agencies. By the mid-1991 50 agencies had been formed with around 200,000 civil servants working in them, this is about half the civil service. Underpinning the agency approach is the use of performance measures as instruments of hands-off managerial control and democratic accountability to Parliament and the public. Each agency is required to publish performance targets in its framework agreement which will be the subject of a quarterly review by its sponsoring department and the scrutiny of parliamentary committees and individual politicians. To this end, the Treasury is co-ordinating 'extensive work on assisting parent departments and agencies to develop a suitable 'portfolio' of output and performance measures'.

In the 1980s most nationalised industries developed an array of PIs in response to pressure from the government (Treasury [1978]) and numerous critical reports from parliamentary committees and the Monopolies and Mergers Commission. It seems likely that PIs will play an increasing role in the regulation of monopolies, whether public or private. It will be encouraged because of the changing climate of growing concern about consumer satisfaction and quality of service. As well as the nationalised industries it would be thought desirable that the various regulating authorities would need to monitor the activities of privatised monopolies.

The PI approach will now be evaluated. Its appeal is that it relies on simple mathematical concepts and that it can be useful in locating extremely good or extremely poor operating relationships. The aspects of a units operations that are out of line with the norm are easily identified and can be easily assimilated and used by managers. But, there are a number of drawbacks with this approach.

The main problem is the partial picture that individual PIs give of overall performance. A single partial productivity measure can be misleading and when dealing with several products and factors, the ratio of one particular component of output to a particular component of input even less meaningful. The multi-input multi-output nature of public organisations means that a range of PIs are presented as a morass of numbers that give no clear indication of true efficiency. Different ratios may give conflicting signals about efficiency. Even given a consistent non-conflicting set of indicators, it is not possible to implement a data reduction technology for aggregating the multiple PIs. Relative weights can be attached to the ratios in an index-number approach. But, this is problematic as in a great many situations outcomes will be politically and publicly sensitive, and therefore disputed, or they may just be undecided or unrevealed. As there is no objective way of assigning relative weights to the ratios it is difficult to conclude which units are inefficient.

A criticism of the way in which PIs have been developed to date is that they focus on inputs and processes, and say little about outputs and objectives. This is a particular problem in the NHS for example where, their PI package contains a lot of detail on resource usage as on throughputs, but contains no information on patient outcomes. This problem probably says more about the public sector than it does about PIs. The relationship between inputs and outcomes is not always obvious and processes are easier to measure. It also reflects the fact that the objectives of public sector organisations are frequently not well defined. Not only is causation hard to follow but the final output may be qualitative in nature.

Another problem with the PI approach is that there is no way of taking into account the effects of endogenous or exogenous factors which affect the efficiency comparison. Exogenous factors can be ***systematic*** or ***random***. For instance, the length of stay in hospital of people in an inner city district may be systematically affected by the socio-economic background or age profile of the local population. This makes it difficult to compare different hospitals. An extreme value may occur because of a random event eg. a local epidemic, bad weather etc. There will be a random element in differences in indicators anyway so some indication of statistically significant differences is required, to avoid waste of management effort in pursuing explanations of chance variation.

The only way possible efficiency improvements can be quantified using PIs is by relating the value of a ratio to the mean across the sample. But, random variation makes it difficult to quantify possible efficiency improvements in terms of reduction in inputs or increases in outputs. In addition due to the fact that there are a range of PIs, different DMUs may have good PIs in some areas and not in others, no reference peer group can be identified as models to improve efficiency generally.

Care must be taken in the construction and selection of ratios. 'Pseudo ratios', the result of dividing items that are not logically related must be avoided. They may be mathematically related to real ratios, but do not measure any underlying reality. For instance, the ratio of hospital porters to hip operations is a pseudo ratio because the inputs and outputs are not directly related.

1.2.2 Regression analysis.

Regression analysis can be used to model the output level of an organisation as a function of the various input levels. The result of this kind of estimation takes the form of an equation.

1.2.1. $y = a + bx + u$

For each observation, y is the variable to be explained and x is the variable explaining it. The u term is the unexplained variation in y. The values of a and b are parameters chosen to minimise the variation of u, that is, to explain as much as possible of the variation of y in terms of x. The method can incorporate more inputs (multiple regression analysis) if x is replaced by a whole set of measured variables each having its own coefficient, b, to be estimated.

Regression analysis can be used to identify technically inefficient DMUs. Application of, say, ordinary least squares to the above equation will give estimates of b, from which one can compute the residual. The residual is the difference between the output that the model predicts for a DMU and its actual output. A DMU with a residual equal to zero (that is, a DMU lying on the regression line) is said to be of average efficiency. A DMU with a positive residual (a DMU lying above the regression line) is said to be of above average efficiency and vice-versa. Thus the residual shows how technically efficient a DMU is because it indicates whether it is providing more or less than the average on the basis of its input usage and the estimated parameters of the production function.

Several studies have used this approach to estimate technical efficiency in public sector organisations. The earliest study of this type was by Feldstein (1967) to examine the (technical) efficiency of 177 large acute non-teaching hospitals in the NHS. Levitt and Joyce (1987) used basically the same approach to analyse police authoritys. They estimated a simultaneous equation incorporating *inter alia*, an equation relating an authorities clear-up rate (an output) to various input variables, including the number of police officers *per caput* the resources per police officer, the crime-mix in the area and various socioeconomic variables. Authorities that were more than one standard deviation above (below) the regression line were judged to be particularly efficient (inefficient). The authors also compared the ranking based on the regression model estimates with the ranking implied by the standard police performance indicator, the crime clear-up rate. Interestingly, the comparison showed very little correspondence at the extremes of the distribution. For example, of the authorities ranked according to the clear-up rate only two were amongst the most efficient authorities. The same approach has been used to analyse the efficiency of local education authorities (LEAs) by Levitt and Joyce (1987) and the Department of Education and Science (DES [1983 and 1984]). Output is measured by examination results. For example, Levitt and Joyce measure output by:

a. the percentage of pupils with five or more higher grades
b. the percentage of pupils with two or fewer of any grade
c. a composite measure of exam performance generated through a primal components analysis of six indicators of exam results.

Input variables include proxies for socioeconomic status and teaching resources available. Residuals of the production function are then used to identify efficient and inefficient LEAs. Running regressions using the three alternative output measures produced three efficiency rankings, there was a degree of agreement between them with some exceptions.

There are a number of shortcomings of using regression analysis as a measure of technical efficiency. The technique does not allow for the use of multiple outputs. Single -equation regression analysis requires that there be only one output or that all outputs be combined into a single indicator of production. Multiple-equation regression models can be used, but then there are multiple sets of residuals and no clear way of interpreting them in terms of efficiency.

Another drawback of regression analysis is its use of a fixed parametric form. The parametric specification of a production function is the equation specifying how inputs are combined to produce outputs. The appropriate mathematical form may be unknown and if measuring technical efficiency, incorrect results produced if an inappropriate functional form is imposed.

The most important criticism of regression analysis as a measure of technical efficiency is that *average* production functions are estimated. The results of a comparison between the actual and expected (average) outputs do not necessarily indicate good or bad performance. The regression model might be misspecified, measurement errors might be present, chance alone can account for at least part of the residual. The number of standard deviations from the regression line could be used as a measure of the level of technical efficiency but due to the above factors the size of the residual cannot really be used as an indicator of robustness. Rather than relating actual output to the average output we need to relate it to the maximum that is potentially available.

This is the idea behind frontier models of efficiency. The starting point for a discussion of such models is the work of Farrell (1957), who first proposed the ratio 0B/0A in Figure 1.2.1. as the measure of the technical efficiency of a firm at point A.

1.3 Production frontiers for efficiency measurement.

1.3.1 Estimating production frontiers.

The standard definition of a production function is that it is 'the maximum flow of output per unit of time achievable for any given rates of flow of input services per unit of time' (Johnson [1960]).[1] However, a production function calculated using regression analysis will have positive residuals as well as a negative ones, which is paradoxical in the light of the definition given above. This occurs because a regression estimates the *mean* rather than the *maximal* output given. To answer certain economic questions, for instance, the efficiency level, the maximum output is relevant. A substantial literature has arisen which attempts to estimate frontier production functions. A variety of approaches has been used. The production frontier can be ***parametric or non-parametric***, depending upon whether a particular functional form has been imposed on the data. ***Statistical or non-statistical***, statistical assumptions can be made about the disturbance term u. Or the frontier can be ***deterministic*** *or* ***stochastic***, depending upon whether variations from the frontier are attributable purely to technical efficiency or technical efficiency and random variation. Surveys of these production frontier models are provided by Førsund, Lovell and Schmidt (1980) and Schmidt (1986).

The frontier technique that we are interested in is DEA (which is deterministic, non-parametric and non-statistical), the method will be explained, as well as the basis for measurement. Also, for the sake of completeness, a brief overview of other frontier production models will first be given.

1.3.2 Parametric production frontier models: deterministic and stochastic frontier models.

The parametric frontier models can be divided up into deterministic and stochastic models and these will be looked at in turn.

Farrell's original paper also suggested a parametric approach to the estimation of the frontier using a

Cobb-Douglas functional form. This idea was taken up subsequently in a paper by Aigner and Chu (1968). The basic model may be written in a simplified form as

$$1.3.1 \quad y = \ln f(x) - u, \quad u \geq 0$$

$$= \alpha_0 + \sum_{i=1}^{n} \alpha_i \ln x_i - u_1, \quad u \geq 0$$

where y is output, x_i are inputs, u is a disturbance term, and α_i are parameters to be estimated (x_i and y are in logarithmic form). On the assumption that all variation from the frontier in 1.3.1 is due to technical inefficiency, the residuals are constrained to be one-sided and thus output is bounded from above, that is $y \leq f(x)$. This kind of frontier is deterministic in the sense that it is assumed that all DMUs operate on or beneath the frontier, but not above it. They all face the same frontier and no allowance is made for exogenous shocks outside the control of the firm which may affect performance, nor for the effects of measurement error or omitted variables. Estimation of the parameters of frontiers of 1.3.1. is possible using linear programming techniques and the residuals can be used to generate a measure of technical efficiency for each of the sample observations.

The advantage of this kind of parametric approach is that it enables the frontier technology to be described in a relatively simple functional form, reducing the amount of computational complexity involved (the Cobb-Douglas functional form does not have to be used, more flexible forms such as the CES or translog could be used). The disadvantage of this approach is the structure imposed on the data may not be warranted. Also, as in the case of Farrells non-parametric approach, there is the same sensitivity to outliers and measurement errors/misspecification, and in the absence of any statistical assumptions about the residual means that statistical tests cannot be conducted.

It is an obvious next step to make some statistical assumptions about u and this has often been the approach adopted in the empirical literature. It is assumed that observations on u are independently and identically distributed (iid), with mean u and finite variance and that u is uncorrelated with the inputs. Estimation of the frontier can usually be by corrected ordinary least squares (COLS), this involves 'correcting' the estimated intercept by shifting it upwards until no residual is positive and one is zero (see Richmond [1974]). Or the frontier can be calculated using maximum likelihood (ML). The likelihood function can be derived and maximum likelihood estimators (MLEs) can be calculated if a specific one-sided distribution for the disturbance term is assumed (eg. exponential or gamma). This was first done by Afriat (1972).

The problem is that the ML estimates will be sensitive to which distributional assumptions are made and theoretical grounds provide very little guidance for preferring any particular assumption, though this problem could be confronted by examining the effects of the possible alternative assumptions about u. Another problem pointed out by Schmidt (1986) is that the range of the dependent variable will depend upon the parameters to be estimated. This violates one of the conditions for ML, that the MLEs should be consistent and asymptotically efficient, thus making their statistical properties uncertain. But Greene (1980) finds sufficient, though fairly restrictive, conditions on the distribution of u for the MLEs to have their usual desirable asymptotic properties.

In a stochastic frontier model, output is assumed to be bounded from above by a stochastic frontier. Deterministic models assume that all variations in performance from the frontier are attributed to inefficiency. In contrast, the stochastic frontier (also called the composed error model), introduced by Aigner, Lovell and Schmidt (1977) and Meeusen and van den Broeck (1977) get around this problem by decomposing the error term into two parts. The model can be written thus

1.3.2 $y = f(x) + v - u$

where $u \geq 0$ so that all points lie on or below the stochastic frontier. One symmetric component of the error, v, is intended to capture random effects of statistical noise factors outside the control of the individual firm. The second one-sided component, u, is intended to capture the effects of inefficiency relative to the stochastic frontier.

Estimation of the frontier is possible using either COLS or ML (see Greene [1980], [1982]). The asymptotic properties of ML now hold because of the symmetric error component. Whichever method is used one ends up with a residual for each DMU, an estimate of the mean of u, but not an estimate of u. What one can estimate however is E(u/v+u), the expected value of u, given the value of the composite error (Jondrow, Lovell, Materov and Schmidt [1982]).

A criticism of the stochastic frontier model approach by Schmidt (1986), is that in order to be able to distinguish the two error components, it is also necessary to make some strong distributional assumptions. In most cases the symmetric error, u, has been assumed to be iid normal, but a variety of assumptions have been made about the distribution of technical efficiency. Studies show that different distributional assumptions can lead to different results in terms of estimated efficiencies. This problem is not overcome by assessing the frontier to be deterministic, as Schmidt points out 'assuming statistical noise not to exist is itself a strong distributional assumption'.

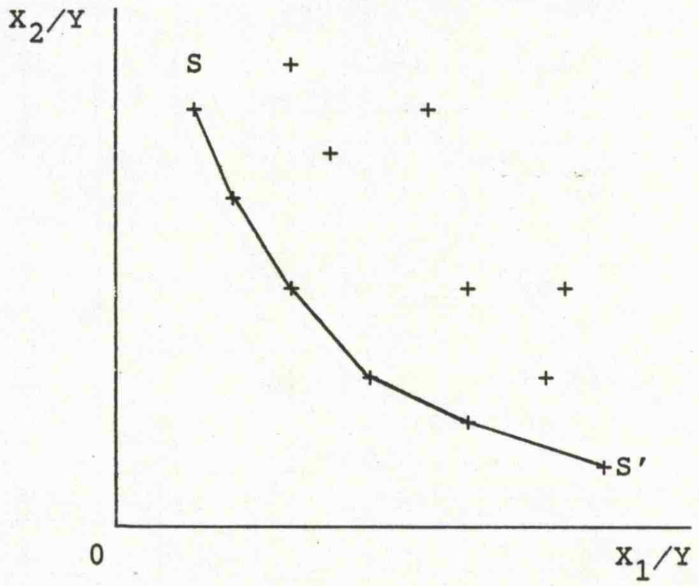
The main advantage of the parametric statistical approach is that it is statistical. It is the only approach that makes any accommodation for noise, measurement error and exogenous shocks beyond the control of the production unit. Without such an accommodation these phenomena are construed as inefficiency. Its main disadvantage is that it is parametric, so a possibly unwarranted structure is imposed on the production function and on the distribution of efficiency as well.

1.3.3.1 Non-parametric production frontier models: Data envelopment analysis.

DEA is a non-parametric approach to estimating the production frontier. It is non-statistical as non-parametric approaches usually are (with the exception of Banker and Maindiratta [1985]). Farrell (1957) first suggested the DEA approach as a method of measuring technical efficiency.

On the same set of assumptions as in section 1.1.2, Farrell describes the set of firms in a given industry by plotting them according to inputs per unit of output for each of the various inputs, producing a scatter diagram. To obtain a standard for measuring the efficiency of the firm under consideration, Farrell fitted a frontier function to the points as a piecewise linear function (see Figure 1.3.1). The line segments linking all the efficient input bundles trace out the efficient isoquant and 'envelops' all the inefficient firms.[2]

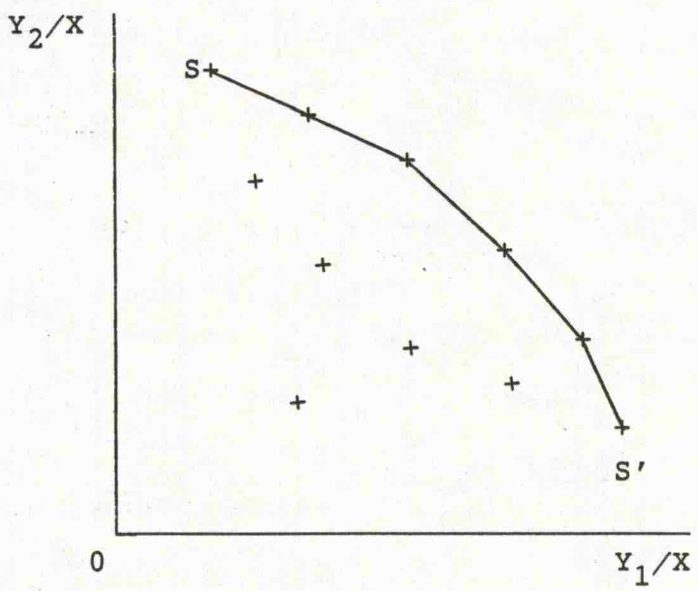
Figure 1.3.1
Efficiency measurement for more than one input.



He assumed that the isoquant would be convex to the origin (***convexity*** means that if two points are attainable in practise then so is any point representing a weighted average of them) and has nowhere a positive slope, the curve SS' being taken as the estimate of the efficient isoquant. Farrell calls this curve the 'efficient production function'. Technical efficiency is measured by the ratio of the observed performance of a firm to a hypothetical firm using the factors in the same proportions, measured in the direction of the origin. This hypothetical firm is constructed as a weighted average of those of the observed firms. This type of frontier function is called deterministic, all the observations must lie on or below the frontier, all deviations from the efficient unit isoquant arise solely due to technical differences in efficiency with no allowance for statistical noise of any form.

The approach can be extended to tackle multiple outputs (see Figure 1.3.2). The DMUs are plotted according to outputs per unit of input for the case of one input and two outputs.

Figure 1.3.2
Efficiency measurement for more than one output.



With multiple outputs the measurement of inefficiency proceeds along much the same lines as before. Inefficient output bundles are separated from efficient bundles by joining adjacent pairs of bundles with a line segment. If the line segment has a non-positive slope and none of the other bundles lies to the North-East of it, the chosen bundles are declared efficient. If they are below this they are inefficient. As before, efficiency is measured as a ratio of the distance from the origin to the DMU being evaluated to the distance from the origin to the frontier.

The frontier is computed using linear programming techniques (these are discussed in Chapter 2). The output of these calculations includes, in addition to an estimate of the technical efficiency of each organisation, the value of the weights used to construct the hypothetical DMUs, as well as indicators of the amount of slack usage in any of the inputs and outputs.

What has just been done is to compare the observed performance of a firm with some postulated standard of perfect efficiency. This is the concept engineers use when they discuss the efficiency of a machine or process. However, although it is perhaps the best concept to use for the efficiency of a single production process it is very difficult to specify a theoretical efficient function for something as complex as the production process of a typical firm or industry. Even the best engineer is likely to overlook some problems and the more complex the process the less accurate is the theoretical function likely to be. Farrell realised this and while acknowledging the validity of defining a standard of absolute efficiency felt it would be more useful to define efficiency in terms of the observed standard. DEA is therefore a ***relative efficiency*** measure. The Farrell frontier is known as ***best-practice*** because it reflects the achievements of the best DMUs in the sample. Each DMU will have a ***peer group,*** a peer group can form the frontier for a number of inefficient DMUs and their efficiency is measured relative to the peers.

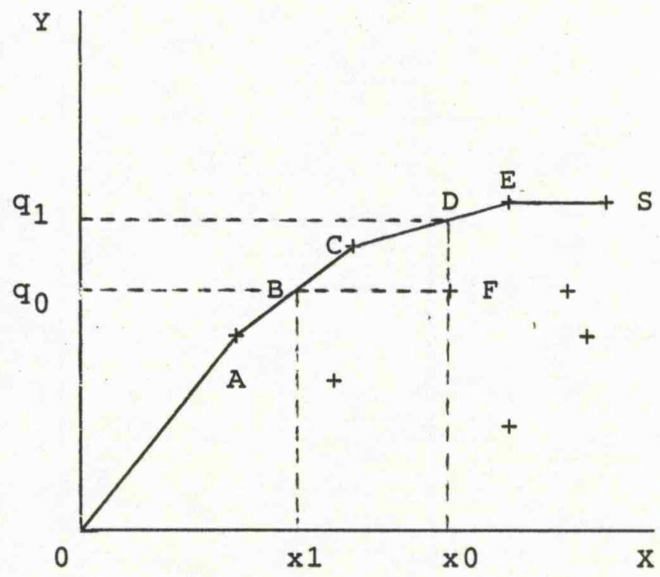
An alternative name for a peer group used in economic literature is a **reference set.** The idea of relative performance evaluation through the use of appropriate comparison or 'reference' sets is not new. The concept of a reference group was introduced by Hyman (1957), and applied to goal formulation and goal attainment by March and Simon (1958) and Cyert and March (1963). In general, when the desirable performance criteria are ambiguous or when cause and effect relationships cannot be specified with precision, organisations utilise reference groups in goal setting and performance evaluation.

1.3.3.2 Input minimisation or output maximisation.

It should be said that technical efficiency can be measured in two ways. Where perfect technical efficiency exists, it is impossible to reduce any input without reducing at least one output or to increase any output without increasing at least one input. The relative efficiency of a management unit can be measured in terms of an efficiency score. This can be defined as either the ratio of its actual output, or where there is more than one, some weighted average of them to its expected output (given its inputs and after allowing for factors outside its control), this is ***output efficiency***. Or alternatively, the ratio of its expected input or, where there is more than one, some weighted combination of them (given its outputs and circumstances beyond its control), to its actual input, this is ***input efficiency***. The expected output or input would refer to the maximum possible output, given input, or the minimum possible input, given output, predicted from an observation of other management units. The efficiency scores can then be used to order management units in an efficiency ranking.

The fact that the reference set and thus the technical efficiency of each DMU depends on whether output maximisation or input minimisation is assumed can be illustrated as follows. Figure 1.3.3 gives an example of DEA

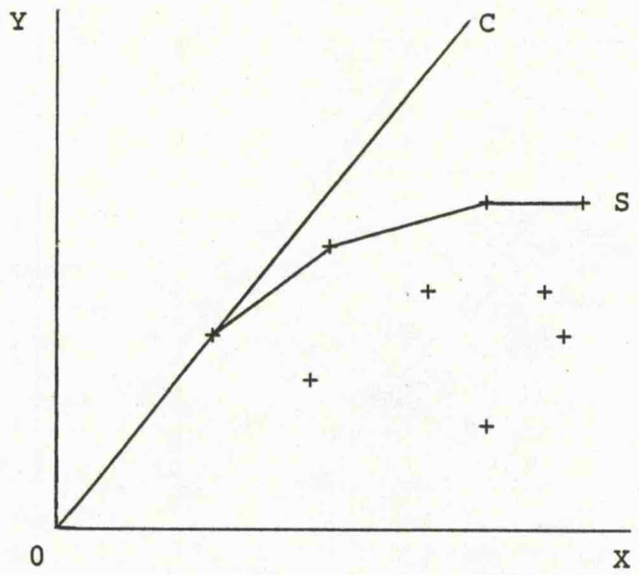
Figure 1.3.3
Input and output efficiency measurement.



for the simple case of one input and one output. The efficient frontier is constructed from a sample of observations by constructing a piece-wise linear locus from those DMUs which envelop the other points, here 0S. Convexity is assumed and so each DMU is measured by comparison with a hypothetical DMU formed as a weighted average of a number of efficient DMUs. If it is assumed that DMUs attempt to maximise output given their input, then F would be compared with point D, a linear combination of C and E. If DMUs are thought to minimise input given output, F would be compared with B, a linear combination of A and C. The technical efficiency rating of F would be $0q_0/0q_1$ under output maximisation and $0x_1/0x_0$ under input minimisation.

The efficiency score of a management unit and its efficiency ranking will in general be different under each assumption. Input and output efficiency scores will only be the same if there are constant returns to scale in the activity being considered. In this case, where a given increase in input leads to an equiproportionate increase in output, the true production frontier would be a straight line through the origin. However, if the constant returns to scale assumption is not appropriate then the results can be very misleading. In the case in Figure 1.3.4. where there

Figure 1.3.4
The effect of imposing constant returns to scale.



appears to be decreasing returns to scale, the imposition of constant returns would lead to the frontier 0C, revealing all but one of the DMUs as inefficient.

Which measurement assumption is used would depend very much on the objectives of an organisation. In the public sector, given the imposed constraints, in general it would be thought the input minimisation assumption would be more appropriate.

1.4 Conclusion.

Previous attempts to utilise comparative approaches to evaluate performance and measure technical efficiency have not been entirely satisfactory. Both the PI approach and regression analysis have their drawbacks. Their shortcomings in general derive from the complexity of public sector organisations, the multiplicity of inputs and outputs and their qualitative nature.

The examination of PIs highlighted the need for a performance measure that could be capable of

1. deriving a single aggregate measure of relative efficiency for a set of DMUs in terms of their utilisation of input factors to produce outputs.

2. being able to handle non-commensurate, multiple inputs and multiple outputs.

3. not being dependent on a set of *a priori* weights or prices for the inputs or outputs.

4. being equitable and defensible

5. handling qualitative as well as quantitative factors

6. being able to adjust for factors outside the control of the units being evaluated

7. being able to provide insights on the possibilities for increasing outputs and/or decreasing inputs for the inefficient unit to become efficient

Regression analysis can achieve all of these things except 2., multiple outputs cannot be accommodated. It satisfies 1. to the extent that it aggregates inputs and can say what output should be and how many standard deviations from the regression line actual output is. But, it is an average not a frontier model of the production function and due to statistical noise cannot really tell us much about the level of technical efficiency. In addition, regression analysis imposes a parametric form on the data which may not be warranted. Frontier production models whilst improving on regression analysis also suffer from this problem.

DEA fulfils all of the above criteria for a good performance measure. It is non-parametric and so does not use all of the available information, the results are therefore less precise but avoid the danger of distorting the evidence by imposing the wrong parametric form.[3]

That is not to say that DEA is free from problems and these are as follows.

1. As DEA measures relative rather than absolute efficiency it does not take into account ***structural efficiency***. DEA measures the extent by which the organisation keeps up with the performance of its own best DMUs. If they themselves are not efficient then the organisation is structurally inefficient. This has implications for assessing the level of efficiency of relatively inefficient DMUs. The DMUs deemed to be best-practice need to be examined by observation to decide whether they are absolutely efficient.

2. DEA as with any other performance measure assumes all inputs and outputs to be specified and measured. But it also has to assume that the direction of causation is known, which is not always known a *priori*. DEA is entirely mathematical, it simply calculates the ratio of specified outputs to specified inputs (or vice-versa). It does not test whether or not there is any statistically significant relationship between the inputs and outputs, it assumes they are causally related. In practise an assumed relationship can be difficult to validate, if at all.

3. Since the frontier is constructed from a sub-set of the data (the most efficient management units) it is very vulnerable to ostensibly efficient outliers that could be extreme observations or measurement errors. DEA searches for the best possible weights for a DMU, if it is superior to any other in just one dimension then it will be deemed to be efficient. Whilst DEA is susceptible to outliers it is no worse than other forms of frontier estimation as all frontier methods will use an outlier to construct part of the frontier. This has greater implications than with regression analysis because an outlier does not affect the regression line so significantly. The regression line is an average of the observations and so an outlier will have a broad but small quantitative impact.

Whilst these problems exist they do not negate the superiority of DEA over other measures. They will be borne in mind and account taken of them or discussed directly in the course of this thesis. Structural efficiency is dealt with as a problem in an applied context in Chapter 5. Methods of model validation are examined in Chapter 3 but, in the two case studies that take place in this thesis, as will be seen, the models are not complex. This means that causation is reasonably apparent and in any case, variables are often defined by the available data. The problem of outliers is examined in Chapter 6.

Footnotes.

1. The same interpretation applies to the cost function. Duality theory establishes the relationship between production and costs. For given factor prices the cost function can be interpreted as a frontier function because it is impossible to achieve costs lower than the maximum requirements implied by the production frontier. Whilst frontier cost functions can be estimated these will not be discussed in this study as we are concerned with technical rather than economic efficiency.

2. The term that will be used to denote an organisational entity is a decision-making unit (DMU). This definition is used in order to emphasize that interest is centred on decision-making by public sector organisations. In this context data is not readily weighted by referring to market prices. This is in preference to the terms, firms or plants, concerned with input and output decisions in the economics literature (eg. Coase [1937]). Borrowed and adapted from this literature are ideas that will be used, such as the production function and related concepts such as duality. This is not to say that the use of these concepts are wholly perfect deriving as they do from the neo-classical theory of the firm. Williamson (1985) regards the business firm as a governance structure rather than a production function.

3. On a theoretical basis DEA would appear to be the most satisfactory performance measure. Some authors have attempted to determine this empirically. DEA is compared with ratio analysis by Sherman (1981), with regression analysis by Banker, Conrad and Strauss (1986) and with both by Bowlin, Charnes, Cooper and Sherman (1985) and Levitt and Joyce (1987). In general the results were favourable to DEA. It is not proposed that a comparison of performance measurement techniques will be conducted here. DEA has advantages over other measures and in any case a comparative study using Counters data has already been conducted by Kirthisingha, Shutler and Land (1987).

Chapter 2.
The estimation of DEA models using linear programming.

2.1. Introduction.

In Chapter 1 the DEA technique was illustrated graphically for the case where there are only two inputs or outputs. But, when there are more than two quantities involved, as there always are in such applications, this cannot be done. They can only be displayed graphically in two dimensions and so another method is necessary to compute the DEA results. This can be achieved using linear programming techniques and these are described in this chapter.

The basic model is set out as a fractional program and this is described in section 2.2.1. Whilst in its fractional form the program can be thought of as the conceptual DEA model, the linear program is that used in the actual computation of the efficiency ratio. This program is called the primal it is described in section 2.2.2, its mathematical counterpart is the dual program and this is explained in section 2.2.3. Section 2.3 relaxes the restrictive constant returns to scale assumption of the basic DEA program and allows other types of technologies to be estimated.

2.2 The DEA program.

2.2.1 The fractional program.

The efficiency of the ith decision-making unit is defined as follows.

2.2.1.

$$\text{Efficiency} = \frac{\sum_{r=1}^{r=t} u_r \, y_{rj}}{\sum_{i=1}^{i=m} v_i \, x_{ij}}$$

y_{rj} = the amount of the rth output of the jth DMU
u_r = the weight given to the rth output

x_{ij} = the amount of the ith input for the jth DMU
v_i = the weight given to the ith input

Efficiency is thus defined as a ratio of the weighted sums of the outputs to the weighted sums of the inputs.

The relative efficiency of a set of decision-making units is obtained by treating the weights u_r and v_i as variables and maximising the efficiency of the unit subject to the efficiencies of all the units being constrained to be less than 1.

Thus to determine the efficiency of unit j, 2.2.2. MAX.

$$\frac{\sum_{r=1}^{r=t} u_r \, y_{rj}}{\sum_{i=1}^{i=m} v_i \, x_{ij}}$$

S.T.

$$\frac{\sum_{i=1}^{r=t} u_r \, y_{rj}}{\sum_{i=1}^{i=m} v_i \, x_{ij}} \leq 1 \quad j = 1 \ldots . j_0 \ldots n$$

The selected weights cannot be negative

$u_r \geq 0;\ r=1, \ldots\ldots, t$
$v_i \geq 0;\ i=1, \ldots\ldots, m$

2.2.2 The primal program.

The former program is a fractional linear program which in its original formulation is both non-linear and non-convex. Charnes and Cooper (1962,1973) have advocated the use of a transformation to convert the fractional program into an ordinary linear program. By solving a series of linear programming optimisations, one for each DMU, DEA is able to identify those DMUs that are efficient, and the remaining inefficient DMUs along with their efficient reference points. The program can be solved by one of two linear programming formulations, these are derived from Ganley (1989) as they relate to the Cubbin software.

The first formulation constrains the weighted sum of the inputs to be unity and maximises the outputs that can then be obtained. For each DMU solve

$$2.2.3 \quad \text{MAX.} \quad hk = \sum_{r=1}^{r=t} u_r \, y_{rk}$$

$$\text{S.T.} \quad \sum_{r=1}^{r=t} u_r \, y_{rj} \leq \sum_{i=1}^{i=m} v_i \, x_{ij}, \quad j=1,....k....,Z$$

$$\sum_{i=1}^{i=m} v_i \, x_{ik} = 1$$

$u_r \geq 0$; $r=1,.....,t$.
$v_i \geq 0$; $i=1,.....,m$.

The second formulation constrains the sum of the weighted output at unity, and minimises the inputs needed. Its formulation is as follows

$$2.2.4 \quad \text{MIN.} \quad h_k = \sum_{i=1}^{i=m} v_i \, x_{ik}$$

$$\text{S.T.} \quad \sum_{i=1}^{i=m} v_i \, x_{ij} \geq \sum_{i=1}^{i=m} u_r \, y_{rj} \quad j=1....,k....,Z$$

$$\sum_{r=1}^{r=t} u_r \, y_{rk} = 1$$

$u_r \geq 0$; $r=1,.....,t$
$v_i \geq 0$; $i=1,.....,m$

A strict positivity requirement on the weights was introduced by Charnes, Cooper and Rhodes (1979) as a correction to their first model with non-negative weights in (1978). Charnes, Cooper and Rhodes (1979), restricted the input and output weights such that

2.2.5 $$u_r \geq \varepsilon, \; r=1,\ldots\ldots,t \qquad \varepsilon>0$$
$$v_i \geq \varepsilon, \; i=1,\ldots\ldots,m \qquad \varepsilon>0$$

Where ε is an infinitesimal constant eg. 10^{-6}. This 'lower bound constraint' (Lewin and Morey [1981]) was introduced into the primal because in the situation where there are no lower bounds, under certain circumstances unity efficiency ratings could be implied in the fractional program for branches with non-zero slack variables. It could then be the case that further improvements in efficiency are possible (see Boyd and Fare [1984],and Charnes and Cooper [1984]).

2.2.2 The dual program.

In the theory of linear programming every linear program has a companion linear program which is called its dual. The dual problem has its set of variables and set of constraints and its own objective form expressed in terms of those variables. When the simplex method is applied to a linear program, not only is its optimal solution obtained but also the optimal solution to the dual problem. Thus, with no additional computation, whenever the primal is solved, so also is the dual, that is, the values of the dual variables are obtained that optimise the dual objective while satisfying the dual constraints.

Computation of the efficiency score is achieved using the duals of the primals 2.2.3 and 2.2.4. The dual constructs a piecewise linear approximation to the true frontier by minimising the quantities of the m inputs required to meet stated levels of the t outputs. That is

2.2.6 MIN. $$h_k = \Theta_k - \varepsilon \cdot \left(\sum_{i=1}^{i=m} S_i + \sum_{r=1}^{r=t} S_r \right)$$

$$\text{S.T.} \quad x_{ik} \cdot \Theta_k - S_i = \sum_{j=1}^{Z} x_{ij} \lambda_j \qquad i=1,\ldots.,m$$

$$y_{rk} + S_r = \sum_{j=1}^{Z} y_{rj} \lambda_j \qquad r=1,\ldots.,t$$

and $\lambda j \geq 0$, $j=1,\ldots.,k,\ldots.,Z$ (weights on DMUs)
$Si \geq 0$, $i=1,\ldots.,m$ (input slacks)
$Sr \geq 0$, $r=1,\ldots.,t$ (output slacks)
(Θ_k is indeterminate in sign)

The objective function of the models attempts to find a minimal value for an 'intensity' factor Θk which indicates the potential of a proportional reduction in all the inputs of DMUk. In addition, it seeks the largest slack values in all input-output dimensions. Thus, it finds the reference point on the empirical production frontier which portrays DMUk in the efficiency characterisation that gives it the lowest technical efficiency rating. The constraints of the model represent the envelopment principle explained earlier.

In the dual problem, the kth branch is relatively efficient if, and only if, the efficiency ratio Θk^* equals unity and the slack variables are all zero. That is if,

2.2.7 $$\Theta k^* = 1 \text{ with } S_i^* = S_r^* = 0 \text{ for all i and r,}$$

the asterisk denotes optimal values of the variables in the dual program. The branch in question must be operating at the end point of a negatively sloped facet of the frontier so the conditions in 2.2.7 are conditions which define best-practice.

It should be noted that the shadow price interpretation of the choice variables is confined to the primal since the dual calculates weights (λj) on branches rather than on inputs and outputs. Also the dual weights are non-negative.

The dual of the output maximisation program will now the given for the sake of completeness.

$$2.2.8 \quad \text{MAX.} \quad \beta_k = \theta_k - \varepsilon \cdot \left(\sum_{i=1}^{i=m} S_i + \sum_{r=1}^{r=t} S_r \right)$$

$$\text{S.T.} \quad \beta_k \cdot y_{rk} + S_r = \sum_{j=1}^{Z} \lambda_j \, y_{rj} \quad r=1,\ldots\ldots,t$$

$$x_{ik} - S_i = \sum_{j=1}^{Z} \lambda_j \, x_{ij} \quad i=1,\ldots\ldots,m$$

and $\lambda_j \geq 0, \quad j=1,\ldots\ldots,k,\ldots\ldots,Z$
$S_i \geq 0, \quad i=1,\ldots\ldots,m$
$S_r \geq 0, \quad r=1,\ldots\ldots,t$

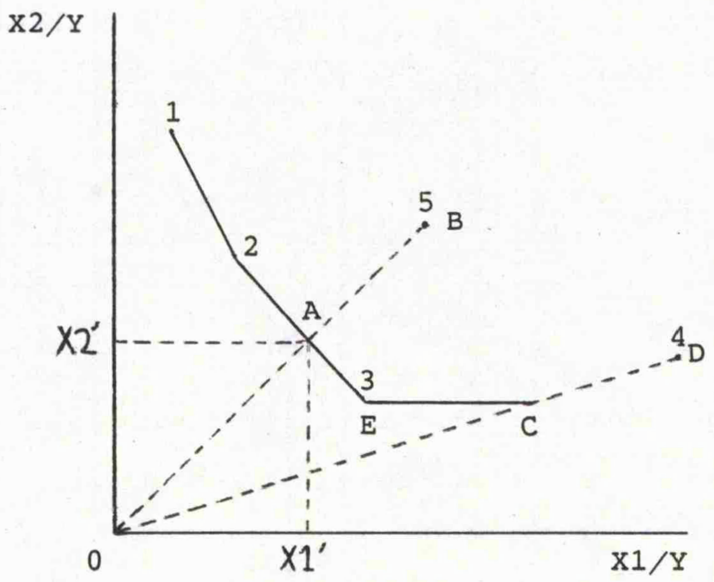
As before, the dual is the program used in the computation of the DEA efficiency ratio, but this time it determines the output efficiency of a branch k with inputs given.

The dual program 2.2.6 will be the model that is used to calculate all of the results in this thesis. There will be some adjustments to the basic model to take into account returns to scale and these will be explained in section 2.5.

The reasons for choosing the input minimisation form of the model will be explained in Chapter 5. It is useful at this stage to give a graphical illustration of the dual. This will facilitate an explanation of how inefficient DMUs can achieve best-practise.

A DMU is defined as being technically efficient in its use of inputs if no other DMU or linear combination of DMUs is producing equal amounts of outputs for less of at least one input. This definition equates with the formal efficiency conditions given in 2.4.2., that is, a DMU k is efficient if the efficiency ratio is unity and the slack variables are zero. Figure 2.4.1. illustrates a

Figure 2.2.1
The dual Technology.



hypothetical frontier technology with 5 DMUs producing a single output, Y, from 2 inputs X1 and X2. DMUs 1, 2 and 3 are best-practise, they have unity efficiency ratios and zero slacks in the solution to the dual. Taking as an example, the solution of the dual for DMU 2

$\theta 2^* = 1$ (* indicating optimal values of the variables)

the constraint being

Input 1. $x_{12} \cdot \theta_2^* - 0 = x_{12} \cdot \lambda_2^*$
Input 2. $x_{22} \cdot \theta_2^* - 0 = x_{22} \cdot \lambda_2^*$

and for output

$$y_{12} + 0 = y_{12}$$

The left-hand side of the constraints defines the possible reduction in inputs to enable a DMU to achieve best-practice. In this case there is none because actual performance on the right-hand side of the constraints is equal to the left-hand side. Best-practice implies $\lambda2*$, the peer group drops out of the RHS of the constraints and for an efficient DMU its peer group is itself because, $\lambda_2^* = 1$ and λ_j^* and $\lambda_j^* = 0$, $j \neq 2$.

The quantity of reduction in inputs or increase in outputs to enable a DMU to achieve best-practice is called the ***target***. This concept can be illustrated by examining the inefficient DMUs 4 and 5. They are inefficient relative to frontier performance, that is, for the same level of output, it is possible to find a DMU or a linear combination of DMUs, which are using at least one of the inputs.

DMU 5, has an efficiency ratio 0A/0B which is less than unity. This is because a linear combination of DMUs 2 and 3 is producing at least as much output as 5 with less of X1 and X2. The efficiency ratio can be used to define a target for DMU 5 so that it can become efficient and move to a position on the frontier, that is,

$$(0A/0B) \cdot 0B = 0A$$

The target vector 0A is that which needs to be aimed for to achieve best-practice. DMU 5 can be efficient and maintaining its current output by reducing inputs to X1' and

X2'. It has been suggested that targets can be achieved by reference to peer performance (in this case DMUs 2 and 3). How these reductions are actually to be achieved in practise is discussed in Chapter 4.

The peers are those DMUs that have non-zero weights in the optimal solution in the dual. For DMU 5 the solution is

$$\theta 5^* = 0A/0B < 1$$

the constraints are

Input 1 $\quad x_{15} \cdot \theta_5^* - 0 = x_{12} \cdot \lambda_2^* + x_{13} \cdot \lambda_3^*$
Input 2 $\quad x_{25} \cdot \theta_5^* - 0 = x_{22} \cdot \lambda_2^* + x_{23} \cdot \lambda_3^*$

and on output

$$y_{15} + 0 = y_{12} \cdot \lambda_3^* + y_{13} \cdot \lambda_3^*$$

Target performance for 5, x_i, 5 . θ_5^*, i = 1,2, is equal to a linear combination of performance at DMUs 2 and 3 where λ_2^*, $\lambda_3^* > 0$ and the weights on the other branches are all zero: $\lambda_j^* = 0$, $j \neq 2,3$.

Note that there are constraints on inputs and outputs in the dual. The input constraints define a proportionate decrease in inputs given by the efficiency rating θ_k^* with additional reductions given by non-zero input slack variables, S_r^*, r = 1,....,t. In the input minimisation dual, the output constraints do not only include a proportionate adjustment and are only of importance if any of the optimal output slacks S_i^* are non-zero. DMU 5's solution has all input and output slacks equal to zero, but DMU 4 has a non-zero slack on input X1. The efficiency ratio for DMU 4 is 0C/0D which is a proportionate decrease in both inputs, however, at point C DMU 3 is producing the same output for less of X1 and the same amount of X2. Thus, DMU 4 is not fully efficient until it reduces input X1 by the horizontal distance C to E. This distance is given by a non-zero slack S_1^* in the solution of the dual for DMU 4,

$$\theta_4^* = 0C/0D$$

the input constraints are

Input 1 $\quad x_{14} \cdot \theta_4^* - S1^* = x_{23} \cdot \lambda_3^*$
Input 2 $\quad x_{24} \cdot \theta_4^* - 0 = x_{23} \cdot \lambda_3^*$

and on outputs

$$y_{24} + 0 = y_{13} \cdot \lambda_3^*$$

The target for DMU 4 is a contraction in both inputs given by θ_4^* (except that only X1 is reduced) and the additional reduction in X1, given by S_1^* DMU 4's peer group is branch 3 alone since its target coincides with its performance. Therefore, $\lambda_3^* = 1$ and $\lambda j = 0$ for $j \neq 3$.

2.3 Models for analysing returns to scale using DEA.

Farrell's notion of efficiency has been measured using two basic empirical approaches. Firstly, the non-parametric approach using linear programming techniques to calculate productive efficiency. Secondly, a non-parametric approach, either deterministic or stochastic, where efficiency is measured relative to a frontier which is estimated statistically. In contrast, economists have chosen to use the parametric approach to measuring efficiency (in order to overcome Farrell's assumption of constant returns to scale [CRS]). Generalisation of the Farrell approach using parametric methods has taken place recently by Schmidt and Lovell (1979), Kopp (1981) and Kopp and Diewert (1982). Grosskopf (1986) states that improvements in this method have decreased the validity of the non-parametric approach among economists because of the restrictive technologies that were employed in early studies. In support of this Forsund, Lovell and Schmidt (1980) say that

'While his [Farrell's] measures are valid for the restrictive technologies he considered they do not generalise easily to technologies that are not linearly homogenous, or to techniques in which strong disposability and strict quasiconcavity are inappropriate'.

More recent work has attempted to make the original DEA program more general and incorporate different types of technologies. Suggestions for programs allowing for different types of scale and disposability have been produced by Fare, Grosskopf and Lovell ([1983],[1985],[1986]), Banker (1984) and Banker, Charnes and Cooper (1984).

Fare Grosskopf and Lovell have argued that by relaxing the assumptions of constant returns to scale and strong disposability of inputs, one could decompose the overall technical efficiency (TE) into pure technical efficiency (PTE) and scale efficiency (SE). This would of course allow a more detailed picture of the ways in which productive efficiency can be increased. Strong disposability means that non-boundary production is possible. In the context of

Figure 2.3.1.
Weak and strong disposability.

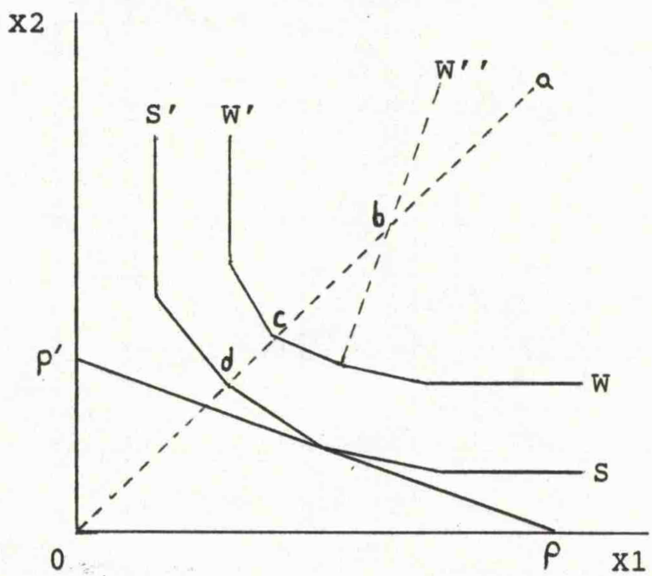


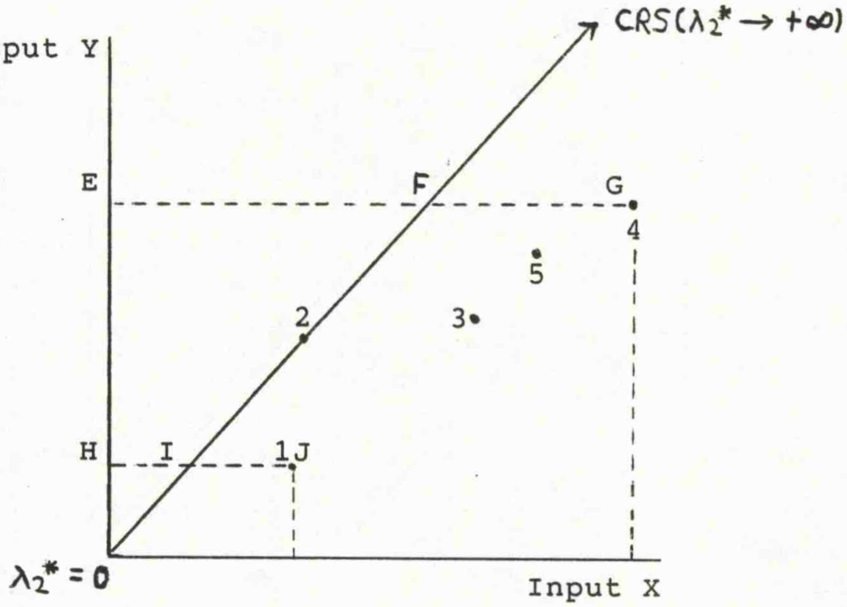
Figure 2.3.1 strong diposability means that if X2 increases holding X1 constant, output will not decrease. Weak disposability means that if X1 and X2 increase proportionately, output will not decrease. These assumptions are illustrated in Figure 2.3.1. by the isoquants WW' and WW'' respectively. If the true technology is weakly disposable (WW'') whilst it is assumed that it is strongly disposable (WW'), then firm b will appear technically inefficient (because it is in the interior of isoquant WW') even though no greater output could have been produced with the given inputs. It is not proposed that disposability will be discussed in any greater detail, (but see Fare, Grosskopf and Lovell [1987] and Fare and Grosskopf [1983]). It was decided that it would be preferable to use a program that exhibits strong rather than weak disposability and as Fare, Grosskopf and Lovell program assumes weak dispoability, it is not used. Banker's variable returns to scale program assumes strong disposability and so the discussion will follow his work.

Before the possibilities for other types of scale technologies are examined it is necessary to look at the construction of the constant returns to scale frontier in more detail. The frontier constructed by Farrell (1957) and Charnes, Cooper and Rhodes (1978) and in the dual program 2.2.6. is defined by identifying the DMU which maximises the ratio of output to input. An unbounded ray is drawn starting at the origin and passing through this DMU which is a point that denotes maximum average productivity. This is displayed in Figure 2.3.2. Examining the solution to the CRS dual corresponding to the ray OCRS in Figure 2.3.2 the solution to the dual 2.2.6 for DMU 2 (suppressing subscripts on inputs and outputs) would be

$$\theta 2^{*} = 1$$

$$x_2 \theta \, 2_{*} = x_2 \, \lambda_2^{*}$$

Figure 2.3.2.
Constant returns to scale.



and $y_2 = y_2 \lambda_2^*$

where $\lambda 2* = 1$ and $\lambda 2* = 0, j \neq 2$

The remaining DMUs have lower average productivity ratios but as they have a higher input-output mix than DMU 2 they do not appear to face a frontier. To calculate the input-efficiency ratios of these DMUs the performance of DMU 2 has to be extrapolated in the appropriate direction using an assumption of 'Ray Unboundedness' (Banker, Charnes and Cooper [1984]). This generates the ray 0CRS in Figure 2.3.2. Computationally it is constructed by varying the weights on the scale efficient branch in the solution to the dual.

For instance, consider the dual solution for DMU 4 which is consistent with Figure 2.3.2.

$$\theta_4^* = EF/EG$$
$$x_4 \theta_{4*} = x_2 \lambda_2^*$$

where $\lambda_2^* > 1$ and $\lambda_j^* = 0, j \neq 2$

That is, the target vector for branch 4, $(x_2\lambda_2^*, y_2\lambda_2^*)$, is a re-scaling of performance at the dominant branch by the factor λ_2^*.

Considering the solution for DMU 1 which has lower inputs and outputs than the scale efficient branch:

$$\theta_1^* = HI/HJ < 1$$
$$x_1\ \theta_1^* = x_2\ \lambda_2^*$$

where $\lambda_2^* < 1$ and $\lambda_j^* = 0,\ j \neq 2$

The target vector for DMU 1 is a re-scaling of performance at the dominant DMU, but for input-output levels lower than scale efficient levels the optimal weight λ_2^* is less than unity. Thus it is apparent that by varying the value of the weights on the scale efficient DMUs (that is, λ_2^* in figure 2.3.2) it is possible to construct a frontier consistent with a constant returns to scale technology. It should be noted that at the origin $\lambda_2^* = 0$ and for higher levels of inputs and outputs $\lambda_2^* \rightarrow +\infty$

Banker (1984) points out that the type of scale can be identified from the CRS dual. This is done via the weights, with lower inputs and outputs than the reference DMU a subject DMU will have a target which is a scaling up of best-practise performance and vice-versa. So that 'in the case when a unique supporting hyperplane passes through an efficient point'

$\lambda_b^* < 1 \Rightarrow$ IRS (Increasing returns to scale)
$\lambda_b^* = 1 \Rightarrow$ CRS (Constant returns to scale)
$\lambda_b^* > 1 \Rightarrow$ DRS (Decreasing returns to scale)

Where are multiple inputs and outputs several DMUs may be scale efficient on at least one variable, so that the scale indicator would be the sum of the optimal weights on each of the DMUs

$$\sum_{j=1}^{Z} \lambda_j < 1 \Rightarrow \text{IRS/DRS}$$

$$\sum_{j=1}^{Z} \lambda_j = 1 \Rightarrow \text{CRS}$$

$$\sum_{j=1}^{Z} \lambda_j > 1 \Rightarrow \text{DRS}$$

where some of the $\lambda_j^* = 0$ for inefficient DMUs.

It has been shown that the position of the frontier is embodied in the constraints from the dual program. An unbounded CRS ray can be generated by an unlimited selection of values of the weights λ_j^*. Thus, if the program restricts the value of the weights this will affect the shape and position of the frontier. This is the idea behind Banker's variable returns to scale program. The addition of a constant

$$\sum \lambda_j^* = 1$$

will exclude the constraint of the unbounded CRS ray because the unlimited vector extension of scale efficiency performance is no longer possible.

Banker's full variable returns to scale (VRS) program is used in this thesis, it assumes input minimisation and locally increasing, constant and decreasing returns to scale. It is,

$$2.3.1 \quad \text{MIN.} \quad h_k = \theta_k - \varepsilon \cdot \left(\sum_{i=1}^{i=m} s_i + \sum_{r=1}^{r=t} s_r \right)$$

$$\text{S.T.} \quad x_{ik} \cdot \Theta_k - S_i = \sum_{j=1}^{Z} x_{ij} \lambda_j \qquad i=1,\ldots.,m$$

$$y_{rk} + S_r = \sum_{j=1}^{Z} y_{rj} \lambda_j \qquad r=1,\ldots.,t$$

$$\sum \lambda_j^{*} = 1$$

and $\lambda j \geq 0$, $j=1,\ldots.,k,\ldots.,Z$ (weights on DMUs)
$Si \geq 0$, $i=1,\ldots.,m$ (input slacks)
$Sr \geq 0$, $r=1,\ldots.,t$ (output slacks)

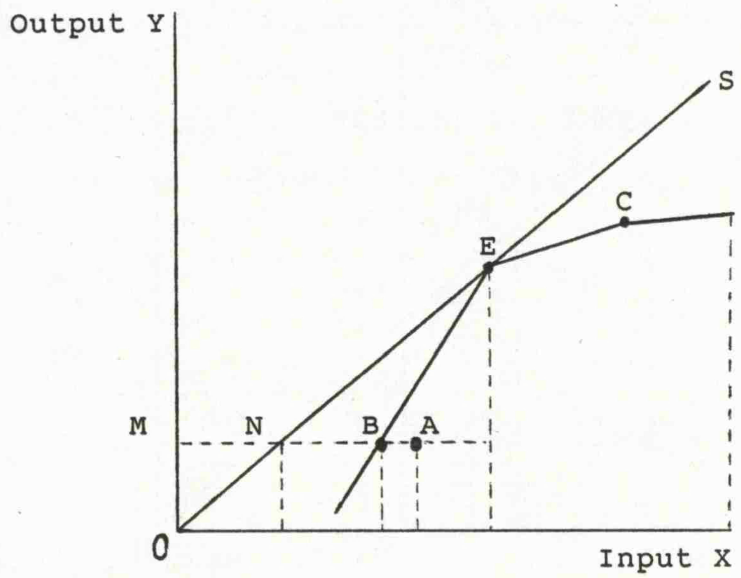
Which is an identical program to 2.2.6 except for the addition of the constraint.

The relationship between the CRS and VRS technologies and the separation of technical and scale efficiencies is illustrated in Figure 2.3.3. The constant returns frontier combines technical and scale efficiency. The variable returns to frontier represents pure technical efficiency. The scale efficiency measure if both frontiers have been estimated for a given DMU is the VRS DEA rating divided by the CRS DEA rating. This will be explained more clearly in Chapter 5 when the analysis will be illustrated with real data.

The constant and variable returns to scale technologies are not the only ones it is possible to estimate. It is possible to estimate a non-increasing returns to scale (NIRS) frontier. It has been implemented by Jesson, Mayston and Smith (1987) and Mayston and Smith (1987). The NIRS frontier is a combination of the CRS and VRS frontiers and would correspond to OEC in Figure 2.3.3. Thus, only constant and decreasing returns are possible. This is achieved in the

Figure 2.3.3.

Technical and scale efficiency.



A represents the DMU being evaluated
B represents a technically efficient reference point with the same (output) scale size
E represents a technically and scale efficient reference point at the most productive scale size

Technical and scale efficiency = MN/MA
Pure technical efficiency = MB/MA
Pure scale efficiency = MN/MB

linear program by replacing the constraint $\Sigma\lambda_j = 1$ in the dual with $\Sigma\lambda_j \leq 1$. Using the NIRS program for efficiency measurement against the frontier over 0E, $\lambda_j^* < 1$ and for the segment EC, $\lambda j^* = 1$. The NIRS frontier will be calculated in Chapter 5, but not by introducing this constraint because only a CRS and Banker VRS program is available to me. An inelegant but effective solution (which actually corresponds to $\Sigma\lambda_j \leq 1$) is to introduce a DMU into the analysis whose input and output values are zeros, this then forces the frontier through the origin to create a NIRS frontier.

2.4 Conclusion.

This chapter has described the linear programming basis for the programs used in this thesis. As to which should be used when and in which context this depends upon the data and what we are trying to find out. All of the programs used exhibit strong dispoability of inputs. The variable returns to scale, input minimisation program 2.3.1. will be used to calculate the results in Chapters 5, 6 and 7. The reasons for choosing the VRS progam to calculate efficiency in Post Office Counters and the Area Electricity Boards is that it measures pure technical efficiency and scale inefficiency is not considered to be a problem (this will be discussed more fully in Chapters 5 and 7). The CRS and NIRS programs will be used in Chapter 5 to calculate scale efficiency. Input minimisation is considered a more appropriate assumption than output maximisation when applying the program to the British public sector (again this will be discussed in more detail in Chapters 5 and 7).

Chapter 3.
Conducting a study of Post Office Counters technical efficiency.

3.1 Introduction.

This chapter provides the background to understanding the discussion of Post Office Counters in subsequent chapters. Issues involved in the privatisation and liberalisation of POC are examined and the light an efficiency study would throw on them. The suitability of POC for a DEA study is examined and the information that would be needed to conduct it defined. This is ascertained through an examination of other Post Office efficiency studies, the structure and organisation of POC and the data that it currently generates in the running of the business. The purpose of such a study will then be explained.

Section 2 looks briefly at the history of the Post Office, how Counters came into being and its prospects for liberalisation and/or privatisation. Also, some past Post Office studies are examined, to gain information on the previous efficiency of Counters. In section 3, to gain insight into how a DEA of POC should be conducted in terms of the desirable inputs/outputs an application procedure for DEA is explained. To facilitate the discussion section 4 describes the structure and management of the Counters business and the service it provides. In order to glean more information about the conduct of the study and the selection of inputs and outputs, section 5 examines other DEA studies of post office networks, similar organisations, that is financial institutions such as banks and building societies, and also the way in which Counters measures its own performance. Section 6 brings this information together to evaluate the suitability of POC for a DEA and details the data needed for the analysis.

3.2 The Post Office: an overview.

3.2.1 The history of the Post Office.

Until quite recently the Post Office was and had always been the monopoly supplier of mails and telephone services. Over 350 years ago the decision was taken to give the monopoly on mail deliveries to the Royal Mail. Although competition to supply telegraph services existed for a brief period in the mid-nineteenth century in 1869 the Post Office (a government Department of State) was given a statutory monopoly of inland business. In 1880, four years after Bell patented the telephone, this monopoly was extended to telephone services. The chosen approach was therefore to limit competition and for government to run the industry according to public interest objectives rather than the pursuit of profit.

For a period of about forty years the Post Office granted licenses to private companies and municipal authorities, but the regime of competition and regulation did not work satisfactorily and by 1912 the Post Office had taken over all telecommunications suppliers (except for the municipal authority in Hull). The Post Office remained unique in not following the Morrisonian model of the public corporation, existing as a separate government department. This anomaly was ended by the Post Office Act of 1969 which established it as a separate public corporation, with statutory powers, a financial remit and a ministerially appointed Chairman and Board.

At that time the Post Office still remained responsible for telecommunications as well as for postal and other affairs but the British Telecommunications Act 1981 formally separated it into two independent corporations, with British Telecommunications assuming a separate corporate identity. Thus the Post Office was left on 1 October 1981 with the duty of providing postal and counter services and, through its banking arm National Girobank, banking and money remittance services.

At the same time the 1981 Act reduced the Post Offices exclusive privilege over the conveyancing of letters. The statutory monopoly thus does not now cover the conveyance of letters for which a consideration of not less than a pound is paid, the delivery of Christmas cards by charities, document exchanges, conveyance by air courier, the conveyance and delivery to the Post Office of prepaid letters correspondence flowing between different parts of a body corporate and its wholly-owned subsidiaries, the conveyance and delivery of banking instruments from one branch to another or from a bank to a Government department, and electronic mail.

In a move to implement a change to a more product-oriented type of organisation in 1985 the Post Office reorganised Posts into three distinct businesses, Letters, Parcels and Counters. Girobank has always been a separate entity. This was part of the Post Offices aim to increase its commercial approach, efficiency, competitiveness and responsiveness to the needs of its customers. The reorganisation would facilitate this for each business because it enables them to develop the specialist expertise required by each of the very different markets the Post Office serves. The division into separate operating units is a logical one but can probably be seen as an organisational precursor to privatisation.

3.2.2 The Post Office, liberalisation and privatisation.

The Post Office has been considered a possible candidate for privatisation for quite some time. Beesley and Littlechild recognised this in their 1983 study, prioritising organisations for privatisation they recommend the Post Office. Their criterion for privatisation to go ahead is the 'present value of aggregate net benefits to UK consumers'. This net benefit should be measured 'primarily by lower prices of currently available goods and services (offset by any price increases)', adjusted to take account of 'effects on the level of output, the quality and variety of goods available and the rate of innovation'.

They concluded that the Post Office would benefit from privatisation if appropriately privatised, this depending on whether there is the possibility of restructuring to create multiple (competing) ownership. Privatisation in its present form would not be a good idea because it is a single ownership monopoly supplier of services facing adverse demand prospects. Appropriate privatisation would consist of allowing multiple ownership and liberalising the market by removing the Post Offices letters monopoly. Beesley and Littlechild also say that consumers would benefit from privatisation because the basic distribution network has great potential for development outside traditional Post Office work.

Nothing in the public domain is free from the possibility of privatisation and in a Financial Times (19 October 1986) interview with reference to Post Office privatisation Mrs. Thatcher said that 'the only question is the order and the form'. The sale of the Royal Mail was ruled out by her during the 1987 election campaign though this was probably political expediency. The Post Office monopoly on time-sensitive mail has already been relaxed, and the question arises of whether its monopoly of the letter post should cease, whether or not privatisation occurs. The Chairman of the Post Office has argued against such liberalisation on the grounds that 'cream-skimming' by entrants into profitable business segments and cut-price intra-urban mail services would cause price increases in rural areas and would jeopardise efficiency. This is not just an issue for mails but also for Counters.

The question that should then be addressed given these things are why should government continue to own this business? It is an enterprise competing in a private market and so should be privately owned. There are a number of rationales for privatisation, the main one being that large public sector organisations are inefficient and only privatising them will provide the knowhow and incentive to make them become efficient. But, if the Post Office is operating at optimum efficiency then there is no justification for privatising it. This may be so, but Post Office privatisation is on the agenda and given continuing Conservative government a historical inevitability. Efficiency measurement in the Post Office context is important because it is necessary to produce evidence so that a more reasoned approach can be taken than the position that all public sector organisations are inefficient. If privatisation never takes place performance studies are still important because on an ongoing basis resources must be monitored and controlled and efficiency maintained.

3.2.3 Post Office Counters, liberalisation and privatisation.

What are the issues surrounding the liberalisation and/or privatisation of Counters? It is possible for the privatisation of Counters to take place in its entirety, the Crown Offices could be retained as one organisation and the sub post-offices leased on a franchise basis (it would be necessary to subject Counters to some kind of regulatory control.) The problem with privatisation is that the rural post office network is maintained partly for public service reasons and is cross-subsidised from the more profitable urban network. How could this be maintained after privatisation? These issues will be examined.

It is also possible for liberalisation and the introduction of competition to take place. As to what form this takes depends on whether Counters can be deemed to be a natural monopoly. That is, whether the technical characteristics of production in the situation where there are significant economies of scale, falling average costs throughout the relevant output range as given by demand, mean that production by a single enterprise is most efficient. It would seem though that the services Counters provides are 'contestable' (Baumol, Panzar and Willig [1982]). It has at present been given a monopoly to supply a wide range of government agency services, for example, the payment of pensions and social security. These services fulfil the conditions for a contestable market in that it must be possible for a potential entrant to face no 'sunk' (that is irrecoverable) costs in entering and exiting the market and that the incumbent enterprise is unable to quickly reduce its prices when entry is threatened. It is potentially possible for a competitive supplier to enter the market and take over the job of single supplier (given that the government provides the institutional structure for this to take place. These entrants could be banks and building societies, which already have a network of branches to supply financial services.

The monopoly on government agency services is however, being eroded since many forms of benefit are now paid directly into individual bank accounts at lower cost to the government. The next step in the government's logic must be to invite banks and building societies to tender in competition with the Post Office to provide these other services. But, it is not certain that the banks would be able to compete for this business. Their high street premises are more expensive than the typical post office and carry higher overheads. In 1982 the big clearing banks quoted about 30p as the average cost of a counter transaction, whereas the Post Office's counter services carried out its transactions at an average cost of 22p (though of course the average cost will fall if they increased their volume of transactions by supplying the government agency services). In addition it is questionable whether it is desirable for supply to take place by another organisation given the social benefits of the size of the network and the convenience and accessibility this offers to pensioners and single parents. Even though Counters might expect to retain government business, the liberalisation threat of competition from other financial institutions would remain a powerful spur to maintaining lower costs.

Privatisation on the other hand would not on its own have much of an impact, even on rural services. This is because most rural post offices are sub post-offices. Sub-postmasters are paid on a scale of fees representing the work they do. This is measured accurately because all their materials, such as postal orders, vehicle licences and so on, are supplied from the local head post office.

If the scale of fees remains the same then unless there is a serious fall off in business the situation will remain the same. Even if the scale payments fell due to an end to cross-subsidisation it is unlikely that many would close. It is a coveted job and it is common when a sub-postmaster ceases to work for the Post Office to have to choose between a number of applicants. If a village has no sub-postmaster it is generally because the Post Office refuses to appoint as nobody suitable has been found. If it were the case on social grounds that a sub-post office were thought necessary and no-one was willing to do the job at the rate of scale payments then a direct government subsidy could be paid.

The impact of privatisation on the Crown Post Offices could be somewhat different. Senior (1983) argues that privatisation is necessary to Counters network survival because it will provide a spur to modernisation. He says that the Post Offices Counters services are in danger of becoming obsolete because of the technological revolution in transferring funds. It offers no automated teller machines to dispense cash and provide other services outside counter service hours, it is not experimenting with the transfer of funds electronically at the point of retail sale or with home banking. It was only in 1983 that the first trials began in using electronic registers in Crown offices (ECCO) to record transactions and balance at the end of the day. Senior warns that many Crown Offices are likely to become redundant under the impact of home banking and other technological developments unless they diversify their activities to justify the high cost of prime high street sites.

Since Senior's prediction there have been a number of developments. Counters has diversified some of its activities, for instance, it has developed a limited range of retailing activities, it provides some information services and franchises some services. Plans for automation exist, ECCO has completed its trial and there are plans for an 'Automation Scheme'. The Automation Scheme is planned to follow a pilot scheme currently being conducted in some 250 Crown and Sub-offices in the Thames Valley. It is concerned with the development of a degree of automation of some aspects of the work of the counter clerk at Crown and some sub-offices. It is associated with a data transmission network, controlled from a central site, and will have on-line facilities to some of Counters clients. But, the Monopolies and Mergers Commission Report 'Post Office Counters Services' (1988) criticises both these schemes. It says that it is unlikely to be possible to justify both the Automation Scheme and ECCO, 'we regard the independent development and appraisal of the Automation and ECCO schemes as a failure on Counters part'. Initially these two schemes were distinct projects operating to different time-scales and providing non-overlapping capabilities. But now there is overlap, for instance, balance and summarisation had not been part of the original scheme for automation of counters. If only one of the schemes is considered then the full ECCO scheme may be economically justified but the NPV value estimated for the Automation Scheme is not robust because this would depend upon the continuing business of existing clients, most of whom would not say if they would stay with the Post Office in the long-term. However, recent events do confirm that product innovation and modernisation can be initiated without recourse to privatisation.

3.2.4 Past assessments of the Post Offices performance.

Given then that one of the arguments for privatisation is that it improves performance then this proposition has to be tested. If it is the case that the post office is already efficient then there would be a lot less justification for privatisation taking place. As most of the reasons cited for privatisation are ideological and political rather than economic a study would contribute to the debate. It was concluded in chapter one that technical efficiency was the most useful way of looking at measuring performance using data envelopment analysis. Before it is explained which aspects of the Counters business will be examined, it is interesting to look at some previous empirical work on Post Office services in this country and abroad.

Pryke (1980) conducted a study of nationalised industry efficiency which included the Post Office. He examined changes over three time periods 1968-73, 1973-78 and 1968-78. He concluded that labour productivity had been falling throughout the period. There is a problem in Prykes work in that he assumes that data on changes in labour productivity can be used as indicators of change in technical efficiency. This provides only a partial picture because it does not identify the contribution of capital. Also, the activities of Counters are not examined as the study is of Postal Services.

Molyneux and Thompson have conducted a similar study for the period 1978-1985. They use labour productivity as a measure and record an increase over the period of 2.3%. They also record an increase in total factor productivity (TFP) of 1.9%, though this measure has its limitations in that it does not just measure changes in technical efficiency but changes in technical progress as well. It is also very difficult to adequately quantify capital which is included in the TFP measure.

In addition, the Post Office has also been the subject of scrutiny by government. The first report of its kind was conducted by the Post Office Review Committee in 1977, known as the 'Carter Committee Report' (CCR). Its main recommendation was that the Post Office should be split into two businesses, one comprising of Posts, the other of telecommunications. The reason for this recommendation, which was subsequently acted upon, was the very different management problems each organisation presented. The postal-service was labour-intensive with a low demand for capital investment and a declining volume of business and quality of service, telecommunications was capital intensive, has a need for technical innovation and an expanding volume and range of facilities. Increases in efficiency in the postal service thus requires an efficient use of labour, in telecommunications the key factors relate to efficiency in the use of expensive capital as well as labour.

As to the actual efficiency of the Post Office (we are more interested in Posts because it contains Counters) at that point in time, the Carter Committee Report cites findings of a NEDO (1976) report. Using the partial labour productivity measure it was found that over the period 1960-75 productivity had been static and then fallen. Again, this does not actually tell us very much specifically about Counters. The report does examine Counters briefly but does no more than give its support to the existing situation in terms of approving of the distribution and organisation of the Counters network.

The government has also charged the Monopolies and Mergers Commission (MMC) to conduct efficiency studies of different aspects of the Post Office's business, there have been four such reports. The first related to Mail services in Inner London was published in 1980, the second in 1984 examined the whole letter post service, the third was related to the procurement activities of the Post Office and published in 1986 and the most recent, published in 1988, examined counter services. The 1988 MMC report on Counters had a number of criticisms of performance and the way in which it is measured. They came to the conclusion that working practices and clerical procedures were outdated and inefficient and that staff in Crown Offices could be utilised more effectively if staff scheduling procedures were improved. There is an absence of physical throughput standards for branch offices preventing any meaningful assessment of the performance of its staff. In addition, the performance measurements that are made are based solely on counter transactions which account for less than half the total workload of most offices.

Academic studies to date have concentrated on the Post Office as a whole. The most recent period these studies cover is 1985, since which substantial changes have taken place in the Post Offices organisation. Government studies have tended to concentrate on different aspects of the Post Office, in particular Mail services. The most relevant government report to date was the 1988 MMC report which was critical of performance but noted that many positive changes were taking place. These criticisms will be examined in more detail in Chapter 5 and related to the DEA findings.

3.3 How should a DEA study be conducted?

3.3.1 An application procedure for DEA.

Insight into the problem of deciding which model to use and how the DEA analysis should be conducted can be gained by examining other Post Office studies, DEA studies of similar financial institutions and by examination of the way Counters measures its own performance at present. These will be examined shortly, but also Golany and Roll (1989) have suggested a systematic application procedure for conducting an efficiency study using DEA which is worth examining. They suggest that there are three main phases in the process:

i. Definition and selection of DMUs to enter the analysis.

ii. Determination of input and output factors which are relevant and suitable for assessing the relative efficiency of the selected DMUs.

iii. Application of the DEA models and analysis of outcomes.

These will be loosely followed in conducting this study.

Work corresponding to phases i. and ii. will be conducted in this chapter and a summary of the resulting factors to be included will be put forward in the conclusion. The reasons for the final choice of DMUs and factors, the models parameters and the results, will be given in Chapter 5. A more detailed explanation of each phase will be now be given.

3.3.2 How should DMUs be selected?

For a group of DMUs to be selected for DEA they should be homogenous, and the size and boundaries of the group should be determined. As DEA is a relative efficiency measure for comparing 'like' DMUs then a homogenous set of units has to be identified. The units will have the same tasks and objectives, perform under the same set of 'market' conditions and the factors characterising the performance of all units in the group are identified except for differences in intensity or magnitude.

Determining the size of the comparison group is a problem because there are two conflicting considerations. It is an advantage to have a large number of DMUs because there is a greater probability of capturing high performance units which would determine the efficiency frontier and there will be a sharper identification of typical relations between inputs and outputs in the set. Also, as the number of units increases, it is possible to incorporate more factors into the analysis. But, on the other hand, the larger the number of units in the analysed set, the lower the homogeneity within the set, increasing the possibility that results may be affected by some exogenous factors which are not of interest.

There are two kinds of boundaries to be examined when determining the DMUs to be entered into a DEA. There are the organisational, physical or regional boundaries which define the individual units. Then there are the time periods used in measuring the DMUs activities. Preferably these should be 'natural' ones corresponding to seasonal cycles or budgeting periods. It should be borne in mind that long periods may obscure important changes occurring within them, while short periods may give an incomplete picture of the DMUs activities. Other considerations are, how far back to go without distorting the comparison and whether to create an overlap of data by means of 'window analysis' (see chapter 7 for application and discussion of this).

Having established the set of DMUs to be used in the analysis, it must be borne in mind that there may be DMUs that may be considered as outliers. That is, units or time periods deviating from the general characterisation of the group to be analysed. These need to be separated, possibly with the help of managers.

3.3.3 How should input and output factors be selected?

The next phase is to determine the factors to be included in the DEA analysis as input and output variables. The initial number of possible factors should be wide and anything that affects the DMUs being evaluated should be included. The factors may be fully or partially controllable by the DMUs or they may be 'environmental' and outside the control of the DMUs. Factors could be quantitative or qualitative, with different degrees of difficulty accorded numerical values. The variables could be inputs or outputs or factors placed on either side of a production relationship.

Such a list of possible factors could result in a very large number being identified (Thomas [1985] managed to list 92 factors, which he reduced down to 14). Large numbers of factors will tend to 'explain away' differences between DMUs. This is because, more factors will increase efficiency scores by shifting the compared units towards the efficiency frontier. It would seem advisable then, to accentuate the basic differences among units, to keep the number of factors as low as possible. It is suggested by Golany and Roll that the best list of factors can be determined by going through the following three stages:

i. Judgemental screening

ii. Non-DEA quantitative analysis

iii. DEA based analysis.

i. Judgemental screening.

To reduce the list of chosen factors to a desirable level there could be a critical examination of them by expert decision-makers in the field where the DMUs operate. As the initial list of possible factors is probably large there will be scope for reduction because some factors will be only of minor importance. Other factors may be conflicting or repeating the same information.

A problem at this stage for choosing a set of inputs/outputs is the difference between factors determining efficiency and factors explaining efficiency gaps. For example, labour input may serve to determine efficiency while the scale at which the unit operates may be an explaining factor. Including explaining factors in the analysis may blur the overall picture and reduce the measured distinction between compared units. If this is because explaining factors account for inefficiency, then the differences are being explained away.

Given that data on these factors is available and reliable then maybe some systematic procedure could be applied for structuring the judgemental process. For instance, Delphi-like techniques or varieties or varieties of the analytic hierarchy process (Saaty [1980]).

ii. Non-quantitative methods.

The various factors should now be assigned numerical values. These can be the physical units by which they are measured or their financial value. If the factors can be measured in economic terms it may be decided that some or all of them should be aggregated. This is effectively assigning fixed relations among the weights of the factors, whether this is thought desirable or not depends upon the objectives of the analysis. Another issue that may give cause for concern is the incidence of cases zero values appear in the factors. This could occur if the time period chosen does not correspond to a 'natural' cycle of operation, or for some data gathering reason. It is not a problem as long as there exists at least one input and one output for each DMU which is non-zero, but care should be taken as the computational algorithms may be sensitive to zero values.

Another important point concerning the value of factors is that an increase in any output should not result in a decrease in any output (isotonicity). To achieve this may mean that some factors need rescaling. This can be done through inverting them, or assigning a negative value so that inputs are positively correlated to outputs.

Qualitative factors can also be included and need to be assigned numerical values. This is usually done by locating a measurable surrogate variable which is assumed to bear a known relation to varying degrees of the qualitative factor. There needs to be a high degree of correspondence between variations in the surrogate data and the examined factor, so there will need to be some trial and error in its choice. Once the surrogate factor has been identified it needs to be expressed in a functional form and must comply with the analyses objectives.

The next step requires the list of factors to be classified into inputs and outputs. Resources utilised by units or conditions affecting their operation are typical inputs whilst measurable benefits generated constitute the outputs. However, some factors may be interpreted as in both ways, depending upon the point of view of the analyst. A way of clarifying the problem could be to carry out a series of regression analyses of each factor, one at a time. A weak relationship to inputs and strong relation to outputs indicates a preference towards classifying the factor as an input, while a reverse outcome will point towards viewing the factor as an output.

Regression analysis could be used to make the set of factors defined as inputs and outputs smaller, defining and eliminating less relevant factors. There has been some discussion about using regression analysis to produce a reduced variable set. Charnes, Cooper and Rhodes (1981) imply that variables which are highly correlated with existing model variables can be omitted from further analysis without significantly affecting the DEA efficiency results. This viewpoint has been formalised by Lewin, Morey and Cook (1982), who suggest the use of regression and correlation techniques for validating the relevant set of input and output measures. Their variable selection methodology assumes that the addition of a highly correlated variable will have an insignificant impact upon subsequent DEA results, and this can be omitted from further consideration.

On the other hand Nunamaker (1985) indicates that for selected DMUs addition of a highly correlated variable may alter substantially the DEA efficiency evaluations. Addition of a correlated variable to a regression model will add little to the mean square accounted for by the existing independent variables. DEA, however, is not based upon any 'squared distance from the mean' notions. The existence of high correlation among variables does not necessarily mean that one of them can be excluded without changing the subsequent DEA results. But, it is not possible to know a priori which DMUs will be affected, it depends upon the specific DMU under examination.

Thus strict reliance on regression and correlation analysis as a way of achieving variable reduction is not advisable as unexpected efficiency evaluations may arise for individual DMUs. However omission of a correlated variable may have a lesser overall impact on the results than would some other, less associated factor. Therefore correlation techniques may only be used to assist the decision-maker in selecting a reduced variable set.

iii. DEA-based analysis.

The final step in the process of choosing and refining the list of factors will be the use of DEA itself. Trial runs of models containing different combinations of factors can be compared. Factors which have very small weights, and so have little impact on the efficiency scores, could be dropped. In addition, factors which make little contribution to efficiency scores do not aid discrimination between DMUs and so are not useful. The discriminatory power of the different factors is tested by running the model with a series of combinations of these factors. DMUs can be grouped using the resulting efficiency and factors which do not alter such groupings significantly should be examined closely. Special attention should be given to factors which could not be easily classified as inputs or outputs, as these can be tested to decide where they are most suited.

The selection process for variables can be repeated as many times as it is thought necessary until a final list of factors is decided upon.

3.4 Organisation of the Counters business.

3.4.1 Structure and management of the Counters business.

To facilitate an understanding of the analysis of Counters in this thesis it is necessary to know something of its organisation and structure and the service it provides, this will now be explained.

Counters became a separate business from Letters and Parcels, the other Mails businesses in 1985. It was fully separated below headquarters level in October 1986. Post Office Counters Ltd. (Counters Ltd.) was incorporated into and became a wholly-owned subsidiary of the Post Office Corporation on 1st October 1987. Counters Headquarters is situated in London with large groups of staff outstationed in Chesterfield and Edinburgh carrying out routine finance and personnel work.

The Counters business controls 21,000 local post offices providing counter services throughout the United Kingdom. The annual turnover of the network, measured in terms of cash passing over the counter is over £86 billion. It is estimated that there are 1.3 billion visits by customers per annum, which makes it the largest shop or financial services network in the country.

The Post Office counter network is made up of two kinds of offices. Crown offices, usually the largest offices situated in city and town centres and in major conurbations, are managed and staffed by Post Office employees. Sub-post offices which are often but not always associated with other retail activities, are managed for the Post Office by Sub-Postmasters who are paid according to a scale of fees which is related to the number of transactions carried out in the individual offices. For this reason the sub-post offices are also known as scale payment sub-offices.

Before 1985 Counters was organised on the basis of the ten postal regions (see Appendix 1). After a re-organisation the local offices for both Crown and sub-post offices are now administered by a management structure which divides the United Kingdom into four territories, London, Eastern, Western and Northern (which includes Scotland and Northern Ireland) and into 32 districts. The territories and districts are listed in Appendix 2. A map showing this organisation is in Appendix 3. The territorial tier of management is an 'outstation' of its Headquarters with a restricted role, but the districts are self-contained management units. Each district is divided into a number of areas, typically four, under an Area Manager. The Area Manager is expected to manage an average of nine Crown Offices and 120 sub-offices and is held responsible for the efficient operation of all Crown Offices in the area. Branch Managers have operational control of individual CPOs. The Branch Manager has overall responsibility for the effective management of the office and is required to supervise all aspects of client and customer service, buildings and security, expenditure and accounting and staff scheduling. Figure 3.4.1 shows the organisation of the Counters business. Its present board are all members of the main Post Office Board but in practise the short to medium-term management of the Counters business is carried out by a 'Counters Management Committee'. This committee includes the directors of the Headquarters departments of the business and the Territorial General Managers under the chairmanship of the Managing Director who is the corporate Board member for Counters and as such the only manager of the Counters business on the Board for Counters Ltd.

3.4.2 The Post Office Counters service network.

The network of offices providing services is divided between Crown offices staffed and operated by the Counters business and sub-offices owned and operated by Sub-Postmasters under contract to the Counters business. Sub-offices are further classified as town or rural depending upon location. In April 1987 there were approximately 1,500 Crown offices and 20,000 sub-offices of which approximately 9,000 were in urban areas. It is these offices which will be the production units or DMUs in this study of comparative efficiency.

In urban areas the split of business between Crown and sub-offices is largely the result of history. In the past there were no specific guidelines for determining through

Figure 3.4.1. Organisation of Counters.

POST OFFICE
COUNTERS LTD
(WHOLLY OWNED BY THE)
POST OFFICE CORPORATION)

|

COUNTERS
MANAGEMENT
COMMITTEE

|

COUNTERS
HEADQUARTERS

|

TERRITORIAL MANAGER OF EASTERN TERRITORY	TERRITORIAL MANAGER OF WESTERN TERRITORY	TERRITORIAL MANAGER OF NORTHERN TERRITORY	TERRITORIAL MANAGER OF LONDON TERRITORY
7 DISTRICTS	8 DISTRICTS	11 DISTRICTS	6 DISTRICTS

Note: This chart only shows the main features of the organisation and excludes the individual management reporting lines eg. the General Managers of territories reports directly to the Managing Director of Counters and there is a functional relationship between the staff of the finance, marketing and operational branches in districts, territories and Headquarters.

which type of outlet service should be provided. Crown offices were generally chosen if it was necessary to accept mail for an associated sorting office, if business levels required a 'large' office, or in response to pressure from the local community because of the enhanced status which it was felt that such an office would confer. In 1986-7 Crown offices and urban sub-offices together transacted 89 per cent of total Counters business, 11 per cent being handled by rural sub-offices. Urban sub-offices handled an estimated 61 per cent of Counters business in urban areas or 54 percent of all Counters business.

Although sub-offices handle the larger share of urban business they are on average much smaller than Crown Offices, with an average of only 1 million units each, but only 24 Crown offices are in this range. At the other extreme there are 131 Crown offices (8 per cent) handling over 6 million units each and no town sub-offices in that range. There is however, a substantial overlap in the range of 1 million to 3 million units per annum.

Sub-postmasters are paid on the basis of unit credits. These are related to the time taken to carry out a transaction together with associated balancing and summarisation procedures. The total unit credits earned in a year determine the Sub-Postmasters remuneration on what is essentially a piece-rate payment basis. Sub-Postmasters are also paid various allowances.

Originally the purpose of the network was to provide a means for the general public to transact a wide range of business, exclusively Post Office, and largely mails. During and since the Second World War, however, there has been a substantial growth of agency work for Government departments, for example the payment of pensions, and this now accounts for approximately half of all business. The requirements for a network of offices are therefore not necessarily the same as 50 years ago, although the distribution of offices across the country remains substantially as it was. In 1982 the Post Office set up a study of its network as part of a general programme to reduce costs so as to avoid a decline in business as prices rose. Rural sub-post offices were excluded from the review. The study found that although there had been some reduction in the number of offices since 1945, urban areas were still over-provided with service. The Post Office therefore decided to reduce the number of its offices by up to 1,190 and within that figure to close or regrade 72 Crown offices. As a result of this review during the period 1983-84 to 1986-87, 72 Crown offices were either closed, merged or regraded as sub-offices (during the same period five were opened and two upgraded.) Twenty-five of the Crown offices in London were closed, two by being merged with others. The estimated net ongoing annual savings from the Counters network review were £9.74 million of which £4.88million were Crown office savings.

3.5 How has efficiency measurement of counters services previously taken place?

3.5.1 Why do we need to examine past studies?

A DEA study of a post office network including Counters services has taken place before. In addition, there are a number of DEA studies that examine financial institutions such as banks and building societies which operate in a similar way to a Crown Post office. They both provide over the counter financial services through a branch network. It would be thought that an examination of such studies would provide valuable insight into how to evaluate the Counters network and which inputs and outputs to use in the DEA. So also, would reviewing the methods by which Counters evaluates its own performance and these will also be examined in this section.

3.5.2 Deprins, Simar and Tulkens Study of the Belgian Post Office.

The most relevant study to the type of efficiency study that is going to be conducted is by Deprins, Simar and Tulkens (1984). Deprins, Simar and Tulkens (DST) set out to compare the technical efficiency of the Belgian Post Office, the Regie des Postes, by comparing post offices across one month. They measure technical efficiency using three methods. Firstly, they adjust a Cobb-Douglas production frontier, secondly they compute a Farrell frontier, that is, a non-parametric deterministic convex hull of the data and thirdly, a method of their own is introduced which is based on the sole assumptions of input and output disposability.

The individual DMUs in the DST study are individual postal stations. They differ from the postal services in this country in that mail collection and delivery centres are at the same place as all window financial services in a majority of cases. DST chose to look at the 792 stations (out of a total of 972) that have these characteristics. Sorting centres are excluded from the enquiry. So Mail and Counters services are examined at the same time. There are 137 different measurable outputs but only six are used, these can be summarised as the number of financial window operations performed, the number of items handled for four different types of mail and the number of delivery points. The sole input was labour as this is the main input, 80% of expenditure is on salaries.

DST conclude that there are inherent limitations to the first two methods, at least as they have appeared in this case and favour their own method. The reason they discount the farrell frontier is due to its assumption of constant returns to scale. Whilst DST think this assumption is unsuitable in their context it should not discount its use in other situations. In my own study a variation is introduced which allows for variable returns to scale (an option not available to DST at that time).

DST do not then just look at the raw technical efficiency numbers but attempt to give them some meaning by taking their average and median, and also evaluating the shape of a frequency distribution of efficiency measures, taking into account the number of extremely inefficient observations and the spread of heavily inefficient observations. Averaging the technical efficiency measures gives an overall technical efficiency of 0.89 (and a median above 0.96), with 48% of the offices being fully efficient (and 66% with technical efficiency of greater than 0.88). Extremely inefficient observations are very rare, only about 9 offices (about 1% of them) have a technical efficiency of less than 0.40 and heavily inefficient observations are not that many, there are 60 offices (about 7.6% of them) for which technical efficiency is less than 0.60. They conclude that 'the over-all picture that we come up with is not an unfavourable one, given the size of the enterprise'.

The DST study provides some useful information about how to conduct a study of the British Postal Service. Firstly, with a very large data set, the DEA ratings can only be analysed for their overall technical efficiency by displaying as a frequency distribution with accompanying statistical summary. Conclusions can be reached on the basis of the spread of results though this could be somewhat objective. Secondly, it suggests the kind of data on inputs and outputs that is needed. It is clear that Labour should be the main input as capital is hard to measure and makes a smaller contribution, DST also point out that their results would have been better if they had distinguished between categories of labour instead of assuming it to be homogenous. This is not possible with Counters as just financial window transactions will be concentrated on and we are only interested in one category of labour. The outputs used by DST, excluding those relating to Mails business, are the transactions performed on the counter.

3.5.3 DEA studies of other financial institutions.

Examining other DEA studies of financial institutions can provide useful insight into what information is needed to conduct a DEA study of POC. This is because they provide similar services and exist in similar operating environments. It should be emphasized that though all of these are in the private rather than the public sector, the principles relating to technical efficiency are still the same. The institutions normally examined are banks and building societies. There are not a great number of UK studies because this is quite a recent field of application. All the studies produce results for overall technical efficiency, that is, an analysis with a constant returns frontier, but some examine other aspects of efficiency.

Most of the research has been undertaken in the area of US banking. Sherman and Gold (1985) analyse overall technical efficiency at the branch level, while the analysis of Rangan, Grabowski, Aly and Pasurka (1988) decomposes technical efficiency and scale efficiency. In addition, Aly, Grabowski, Pasurka and Rangan (1990) also examine the issue of allocative efficiency. A Canadian study has been conducted by Parkan (1987), he examines overall technical efficiency of a chartered bank in Calgary. Another straight application has been conducted by Vassiloglou and Giokas (1990) on the Commercial Bank of Greece.

In the UK there have been two studies of financial institutions, both of which are building societies. Field (1990) examines 71 building societies using cross-section data from 1981. After 1985 they were given a new role which makes the societies compete and act as if they are banking institutions. He calculates technical efficiency and its components, as well as a measure of congestion efficiency. These different measures are then related to building society size and it is concluded that there is a significant negative relationship between overall technical efficiency and the size of societies. This means that the expansion of societies is not justified economically. Drake and

Weyman-Jones (1991) also conduct a study of UK building societies but cover new ground because of the increase in competition and product diversification that took place throughout the 1980s. They use data from 1988 on a sample of 76 building societies to calculate overall, technical and scale efficiencies and correlate these to building society size. They too find a significant inverse relationship between size and overall efficiency, signifying excessive numbers of branches. Also, the relationship with size and scale efficiency indicates that they are operating beyond the efficient scale of production.

Our primary interest in these DEA studies of financial institutions is to see which inputs and outputs they use in their analysis. The best way to examine them is to summarise the inputs/outputs used and then see which could be useful in a DEA study of POC. All inputs and outputs used in the studies are summarised in Table 3.5.1. Inputs and outputs are usually measured according to their total value or the number of units used and so how they are classified is also included in the table. This is not to say that there do not exist other valid inputs/outputs, or that they could not have been classified differently, it is just the way that other studies have been conducted.

The inputs/outputs previously used will now be examined to see if they have relevance in the POC context. Labour is of course a major input and it would be thought best that it is measured in terms of the number of hours worked as the best indicator of the size of the input. I do not think office supplies are that important as an input because they are a function of the number of hours worked by staff. Computer equipment is not a valid input at this point because as will be seen later POC is only very partially computerised. Capital is a problematic input because the value of a branches premises will vary according to region. Office space is a better proxy for size of branch because it does not take account of value. But then it also does not take into account the problem of badly utilised space due to characteristics of the building. The number of branches is

Table 3.5.1

Summary of inputs and outputs used in the studies of financial institutions.

Inputs

1. Labour (Full or part-time, classified by number or hours)
2. Office supplies (value)
3. Computer equipment (value or number of terminals)
4. Capital (value of property owned and other such assets)
5. Office space (rental cost or square metres)
6. Deposits (value, there can be different types)
7. Branches (could be offices or agencies)
8. Quality of customer service area (rating 7-35)
9. Marketing activity ranking (rating 0-50)

Outputs

1. Loans (value of different types to different types of borrowers)
2. Deposits (value)
3. Transactions (number, divided into groups on the basis of complexity)
4. Liquid asset holdings.
5. Customer services quality service rating.

not relevant because we will be comparing individual branches rather than organisations. The value of deposits is not a valid input because POC is not a bank, Girobank is the banking arm of the Post Office. Counters does accept deposits and withdrawals from Girobank customers.

The last two inputs were suggested by Parkan and are unique to his study. The quality of the customer service area was evaluated by the Banks offices, so the variable is necessarily subjective and open to manipulation. Though it would seem not to be relevant to most banks and building societies because they will have decoration regularly updated and have a corporate design ethos. Counters offices vary widely in their interior decoration because many of its branches are old and distinctive, office style is considered as an aspect of quality and this will be looked at later on. The marketing activity ranking was the rating by a superior of a branch managers involvement in their branches promotion. A concept not applicable to POC because demand for services is largely endogenous and any service promotion takes place at national level.

Of the outputs only the transactions and quality categories are relevant to Counters, the others are specific to banking institutions. The main function of POC is to act as an agent to provide financial services to individuals on behalf of a client organisation. It also provides services of its own. Its main output then is the transactions it performs to provide these services. These could be categorised on the basis of complexity as suggested by Sherman and Gold. Or, if there are a large number of types, the main transactions that form most of the work could be grouped and the remainder put into a category of their own. It is desirable not to group too many different types of transaction because in the end, the complexity of each type of transaction is different. A customer quality of service survey a is valid method of evaluating service quality, but is not necessarily that relevant to the Post Office setting. Unlike banks and building societies customer contact is limited to dealing with a transaction, services do not have to be discussed. Counters measures quality of service in terms of time taken for the customer to be served (this will be explained in more detail later on).

This is only a preliminary examination of which inputs/ outputs to use in a DEA study of POC based on those used in similar studies. A more complete analysis will be made after an examination of Counters. It could well be the case that some of the data used will be defined by that which is available.

3.5.4 Counters services.

Having looked at the way other studies have been conducted and which inputs and outputs they have used, it is now necessary to examine Counters, how it evaluates its own performance and defines its inputs and outputs. Starting with an explanation of the services it provides as these will be the outputs in a DEA.

About 96 per cent of Counters income arises from cash transactions with members of the public (its 'customers') undertaken on an agency basis on behalf of clients (mainly Government departments and other Post Office businesses). Counters carried out some 2.3 billion transactions in 1986-87, with a total cash flow over counters of £86 billion and a total income to Counters of £732 million.

Counters carries out around 190 different types of transaction, although only six transactions for clients account for about one-half of Counters volume in terms of BTHs.

Counters has classified its main activities into six markets.

a. Payment of Government Benefits, mainly on behalf of Health and Social Security (DHSS), accounting for about 39 per cent of Counters volume. Main benefits are pensions, child benefits and supplementary benefit (paid on presentation of order books) and unemployment benefit (paid by encashment of Green Girocheques). Green Girocheques are issued by the DHSS which negotiate terms with Girobank.

b. Services for letters and parcels, (some 23 per cent of volume) sale of stamps represents the majority of this activity, but some Counters also, for example, provide a collection service for franked mail, packets, parcels and Datapost, and collects revenues from franking machines.

c. Bill Payments, of British Telecom (BT) bills, or using Girobanks Transcash facility, including savings stamp schemes accounting for about 10 per cent of volume.

d. Tax Collection, on behalf of the Driver and Vehicle Licensing Centre (DVLC) and the Home Office (television licenses), including savings stamp schemes, accounting for some 10 per cent of volume.

e. Banking, as agent for Girobank (some 8 per cent of volume). Main activities include receipt of corporate deposits, and personal account deposits and withdrawals.

f. Savings and investment, on behalf of the Department for National Savings (in particular National Savings Bank withdrawals and deposits) accounting for some 5 per cent of volume.

Other activities of Counters (together accounting for less than 5 per cent of volume) include issuing of British visitors passports and a number of transactions for local authorities, local transport undertakings and water utilities. Counters also has a limited number of its own products and services. It issues postal orders, it has developed a limited range of retailing activities (including Postshops) and philatelic outlets, it also provides some information services (including POPOS displays and video screens) and it franchises some services such as photobooths.

3.5.5 Quality of service.

Thus, for the purpose of a DEA the output of the services that Counters provides is the transactions that take place with customers. Another dimension of output that can be included in a DEA is the quality of that output, quality of service needs to be defined.

Overall, Counters sees quality of service to its customers as covering:

a. speed of service
b. office style, and
c. convenience/accessibility of the office network.

Inspections of Crown offices report on office style, which includes such items as office environment, accessibility of information to the public, staff professionalism, attitudes and appearance. Convenience and accessibility will be considered during any network planning study. But, by their very nature only qualitative assessments can be made of their level. On the other hand speed of service can be quantified. This is a measure of the time that customers may be required to wait before their service commences.

Quality of service to clients could be included in this study. It is measured by the performance that is achieved in processing documents. The measure of service is based either on error rates incurred in the completion of the customer transactions or on document transmission times. It can also be concerned with fraud. But, data relating to the quality of performance for clients is considered confidential and detrimental to their interests if published. Therefore it is not possible to use this aspect of performance in the analysis.

The Counters 'speed-of-service' measures are concerned with waiting time and two such measures are in regular use for Crown offices:

a. percentile of the waiting time distribution, determined from a sample of customer waiting times to indicate the percentage of customers that are served within a given time period of joining the queue, and

b. average waiting time, an estimate of the average time within a period that customers spend waiting for a counter clerk to become available to serve them.

a. is a national target level and is expressed as '95 per cent of customers to be served within five minutes'. The five minutes is the time that a customer spends in the queue before the customer's service time begins with the counter clerk (the choice of the particular target level which is applicable to all Crown offices was based on market research which indicated that three out of four customers would find a five-minute wait acceptable). Districts may also operate local quality of service percentile target levels when these are considered to be useful in improving the quality of service performance of particular post offices. The second measure of quality of service to customers is the average waiting time. Counters has no official target for this measure.

Counters quality of service monitoring procedures require each Branch Manager to record by daily sampling the following information of customer arrivals:

a. the number of customers entering the office during the sample period (obtained by taking readings from customer counting devices before and after the target customer is served)

b. the queue length (again, before and after the target customer is served)

c. waiting time before target customer is served

d. the number of counters open and with clerks serving or available for service

e. the number of counter clerks scheduled to be available for serving

f. comments relevant to items a-e.

The procedure requires the Branch Manager to take 22 samples in each week, each sampling requiring the selection of a particular customer. The sampling frame is chosen to cover each weekday with four samples and the remaining two on Saturday. All completed sample data is sent to the relevant district offices for computer analysis and the production of the following statistics for each Crown Office for circulation to their Branch and Area Managers:

a. the average waiting time in seconds

b. percentage of customers waiting five minutes or more

c. percentage of customers waiting three minutes or more

d. percentage of customers waiting one minute or less

e. number of valid samples, and

f. offices achieving borderline to, or failing, the national target (or any variation set).

In response to the information gleaned from its monitoring Counters is engaged in a number of initiatives to improve quality of service. The present quality of service targets are seen as an initial step towards improving customer service. As individual offices progress towards these national targets, factors that affect quality of service in each one will be used to introduce an element of individual targeting.

In addition there is an audit of the sampling forms submitted by Branch Managers to check that they are collecting data in the way laid down by in the instructions. The second stage in this audit is that a mechanism will be provided for determining the most effective action to take to improve quality of service data.

3.5.6 Current methods of measuring efficiency used by the Post Office.

Before the way in which the study is to be conducted is defined it might be useful to look at how Counters measures its performance at present.

To measure the productivity of a Counters office the 'throughput' or output of that office must be known. Counters measures throughput using Counter Work Units (CWUs) and Basic Transaction Hours (BTHs) CWUs are used only locally. The basis of both measures is multiplication of the number of transactions performed by the time taken to perform them. There are two main sources of error in measuring the number of transactions performed:

a. Items not measured, aborted transactions
- general enquiries

b. Items imperfectly measured

i. Assessment of certain items based on an annual count and included in the assessment of an offices output as a standard allowance.

ii. Assessment based on average monetary value of transactions

- this includes most types of stamps.

iii. Items assessed by annual count, but not included in workload. eg. packets, meter postings, undefinable items and poste restante.

iv. Other items: Franchises and POPOS.

The amount of time taken to perform a transaction is determined by measuring its Counter Transaction Time (CTT). CTTs are worked out by sampling an office every day of the year (including Saturdays) except for the four-week period leading up to Christmas. A stop-watch is used to time individual clerks at a counter for a period of 20 minutes and the time taken for each separate transaction is recorded in seconds. The times for each type of transaction are then averaged over all of the offices sampled. There are sources of error in calculating the CTT and this reduces their comparability between offices. Times allotted to transactions are based on 'actual time taken' and do not include time spent answering customers questions, aborted transactions, or waiting time between customers. Whilst strict adherence to defined transaction procedure is checked the effort of the counter clerk is not assessed and no allowances are added for relaxation or contingencies. There are sources of error in calculating some CTTs as they are estimated on the basis of the financial value of the transaction. National samples are used to calculate the average monetary value of transactions involving the sale of stamps, gift vouchers and browser unit items. Average times are calculated for each type of transaction and CTTs for these are expressed as a time per £10 value. But, for outlets with high proportions of bulk sales the present measures of time distorted actual performance levels.

BTHs when divided by staff hours provide the productivity measure used by the Post Office, counter throughput per hour. There are a number of problems with using this measure. The performance measurements that are made are based solely on counter transactions which account for less than half the total workload of most offices. Back-office work hours and procedures need to be measured.

The workload and hours used for both BTH and CWU calculations are based on different accounting periods. Sometimes, therefore, the result is that misleading throughputs are reported. Furthermore, the present system does not easily permit management to distinguish between increased productivity achieved by increasing staff occupancy and that achieved by increasing the number of standard transactions performed per occupied staff hour.

Because of these things and the problems previously cited with measuring numbers of transactions the 1988 MMC report concluded that 'the existing methods of measuring productivity using BTH throughput ratios are not precise enough to make valid comparisons between offices'. Whilst these problems exist because of drawbacks in data collection and methodology. it remains to be seen whether the data necessary for my own study can be extracted from that which is available.

3.5.7 How does Counters determine its labour inputs?

Controlling resources to improve efficiency in the Counters context means calculating the correct level of labour inputs. Counters has two means of determining staff hours required and the numbers of staff to be employed at individual Crown Offices. The first is a full-scale staffing inspection known as a COCSI (derived from Crown Office Counter Staffing Instructions) the second is a relatively short procedure known as 'desk-top review'. The decision to carry out an inspection or review of staff numbers and duties at a particular office is made at area or district management level. Once such a revision has been carried out the Branch Manager has no power to adjust staff numbers and only limited scope for permanent adjustments to the staff schedule.

Desk-top reviews were widely adopted following the Revised Counter and Writing Staffing Arrangements (RCWSA) agreement of July 1986. This approach uses COCSI data but does not when it is considered that the COCSI data is unreliable, that is, out of date due to changes in transaction mix. This approach requires union agreement but if none is reached a COCSI inspection may be necessary.

The main factors taken into account in a COCSI inspection are the number of customers and the average time taken to serve them and it proceeds as follows:

a. General discussion with the local staff representatives - to explain the objectives of the overhaul.
b. Initial survey - an examination of the environment and organisation.
c. Inspection of counter - measurement of the number of customers and customer times including an analysis of time spent away from serving positions and details of 'other work'. The count of customers include an assessment of the non-customer element, i.e. people visiting the office but not requiring service.
d. Calculation of staffing requirements - the facts gathered are analysed and the required number of staff hours calculated.
e. Formulation of duties - work schedules are produced and the staff establishment is calculated.

At an individual office the number of staff hours required for serving at that counter, i.e. counter cover (CC) is calculated as follows.

During a representative week a sample of customer times is taken. An observer is positioned on the official side of the counter, observes, starting from one end of the counter and working through to the other end, five consecutive customers per clerk on a continuous basis throughout the day. The study is specific to local conditions and the timings are of customers so no direct use is made of national average counter transaction times compiled by Headquarters timing staff. The 'total observation time' is divided by the 'total number of observations' to give an average time (in seconds) per customer, the 'customer transaction time' (CT).

CC is then calculated in accordance with the following formula

$$CC = \frac{\text{Customer arrivals per hour (CA)} \times CT \times 100}{3{,}600 \times \text{occupancy level (X)}}$$

The occupancy level of staff is deemed to be 70 per cent. This means that the total time spent by counter clerks in face-to-face contact with the public is scheduled to be no more than 70 per cent. The remaining time constituting the balance of 30 per cent allows for advanced accounting and summarisation work, stock replenishment and gaps between customers. Counters uses COCSI as a procedure for seeking staff reductions. They will continue to use it (subject to certain improvements) the basis on which District Managers negotiate staff changes.

Measures of relative technical efficiency like DEA can help identify those DMUs (Crown Offices) whose technical efficiency is low due to too much labour being used. Whether this information can actually be utilised is another issue. At present the COCSI procedure has to be undertaken before each one begins resulting in a time-consuming process. After calculation of the required number of staff has been performed, any changes to existing staff numbers have to be negotiated with the trade unions. This is something that does not bode well for a method which requires staff numbers to be adjusted flexibly to achieve the same technical efficiency relative to other DMUs.

Something that is strange because it would be thought that they were complementary is that there is no connection between measuring performance (BTH calculation) and calculating the number of staff hours needed (COCSI). The two calculations are made separately, the key to this being that an average time is measured for a counter clerk to perform a task. In the case of BTHs this is based on the time it takes to perform a transaction and is measured nationally. For a COCSI a calculation is made based on the time it takes to deal with a customer and is specific to a local branch. So no use is made of national performance figures to calculate the labour inputs. Differences in performance are thus allowed to exist between Crown Offices and staffing adjustments will only reflect this. More use should be made of national performance figures to calculate the labour inputs needed. To this end COCSIs should be phased out and replaced by a method that does this. DEA can help in this.

3.6 Conclusion.

3.6.1 Defining how DEA can be applied to Counters business data.

Post Office Counters lends itself to a Data Envelopment Analysis study of relative technical efficiency. Before it is explained why, and which DMUs were chosen for the analysis, it is necessary to explain why the group of post offices has to be limited immediately.

The subject DMUs to be used in the analysis will be the 1301 Crown offices in preference to the sub post-offices, this is because:

a. the Crown offices conduct a majority of the counters business.

b. 21,000 DMUs is an unmanageable amount to analyse for a study of this scale (even 1301 might have been too many to handle computationally and because of this a solution could have been to sample them, or just look at a restricted number of territories and districts.)

c. there is very little point in evaluating the efficiency of the sub post-offices because they operate on what is virtually a franchise basis anyway. They are beyond the control of the Post Office in that they determine their own staffing levels and quality of service. Crown Post offices are staffed by Post office employees. Staffing and procedure can be more centrally controlled and adjusted in the light of an efficiency study. An efficiency study can contribute to the process of measurement and control of the service.

The Crown Post Offices will be the subject of this study and are suitable for the following reasons. They are very homogenous but it is believed that wide variations in efficiency will still exist. The CPOs are homogenous in the sense that they perform the same tasks, under similar 'market' conditions and the inputs and outputs characterising the performance of all inputs are the same except for differences in intensity and magnitude. The size of the comparison group of DMUs is self-defining in that it will consist of all CPOs for which there is usable data. This is a large comparison set so there should be clear discrimination and allows a large number of inputs and outputs. But, the larger are the number of units, the lower the homogeneity within the set. The possibility that this increases the likelihood that results may be affected by some exogenous factors will be discussed in Chapter 5, and a possible solution suggested. The time period over which the CPOs are examined should exclude unusual periods like the summer holidays and Christmas, or should ideally be over a full year. Going too far back with the comparison would not be useful and would be unrepresentative if efficiency was changing. If there is insufficient data, or the results over time are of interest, then the data can be disaggregated to create more DMUs. This possibility is examined in a different analytical context in Chapter 7.

Given that Counters CPOs are suitable for an efficiency study using DEA it is necessary to define what data on input/output factors is needed. Due to the simplicity of the transactions process, very few variables need to be included in a DEA analysis of POC. Of the possible factors discussed earlier, very few are relevant to the POC context, due to this, judgemental screening of those remaining is not a difficult task. They were reduced to a small number, bearing in mind that they should contribute directly to the objectives of the POC and are each unique. Due to Counters own internal evaluation of efficiency there are large amounts of data on inputs and outputs and the defined data is available. The main factors are outlined below.

a. **Inputs:** Inputs into the counter service include capital and labour but Labour will be used in the analysis as the main input. This is because labour costs represent over two-thirds of postal business expenditure (as against 80% in the Deprins, Simar and Tulkens study) and the complexity of deciding how much capital is used by each individual Crown Office is too great. Data relating to branch size of any kind is not available. Using only labour as the main input is a limitation but there is no alternative.

   Labour inputs consist of the number of staff hours worked. It is not possible to divide this up into counter staff hours and back-office staff hours because back-office hours are not quantified. The only data available is counter staff hours, but these are the most important because they are the inputs into the transactions process.

b. **Outputs:** These will consist of the number of transactions undertaken as well as a quality dimension. As has already been stated Counters carries out 190 different types of transactions. It would be impracticable to compute efficiency on the basis of this number of different transactions and would give undue importance to transactions that are rarely used. Thus, it is best to concentrate on aggregates of the major categories. The innumerable other transactions are amalgamated into a category of other 'services'.

Quality of service it was concluded would be deemed to mean quality of service for customer arrivals measured by speed of service. The statistics on this produced by counters that could have been used in the study are for:

a. the average waiting time in seconds
b. percentage of customers waiting five minutes or more
c. percentage of customers waiting three minutes or more
d. percentage of customers waiting one minute or less.

The DMUs and inputs and outputs actually used in the DEA of POC were in the end defined by the data made available to me, though they correspond to the desired factors outlined above. These will be explained in more detail in Chapter 5.

3.6.2 How is DEA useful in evaluating Post Office technical efficiency?

i. Information for the management and control of individual DMUs.

DEA can help with the process of management and control of Crown post offices. Crown offices are relatively small units with five to six staff on average within a range of three to 30 staff. Many Crown offices are similar in size to sub-offices which are managed by Sub-Postmasters together with their private businesses. A Sub-Postmaster runs his Counters business without the detailed information which Branch Managers have. On the other hand he must balance the need to satisfy his customers against his profits. If he does not succeed his business will fail.

Area and Branch Managers face no pressure of this kind. The statistics they receive are a poor proxy for direct commercial pressure. The Branch Managers job of scheduling staff to match the pattern of customer arrivals is complicated. The Sub-Postmasters decision is a decentralised one and staffing levels will be much more tightly controlled. In a Crown Post Office mistakes in matching the pattern of staffing to customer arrivals will result in a reduction in quality (speed of service) if there are not enough staff, or a reduction in the number of transactions carried out per staff member if there are too many staff. This inefficiency in individual DMUs can be revealed by DEA.

The necessary organisational adjustments can then be made in those offices that are identified as being inefficient. A management model and the means by which efficiency can be improved using DEA are presented in Chapter 4.

ii. Overall Technical Efficiency.

Conclusions can be drawn about the overall efficiency of the Post Office Counters service from the results of the DEA analysis or on a smaller scale, the efficiency of a region. There are a number of ways of drawing conclusions about overall efficiency from the spread of technical efficiency results of all DMUs (though deciding whether the conclusion is favourable or not is still somewhat subjective.) These were suggested by Deprins, Simar and Tulkens:

a. the average of the technical efficiency measures
b. the median of the technical efficiency measures
c. evaluating the shape of a frequency distribution of efficiency measures, taking into account the number of extremely inefficient observations and the spread of heavily inefficient observations.

Though as has been pointed out on a previous occasion the block of DMUs could be deemed to be efficient overall but in reality there could be structural inefficiency. The DMUs are only efficient relative to the most efficient DMU and this may by casual observation be deemed to be inefficient.

Chapter 4.
The Use of DEA as a Performance Measure to Aid Management Control and its application to Post Office Counters.

4.1 Introduction.

The aim of this chapter is to evaluate whether DEA can be used as a performance measure in a management control system in the public sector. It is also intended to examine how it can be used to improve performance, and its relevance to Post Office Counters. Section 4.2 outlines the management structure and objectives of public sector organisations and describes the traditional model of management. This control model is usually used in the private sector and is gaining popularity in the public sector. The model incorporates profit as a performance measure, this is obviously not always measurable or relevant in the public sector so section 4.3 establishes whether DEA is an adequate performance measure as compared to profit. Section 4.4.1. examines whether DEA can be used as a performance measure in a control system and 4.4.2 decides whether this is possible in the context Post Office Counters. As the purpose of a control model is to improve performance, section 4.5 shows how DEA can do this. Section 4.6 looks in detail at DEA and its use in Post Office Counters, how the existing control system works, the problems with it, the way that DEA is used at present and if it could be used more effectively to improve performance.

4.2 Management control.

4.2.1 Management control in the public sector.

Management control in its broadest context can be defined as the process by which management assures that an organisation carries out its objectives effectively and efficiently. The traditional management control model is the focus of attention, management control is executed in a repeating process of defining objectives, drawing up budgets and plans, performance is measured and this information is used to reassess the objective and measures needed to reach it. It is the performance measurement part of this process which has the most importance to this study because Data Envelopment Analysis can be used to fulfil this role. There are other types of process that are found in organisations and they need to be differentiated from management control. That is, ***operational control*** which is the process of assessing that specific tasks are carried out effectively and efficiently, and ***strategic planning*** which is the process of deciding on the goals[1] of the organisation and on the broad strategies that are used in deciding these goals.

An organisation should be both efficient and effective. ***Effectiveness*** is measured by the extent to which its outputs accomplish its goals. The more the outputs contribute to the goals, the more effective the unit is. ***Efficiency*** is measured by the relationship between inputs and outputs, that is, the ratio of outputs to inputs. An efficient unit is one which produces a given quantity of outputs with a minimum consumption of inputs (or vice-versa). Effectiveness is always related to an organisation's goals, efficiency, per se, is not related to goals. An efficient organisation is one which does what it does with the lowest consumption of resources, but if what it does is an inadequate contribution to the accomplishment of the organisational goals it is ineffective.

The dominant goal, or at least one of the major goals, of private sector organisations is earning profits. This is not usually the case in the public sector. As most descriptions of the management control process are framed around the primary objective of earning profits this can create some problems for management control in the public sector. As Anthony and Young (1984) argue 'the basic control concepts are the same in both profit-oriented and non-profit organisations, but because of the special characteristics of non-profit organisations, the application of these concepts differ in some important respects'. A problem that may be created by these differing characteristics is that it is possible that the responsibility structure may not coincide with the control structure. A public sector organisation may have an objective that transgresses the responsibility boundaries between other bodies, making control that much more difficult. Another problem with the application of these techniques of management control is the definition of objectives, and outputs and performance measures by which success can be quantified. This is obviously not a problem in the private sector because the objective of profit is also its own easily monitored performance measure, but in the public sector it can be a problem. Fortunately, neither of these factors are a problem in the Post Office Counters context.

4.2.2 Management structure and objectives of public sector organisations.

Public sector organisations can generally be distinguished as having hierarchical structures composed of responsibility centres, units, sections, departments and divisions. A ***responsibility centre*** is simply an organisation unit headed by a responsible manager, each centre having control of inputs and outputs. Except at the unit level, each responsibility centre consists of aggregations of smaller responsibility centres. The function of executive management is therefore, to plan, co-ordinate and control the work of these, more or less autonomous responsibility centres. Responsibility centres are classified according to the degree to which their inputs and outputs are measured in monetary terms. Most of the responsibility centres in the public sector, excluding trading organisations such as the nationalised industries can be described as ***expense centres***. That is, the management control system has only measured the expenses (or inputs) incurred by a responsibility centre rather than the monetary value of its outputs.

A responsibility centre exists to accomplish one or more purposes, these purposes are its ***objectives***. The main objective of private sector organisations is earning profits, decisions made by their managements are intended to increase (or at least maintain) profits. The public sector is primarily composed of non-profit making organisations. For such organisations, it is a matter of producing the best possible service given the limited resources available. Central government departments like the National Health Service depend directly on the exchequer for all their funds. If charges are made for certain services they are usually nominal and have little impact on the level of services offered. Also, so-called autonomous bodies are subject to a high degree of central government control. Local authorities receive by far the largest proportion of their funds via the rate support grant. Whilst they can raise revenue locally, primarily by levying rates, the extent of their powers is strictly monitored by central government. Even public sector business entities, such as the nationalised industries, find that even with a very dominant profit or financial target, account has to be taken of wider exogenous implications. Central government has from time to time, legislated on such matters as their pricing policies, external funding limits and trade with certain overseas countries.

In general the success of what a public sector organisation produces should be measured by the amount it contributes to the public welfare. The effectiveness of this objective can be deduced by examining its outputs. Measuring outputs for many public sector organisations is difficult because their outputs are qualitative rather than quantitative. This problem, has more recently received greater attention but while some areas are capable of providing meaningful output measures, it is proving difficult to develop appropriate measures in others. In addition, many non-profit organisations have multiple objectives served by several different outputs rather than a single measure.

4.2.3 A traditional model of the management control process.

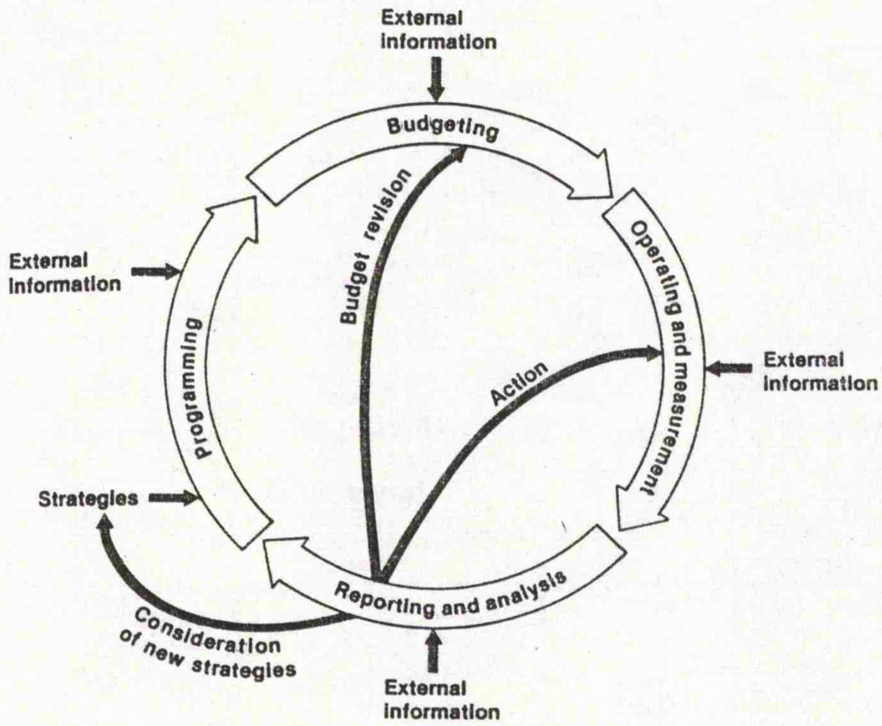
Anthony (1977) describes control as the ability 'to direct oneself and ones work', and as 'the process which ensures that goals are achieved or that plans are met'. A control *system* can be defined as 'a set of mechanisms which are designed to increase the probability that people will behave in ways that lead to the attainment of organisational objectives' (Flamholtz [1979]). The essential aspect of any control system is thus its effect on behaviour (Lawler and Rhode [1976]).

The main aspect of any organisational control system is the correspondence between the control strategy and a reward system (Merchant [1985]). Ouchi (1979) identifies three basic control strategies, output control, behaviour control and clan control. Output control involves the establishment of objectives and goals. Employees are then rewarded on the basis of performance relative to their goals. Behaviour control involves the measurement of processes or behaviours, rather than of outcomes. Employees are rewarded on the basis of their actions relative to a prescribed process, regardless of the outcome. Finally, clan control can be achieved by developing a social structure which minimises the differences between individual and organisational goals.

These control strategies need not always exist in pure form. Most control systems use a combination of all three control strategies. They also employ some form of output-based control as one aspect of their control strategy and performance measurement as a part of the associated incentive and reward system. Output-based control is the basis for the traditional management control paradigm.

Given that in the public sector organisations maximise a variety of objectives in the public interest, and that the main objective of private sector organisations is to maximise profit the management control process is the same. Anthony (1984) suggests that there are four components to a formal management control system. These phases are programming, budgetting, operating and measurement, and reporting and analysis. As indicated in Figure 4.2.1 each of these activities

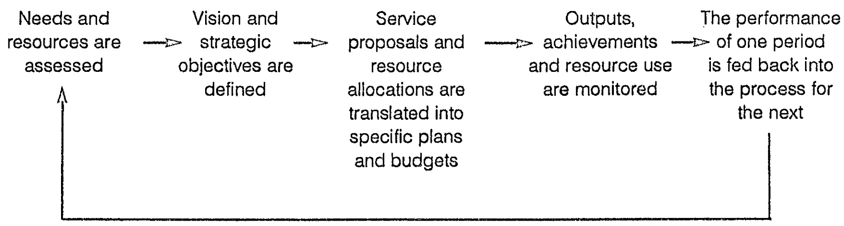
Figure 4.2.1. Phases of Management Control.
(Source: R.N. Anthony et al: Management Control Systems, [1980])



leads to the next. They recur in a regular cycle and together they constitute a closed loop. ***Programming*** is the process of deciding on the programmes that the organisation will undertake and the approximate amount of resources that are to be allocated to each programme. ***Budgetting*** is expressing a plan in quantitative terms that covers a specific period of time. In the ***operating and measurement*** phase records are kept of resources actually consumed and of revenues actually earned. The ***reporting and analysis*** phase is used as a basis for control. Reports are derived from an analysis that compares actual performance with planned performance and attempts to explain the difference. Based on the formal reports, and also on information received through informal communication channels, managers decide what if any action should be taken. If an action is deemed necessary a new plan is set forth in the budget and this leads to a new planning process.

This model of management control is the traditional management control model in the private sector. As a concept it has also found widespread popularity in the public sector. The National Audit Office Report (1986) on the 'Financial Management Initiative' shows clearly how the FMI was based on the principles of the classical control paradigm. Ministries were expected to identify clear objectives for their programmes, establish well-defined responsibilities for carrying out the tasks and measuring performance against those objectives. Similarly the Audit Commission's handbook on 'Performance Review in Local Government' (1986) uses the same traditional control loop. As can be seen in Figure 4.2.2 it follows the same path.

Figure 4.2.2. The traditional management control paradigm. (Source: Audit Commission: Performance Review in Local Government [1986])



Despite its mechanistic and rigid structure the traditional control paradigm of management cannot be easily dismissed. However, its limitations are acknowledged as the Audit Commission Report (1986) states, 'As a concept it is very simple, unfortunately implementation is not so easy. It depends on the acceptance for the need for the management of output, and the ability to control and plan inputs, both of which may have to be carried out in a highly charged political climate and with changing levels of resources'. In addition, the National Audit Office Report (1986) indicates that in implementing the FMI, 'Central government departments drew attention to difficulties created by imprecise, broad, policy objectives; the multiple objectives of programmes, and the difficulty of distinguishing the effects of a programme from other factors'. Despite problems with implementation the model is accepted.

There is another factor to be considered in deciding whether this model of management control is really suitable for the public sector. Criticism can be levelled at the traditional control loop because of its mechanistic approach. This model of control is often seen by organisation behaviour experts (eg. Kanter [1983] Morgan [1986]) as a direct descendent from Taylor's (1911) scientific management principles. According to these principles the concepts of controlling a machine are applied to the management of organisations. A task is set, the detailed systems and methodology are worked out, subordinates carry out the task in the way prescribed from above and output is measured against the task objective. This is often criticised as taking a dehumanised approach to organisational management such that subordinates have tasks heavily prescribed and are allowed little room for initiative. Hence, it is argued the organisation loses capacity both to observe and adopt to the changing environment, often seeing the need to change too late. While organisation behaviour experts may be right to warn of the consequences of using such a rigid model of management control, this view of a machine-like bureaucracy is but a caricature of the nature of most public sector service organisations. Obviously, some services are machine-like, those that produce routine outputs and a large part of lower level administrative tasks. But, much public service provision involves professional judgement and local initiative where it is difficult to measure outputs in terms of relating them directly to broad level objectives of whole programmes. Also, whilst those lower down in an organisation may be doing what they are told as implied by the scientific management model this is not the case for the thinking being done at the top. Thus the control loop may disguise much of what actually happens when it is applied. This does not mean that it is not useful, it is just that in practical application it is not actually used in such a rigid and stereotyped way as it appears when displayed in a diagram.

The classical model of management control was postulated with the private sector in mind, is it then applicable to the public sector? From the point of view of achieving control through a performance measure it is the final stages of measuring and reporting that are important. Organisations should be trying to achieve effectiveness and efficiency. In a profit-oriented organisation the amount of profit provides an overall performance measure of both effectiveness and efficiency. The level of profit can be reviewed after the reporting phase of management control and a decision made as to whether changes are necessary. Public sector organisations objective is not usually to maximise profit, as pointed out above they may have imprecise and varied objectives and many outputs (if they can be measured) and have no single performance measure. To achieve control outputs must be measurable so that actual performance can be compared to planned performance and an assessment of progress reached and new targets set. In this output-control strategy in the public sector even if objectives are clearly specified they will be associated with multiple outputs producing many performance measures and no single measure will exist. For this model of control to work in the public sector it would be advantageous for such a measure to exist. It has been posited that Data Envelopment Analysis could fulfil the role outlined. Before this idea is examined it is useful to look more closely at the role of profit as a measure in private sector organisations.

4.3 The need for a performance measure with the advantages of the profit measure in the public sector.

In the private sector profit has a number of advantages as a performance measure, representing an overall measure of both effectiveness and efficiency. An important objective in a profit-oriented organisation is to earn profits, and the amount of profits is therefore an important measure of effectiveness. Since profit is the difference, between revenue which is a measure of output, and expense, which is a measure of input, profit is also a measure of efficiency. It has been stated by Anthony (1980) that 'the absence of a satisfactory, single, overall measure of performance that is comparable to the profit measure is the most serious management control problem in a non-profit organisation'. In order to assess the significance of this statement it would be useful to find out why Anthony thinks profit is such an incomparable performance measure. Anthony lists the profit measure as having the following advantages:

a. It provides a ***single criterion*** that can be used in focusing the considerations involved in choosing among alternative proposed courses of action. Thus, profit provides a focus for decision-making.

b. It permits a ***quantitative analysis*** of those proposals in which benefits can be directly compared with costs. Such an analysis is possible when the objective is profitability, for profit is the difference between expense and revenue, and revenue is equated to costs.

c. It provides a single ***broad performance measure.*** Profitability provides a measure that incorporates a great many separate aspects of performance within it. It reflects the efficient use and minimisation of costs for a range of inputs and the maximisation of sales to generate revenue. Profitability provides managers with a current, frequent, easily understood signal as to how well they are doing and provides others with an objective basis for judging managers performance. The best manager is not the one who generates the most sales volume, or the one who uses labour most efficiently or the one who makes the best use of capital..

Rather the best manager is the one who does best on a combination of all these activities, and profitably incorporates these separate elements.

d. It facilitates ***decentralisation***. Because the goal is clearly understood and the performance of many individual managers can be measured in terms of their contribution to that goal decisions can be delegated to lower levels in an organisation. The management control device associated with such delegation is the *profit centre*, which is a division or other operating unit whose manager is responsible for both revenue and expenses.

e. It permits ***comparisons of unlike units***, performance can be compared amongst responsibility centres that are performing dissimilar functions. Assuming that the accounting rules used to measure profits are similar, and that the amount of assets employed is properly taken into account in measuring profitability, then the performance of a department store can be compared with the performance of a paper mill. Profitability therefore not only provides a way of combining heterogeneous elements of performance within company, it also provides a way of making valid comparisons among organisations that have the same objective, even though the size, technology, products and markets of the companies are quite different from one another.

In the public sector DEA can provide a single measure of efficiency (effectiveness will be examined later) and fulfil the role that profit does in the private sector. It can also equal the advantages that profit has. Examining the benefits of the profit measure point by point and comparing them to see if DEA can provide the same advantages, DEA does quite well. But, it has to be remembered that its usefullness lies in the public sector. DEA is also, like profit, a single criterion measure but it could not provide help as an aid to investment decision-making because it tell us more about efficiency than effectiveness. Also, in the private sector profit is the objective and in the public sector it might well be social benefit. A cost/benefit analysis might be a more useful measuring tool in the public sector.

DEA can provide a quantitative analysis but not one in which benefits can be directly compared to costs, so it is not that useful for the same reasons as it is not an aid to investment decision-making. It cannot quantify benefits and costs in the public sector as these will not be in monetary terms and need to be the subject of a social cost/benefit analysis.

DEA does provide a broad performance measure. It is able to handle multiple inputs and conflicting multiple outcome measures and environmental factors outside the control of the organisation being evaluated. The measure can be evaluated by managers and the people who judge their performance and provides a frequent and simple signal. The problems with DEA as a broad performance measure will be examined later.

If objectives are clearly defined then DEA can be an aid to decentralisation. Each responsibility centre will have a manager whose individual performance can be measured relative to other centres. As the centres performance can be measured it is safe to delegate decisions on resource allocation from management to them.

The last stated advantage of the profit measure, the ability to compare unlike units, is probably where DEA is most useful. It too can compare the performance of heterogeneous units because it takes into account many different input/output mixes and environmental variables which are not controllable by management. DEA differs from the profit measure in that it cannot be used to compare units that are producing different products because it needs to have the same inputs and outputs for each unit. Though, as it is a relative measure it is good for comparing unlike units within the same organisation. This would be its main use in the public sector.

Whilst Anthony thinks that the profit measure is a very good measure in a great many situations and is better than any alternative, it does have several limitations:

a. A company may have **multiple objectives** of which profit is only one, and profit may not be a measure of all these objectives. Long-term survival must be an objective and means that potentially profitable opportunities may be forgone if they bear too much risk. Another profitable area may be investments that could be considered unethical or immoral by sections of society. Prestige and a good image may also be valued by the company and the people associated with it.

b. Profit does not take into account **social costs and benefits**, at best it measures the success of a company as an economic entity. The impact on society which a company has cannot be measured by the profit measure because they do not make up part of the companies expenses or revenue. The effect of say, the environmental cost to society of a companies pollution or the benefit to society of its training programmes have no immediate effect on profit.

c. The profit measure measures current performance and takes no account of the **long-term implications** for profit. Managers can make decisions that mean that current profits look good but they are detrimental to future profits. A reduction of investment in things that improve long-term performance like research and development, advertising and training, reduce expenses and increase profits but reduce the capacity to increase profits in future years. One of the important caveats in the operation of a management control system in a profit-oriented company is that undue emphasis must not be placed on measures of current profitability.

d. The profit measure provides an **inadequate basis of comparison** for companies to compare how they performed year on year and how they performed compared to other companies. This is because there is no way of assessing what the potential performance is and if this potential has been fulfilled.

e. A company measures and reports its profits in accordance with generally accepted accounting principles, but there is a wide **latitude in accounting** in measuring the profits of a given firm. This decreases the validity of comparisons among firms.

f. There is a problem with the **inadequacy of accounting** because it does not permit measurement in a way that conforms strictly to the economic facts. The most important difference between accounting and economics is that in accounting assets are carried on the books at their historical cost, whereas the real value of these resources at the time of consumption is their opportunity cost.

g. Given these limitations, profit can measure the performance of a whole company and its operating divisions. But, there is **inability to measure certain segments** because the measurement of profitability is not feasible for many responsibility centres. It is not feasible to measure the revenues that accrue to such responsibility centres as research and development, legal, personnel, accounting, finance, administration and other departments. Profit provides neither a focus for the analysis of proposals, nor a measure of performance.

DEA does not have some of the limitations of the profit measure. The problems with the profit measure in the private sector are not encountered by DEA in the public sector or they are not relevant. DEA is better than profit as a performance measure because it can assimilate multiple objectives. If an output can be assigned to an objective then it can be incorporated into DEA. However, the possible objectives mentioned above are difficult to quantify and so it may be difficult to do in practise.

Profit alone cannot quantify the true returns of a public sector investment, only a social cost/benefit analysis can quantify the true returns of a public sector investment. But, social costs and benefits can be incorporated into a DEA to compare the technical efficiency of a range of existing projects. They could be treated as outputs, though social cost would have to be inverted in the analysis.

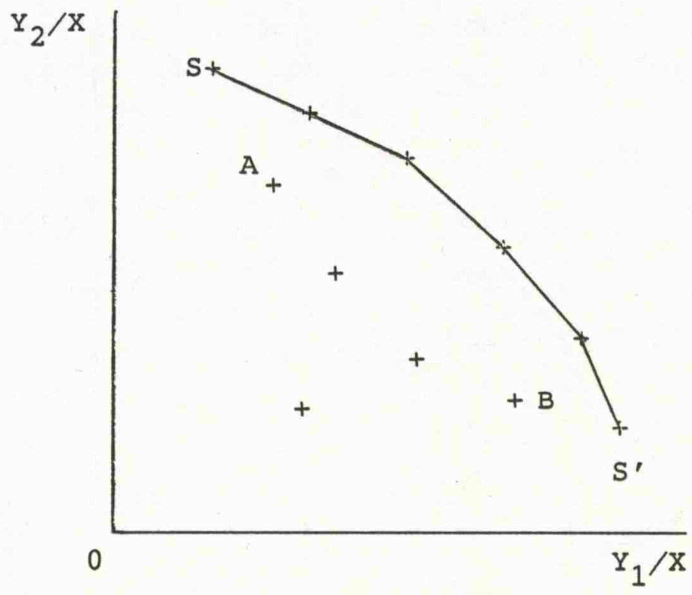
The long-term implications of current decisions cannot be taken into account by DEA because like profit it only examines the current situation.

DEA provides an adequate basis for comparison year on year and between companies, if they produce the same outputs. Comparison is obviously possible between companies which are treated as DMUs by DEA, as measurement takes place relative to each other. Year on year comparisons can be made by doing a DEA of DMUs in different years. Potential performance is determined by examining the slack variables and establishing targets.

The same problems that the profit measure has with conventions and adequacy of accounting are not problems with DEA because there are problems that are specific to the measurement of profit. But, DEA has the same drawbacks as profit in that comparability between firms might be limited, firms cannot be compared in the DEA context if they produce different outputs, and DEA does not have to take into account historical costs.

DEA cannot summarise the performance of a whole company as profit can, but it can measure performance in the segments that profit cannot. Departments that do not produce profit directly usually have some measurable outputs, and providing the same functions are conducted throughout a company relative performance can be measured. Nevertheless, in the private sector the profit measure has a number of advantages, representing an overall measure of both effectiveness and efficiency. In the public sector Anthony (1980) states that 'many non-profit organisations have multiple objectives, and there is no feasible way of combining the measures of the several outputs, each of which is intended to accomplish one of the objectives, into a single number that measures the overall effectiveness of the organisation'. There is some debate about the possible use of DEA as a measure of effectiveness in the public sector. Most of the literature deals with DEA as a measure of technical efficiency but Lewin and Minton (1986) have suggested DEA represents an approach to dealing with the more difficult issue of effectiveness and they relate this to organisational design. Epstein and Henderson (1987) point out a problem with this position. Two units that are equally inefficient may be operating in maximally different regions of their input and output spaces. Equally efficient does not necessarily mean equally effective, for example, in Figure

Figure 4.3.1.
Organisational effectiveness.



4.3.1 units A and B are both approximately equally inefficient, yet they are operating with very different output-mixes. The appropriate level of managerial concern, and the appropriate corrective action to achieve effectiveness for these two units may therefore be very different. Outputs have to be commensurate with objectives for DEA numbers to be a measure of the degree of effectiveness of the DMUs. But, if the best-practice unit is not effective which is quite possible due to variation in outputs and output-mix then the numbers are not a measure of effectiveness.

However, profit also embraces efficiency and there is a single criterion measure for this concept in the public sector. Data envelopment analysis can do for the public sector what profit does for the private sector in providing a single criterion of organisational efficiency.

4.4 Post Office Counters, control and DEA.

4.4.1 DEA and its use as a management tool to aid performance measurement in a control system.

As has been shown, DEA can be useful as a performance measure, the question that needs to be examined is whether it is suitable as a performance measure in a control system which would normally use profit as a measure. DEA is evaluated in this role by Epstein and Henderson (1987). Whilst not actually stating explicitly that the theory of control they are using is the traditional control model their description corresponds to it. They state that the control function is necessary to achieve organisational goals and objectives and that, 'Central to any control mechanism is the concept of feedback (Beer, 1966). One important source of feedback is provided by performance measures (Ouchi, 1979, Anthony 1965)'. This is the control loop of the classical management paradigm.

It was stated earlier, that the key aspect of a control system was its effect on behaviour. In order to have that desired effect on behaviour Vancil (1973) suggests that a control instrument must incorporate performance measures that satisfy two criteria, goal congruence and perceived fairness. In addition Epstein and Henderson also propose that the cost of computation, defined to include all computation and cognition involved in both the derivation and interpretation of the measure, constitutes a third major criterion. DEA will now be evaluated by these criteria on its suitability as a performance measure in an output-based control strategy.

a. Goal Congruence.

Goal congruence will exist in a control system to the extent that maximisation of the reported performance measure is consistent with that behaviour which best supports the overall goals of the organisation. A critical determinant of goal congruence is therefore the behaviour of the performance measure with respect to desired organisation outcomes. The measure must behave such that more preferable outcomes are reflected in higher values of the measure. But, organisational preferences are seldom unambiguous and well-behaved, it is therefore important that the performance measure appropriately reflects the existing level of preference ambiguity. DEA is able to do this because of its incorporation of multiple inputs and outputs without a priori weighting. Also, each variable is of equal value and has equal prior opportunity to influence reported efficiency.

However, DEA does have characteristics which might subvert goal congruence so that maximisation of the DEA measure may not support the overall goals of the organisation. The DEA measure is calculated as a ratio of weighted outputs to weighted inputs. In order for a ratio to be meaningful and well-behaved both the numerator and denominator must be expressed in ratio data. Non-ratio scale data in a DEA model can significantly distort the resulting efficiency measures. For example, the result of adding an intercept, of changing the zero point, or otherwise transforming the scale of any variable will normally change the reported efficiency in DEA model. DMU managers may strive to maximise a performance measure, the behaviour of which is heavily influenced by the scale used to measure one or more of the variables. This would not result in goal congruence.

Another characteristic of DEA which might affect goal congruence is the condition that all variables in the model be of roughly equal importance. This is because DEA allows all variables equal potential opportunity to influence reported efficiency. A DMU which produces a large amount of a relatively unimportant output or which uses very little of an unimportant input, would be rated as efficient on the basis of its performance with respect to an unimportant variable. The tendency of managers to try and maximise reported performance measures is well documented (eg.Prakash and Rappaport [1977]). It would be expected then that if DEA is used they will try to find the best way onto the frontier. A way of doing this might be to reduce an input to zero and become efficient in that dimension. Another effect of maximising the DEA measure might be to encourage too much attention to be paid to relatively unimportant variables. So dysfunctional incentives might be created by the use of DEA as a performance measure to achieve control.

b. Perceived fairness.

The other characteristic that DEA must have as a control instrument is ***perceived fairness.*** Perceived fairness reflects the attitudes and beliefs of the measured individuals or agents, with respect to the fairness of the measure. Perceived fairness is achieved through four criterion, appropriate standards of comparison and measures which are complete, objective and controllable by agents.

Measures become meaningful when compared to an ***appropriate standard.*** *Thus they should reflect existing constraints and circumstances. A way of achieving appropriateness is to use standards which are realistically attainable in the current time period.* ***Completeness*** must exist to the extent that all relevant variables are included in the measure. Measures should be ***controllable*** by the agent and relevant variables included so that variation in the measure should be attributable to decisions and behaviours of the agent. Variations in the measure which are due to factors beyond the agents control like random noise, should therefore be minimised. ***Objectivity*** should exist to the extent that the behaviour of the measure reflects actual observed outcomes rather than arbritrary decisions in the measurement process based on attitudinal or judgemental assessments.

How does DEA meet each of these criteria to achieve perceived fairness as part of the control strategy. There is an ***appropriate*** standard in that DMUs are compared to a hypothetical best-practice unit. But, there is a problem here because DMUs can only be compared if there is constant linear substitutability between any pair of inputs or any pair of outputs. That is, if the hypothetical comparison unit for any inefficient, enveloped DMU, is formed as a convex combination of a set of observed DMUs, the implicit assumption is that any convex combination of any set of efficient units is attainable, and is therefore an appropriate comparison basis for the units enveloped by that set. But, if certain pairs of inputs or of outputs are substitutable only in discrete 'chunks' then a unit may be unfairly rated as inefficient as a result of being compared to a hypothetical comparison unit which is in fact not realistically attainable. This occurrence was reported by Bitran and Valor-Sebatier (1985) in a study of 160 hospitals he found that DEA tended to underestimate true efficiencies due to its failure to recognise the existence of 'lumpy' outputs. So to achieve perceived fairness continuous substitutability is necessary.

Another condition for perceived fairness to be achieved is **constant returns to scale**. The set of DMUs which are selected to form the reference set depends upon only the mixes of inputs and outputs used by the DMUs and not on the relative scales of operation. The best-practice DMUs and the efficiency ratings are therefore determined without regard for the scale of operation. But, under conditions of increasing returns to scale larger units will tend to be rated as more efficient and if they are best-practice this comparison is unfair.

DEA is compatible with perceived fairness in that it is ***complete***, all measurable inputs and outputs can be included. This counters the complaint by DMU managers about standard performance measures that important variables, like uncontrollable environmental, ones were not taken into account by the measure.

Another criterion to be met to achieve perceived fairness is ***controllability***. That is, the ability of the agent to affect the performance measure through control of included variables. This is achievable by DEA because the measures can be adjusted for non-controllable variables, therefore avoiding penalising or rewarding managers for circumstances which are beyond their control. However, there are two areas of concern relating to this criterion. The complexity of the DEA measure could be such that the connection between changes in the values of the measures and changes in reported efficiency is not always apparent. If agents do not understand the behaviour of the measure, they cannot control it effectively. The other area of concern relating to controllability results from the non-robustness of the measure with respect to outliers or measurement error. This undermines its credibility because a DMU may be given a grossly unfair performance rating on the basis of a single outlier.

The final aspect of perceived fairness is ***objectivity***. The DEA efficiency measure depends upon the selection, definition and measurement of the variables in the model. If these are given then DEA is objective in that the method of weighting is mechanistic and replicable. There are two concerns here. Unenveloped units do not have a full reference set, their hypothetical comparison unit is therefore constructed somewhat arbitrarily, Secondly, there is a problem regarding the scaling of data. Sometimes there may be important variables which do not lend themselves to representation by ratio scale data. For instance, the variable 'information technology' cannot be meaningfully represented on a ratio scale. It is therefore necessary to adopt an arbitrary scale with which to measure the variable. This arbitrarily chosen scale will favour some DMUs and penalise others, thus compromising the criterion of objectivity.

c. Low computational cost.

The final characteristic that a performance measure in a control system should have, suggested by Epstein and

Henderson is that the cost of computation should be low in ***derivation*** and ***interpretation***. DEA, by reducing a large number of variables to a single measure of relative efficiency, provides a performance matrix that is relatively easy to interpret and provides an indication on the level of improvement possible. The existence of a hypothetical best-practice unit which is a convex combination of observed DMUs is also helpful because managers can identify the observed DMUs which make up the best-practice frontier. However, whilst the interpretation of the measure is straightforward and low cost its derivation may not be. The cost of computation is relatively high because of the complexity of the DEA algorithm, the need for data collection and transformation, and report formatting.

4.4.2 Can DEA be used as a performance measure in a control system in the context of Post Office Counters.

DEA may be an appropriate performance measurement instrument under the conditions outlined. However, violation of these conditions may seriously compromise the congruence and fairness of the resulting performance measure. The appropriateness of DEA as a performance measure will vary according to its application. It will now be examined whether the conditions hold when DEA is applied to Post Office Counters data.

a. Goal congruence.

The scale of the data used in the DEA analysis should not present any problems. The labour input is expressed in terms of hours worked by staff and the different types of traffic in terms of volume. The quality variable has been transformed so that more preferable outcomes are reflected in higher values of the measure. The variable, average waiting time in seconds, was inverted by subtracting from the maximum figure given. This now means that the number is a measure of how long each customer doesn't have to wait in seconds out of a maximum amount of time it is possible to wait. This remains useful as it is still expressed in a ratio scale.

Another area in which goal congruence may be compromised is if all variables in the model are not of roughly equal importance. This is the case in the Post Office Counters instance where the labour input and quality variable are more important than the transactions outputs. Fortunately in this case there is not a problem because the different types of counter transactions which are classed as outputs cannot be increased because they are non-controllable and dependent upon the level of demand. The relative importance of two variables does give scope for dysfunctional activity to take place. It is not unknown for a Branch Manager to manipulate average waiting time figures. This could take place at the expense of a more efficient use of labour inputs. Though it could be said that this problem may arise in any performance measurement system.

b. Perceived fairness.

A performance measure needs to appear to be fair to work in a given situation and this has a number of dimensions in the POC context. The criteria of an appropriate standard of comparison would appear to be met. There is constant linear substitutability for inputs and outputs. Outputs are in discrete units though due to their large numbers there is not a problem because a change of one unit is very small. Inputs come in discrete 'chunks', each worker contributes a block of hours worked. But these can be made quite small if part-time workers are employed and the number of hours worked in a CPO is finely adjustable if overtime is taken into account.

There could be a fairness problem if constant returns to scale does not exist because different DMUs will work at different scales of operation. But, this will not be a problem in my study because the VRS program is used which just measures technical efficiency.

Completeness of measurement could be a problem because not all inputs and outputs can be included (eg. capital). But, it is not a problem to the extent that labour is the key input and our focus of attention. Other possible inputs and outputs to the transactions process are discussed in Chapter 5.

Controllability should not really be an issue in the POC context. Managers do have the power to affect performance because they do not have control over the key variables of labour inputs and quality output. This does not seem to have been subverted by the complexity of the technique as managers are using it at present. Non-robustness, as was noted, is a problem, but distortion by outliers can be overcome if one is aware of the problem. For instance, a North London CPO is rated as being technically efficient due to its very large bulk sales of stamps

The objectivity of the measure should not be at issue when applied to CPOs. Unenveloped units, though having an arbitrarily constructed reference set, will have their peers represented in their local area if examination is confined to the Area level. Scaling of the data is not a problem because there are no qualitative inputs and outputs. Waiting times could be qualitative but if a set of criteria is laid down it is measurable.

c. Low computational cost

Computational cost should be low if examination is conducted at Area level. The derivation of the results, if conducted in this way, will be straightforward because the data collection, computation and formatting of results is only on a small scale. Interpretation should be a simple affair because of the low number of controllable variables, only labour and AWT are important to the results.

4.5 How can improvements in performance be made using DEA in a control system.

DEA can be used in a control system as a performance measure. Actual performance as measured by DEA can be compared to planned performance and steps taken to improve it at the next phases of programming and budgetting if it is not acceptable. As well as being a performance measure that provides technical efficiency ratings, DEA can actually help to improve performance through the information that it yields. This can be done in three ways. Firstly, efficient units can be identified and the degree by which inefficient units are inefficient, established. Secondly, DEA can show an inefficient DMU its best-practice peer group of efficient DMUs. It has been argued by several authors (Bowlin [1987], Charnes Cooper and Rhodes [1981], Dyson, Foster and Thanassoulis [1987]) that an inefficient organisation should make comparisons with best-practice in order to extract and transfer relatively better managerial procedures to improve its performance.

The third way that DEA can help to improve performance is by quantifying for an inefficient unit, the values of inputs and outputs which, in principle, the unit ought to be able to achieve. That is, the reduction in inputs or increase in outputs that is theoretically possible relative to other DMUs. These can be used as targets of achievement for the unit concerned[2]. However, the use of such targets may not always be feasible in practise, or inputs and/or outputs may not be under the units control. The adjustments to inputs and outputs suggested by DEA are not necessarily the only ones which would make the corresponding unit obtain a relative efficiency of 1. In general, an infinite number of such adjustments are possible. But, a DMU does not need to alter its relative mixes of inputs and outputs, only their individual levels. Also under the proposed strategy, no input needs to be increased or output decreased.

Whilst DEA is able to suggest targets for improvement of technical efficiency ratings it is not apparent how they are to be achieved. As suggested, it can be useful for inefficient DMUs to look at the operating practices of efficient units that are in their peer group. However, Dyson, Foster and Thanassoulis (1987) suggest that before units identified by DEA as efficient can be used as examples of good operating practices or in setting performance targets, they must be investigated further to gain a better insight into their performance. The problem is that inefficient units are quite clearly inefficient, despite the fact that the weights on the inputs and outputs are selected to the best advantage of those units. Whereas, the result is not so strong for identifying a relatively efficient unit. This is because some units are relatively efficient only by placing all their weighting on a very small subset of the inputs and outputs. It is possible, therefore, for a unit to appear efficient simply because it has an unusual pattern of inputs and outputs, this allows it to use flexibility on choice of weights to distinguish itself from the others. Given this, DFT think that the following points should be examined before using a technically efficient DMU as a guide to best-practice:

a. Which aspects of a units performance contribute to its efficiency rating?

DEA allows each unit to select the weighting structure for the inputs and outputs which would make the unit appear at its most efficient in comparison to the other units. This could mean that a unit assigns such low weight to certain inputs and outputs as to virtually ignore them for assessing its efficiency. Thus, an efficient DEA unit could be efficient only in those of its operations which generate the outputs or utilise the inputs which were actually taken into account in determining the unit's efficiency rating. An inspection of the optimal set of weights for a unit would reveal which of its inputs and outputs contribute to its efficiency rating.

However, the relative importance of a units inputs or outputs in determining its efficiency rating cannot be ascertained by a mere inspection of its optimal weights. A larger weight does not necessarily mean that the unit produces the corresponding output or utilises the corresponding unit more efficiently than outputs and inputs with lower weights. This is because the magnitude of each weight is also dependent on the scale of measurement of the corresponding input or output. A clearer picture of the relative importance of inputs and outputs to a units efficiency rating is obtained by an examination of its *virtual* inputs and outputs. The virtual output(input) attributable to a given output(input) is the product of that output(input) and its corresponding weight. The virtual inputs and outputs attributable to each input and output show exactly how the efficiency rating of the corresponding unit is derived. They appear as a proportion of how much the input or output has contributed to the efficiency rating.

b. Does the unit, through examination of its efficiency ratings show well-rounded performance?

An efficient unit has well-rounded performance if all aspects of its performance are taken into account rather than just a small subset of them. Otherwise, its relative efficiency may be simply reflecting an uncommon output profile in one or more of the outputs. But, it should be remembered that in general there exist alternative weighting structures that give a unit its maximum efficiency rating. Some of these weighting structures may show that the unit has a better rounded performance than the initial weighting structure might have implied. A DMU that shows well-rounded performance, is a better candidate for examination of its organisational practices by an inefficient DMU, than one in which the technical efficiency rating is dominant in a single dimension.

c. On which aspects of performance does an efficient unit do best?

Units having the bulk of their virtual input and/or output contributed by one input and/or output could be performing their operations relating to that input and/or output more efficiently than other units. But, the unit may be achieving a high score in one dimension, if that dimension is an input this high score may be achieved by starving the DMU of resources and reducing an aspect that may not be taken into account like quality. Or for an output, by devoting most of its resources to production of that output, rather than by performing all relevant operations efficiently. The virtual inputs/outputs can pinpoint these possibilities but further investigation outside the DEA context is necessary. A DMU that has a high virtual input/output in one dimension may be equalled or bettered in that respect by a DMU that has well-rounded performance. Again investigation outside the DEA context is necessary.

Having established that a DMU is technically efficient because it performs well in one area then it must be examined why this is the case. If it is misuse of resources, or just the case that it is particularly good in this respect, then this must also be established by examination.

d. Can a relatively efficient DMU improve its efficiency further?

Even if a DMU is established to have good all-round performance it should always be borne in mind that technical efficiency as measured by DEA is relative efficiency. This means that DMUs with an efficiency of 1 may not be efficient in absolute terms. There could well be room for improvement in the technically efficient DMUs themselves. Whilst we are looking at the efficient DMUs to see whether they are suitable for observation of their organisational practices, or using them to set performance targets to improve the technical efficiency of inefficient units, it must not be forgotten that closer examination of the means by which they improve their technical efficiency could be useful. Though, of course it makes sense to audit those DMUs that are definitely found to be relatively efficient and do not have well-rounded performance.

It has been established that DEA can be used to improve performance in the context of a control system. It can do this by showing which DMUs are efficient and inefficient and by helping to set performance targets for the inefficient units. These units can be shown improved organisational practices to meet these targets by examining the units that form their best-practice peer group. Before an efficient unit can be used in this way though, it must be established through the use of virtual inputs and outputs, which aspects of performance are worthy of further investigation for identifying good operating practices. Even relatively efficient units can be improved further and this can be done by examining whether aspects of their performance can be strengthened. It will be examined later whether Post Office Counters is utilising the full potential of DEA by using it to improve performance through the ways that have just been described.

4.6 Post Office Counters, control and DEA in practise.

4.6.1 How does the Control System Work at Present in Post Office Counters?

Counters control strategy will now be outlined and its correspondence to the classical control paradigm examined. The structure of Counters has been explained in the chapter 3, but in essence there are 4 territories containing 32 distrcts, each individual territory being presided over by a general manager. Each district is divided into a number of areas, typically four, under an Area Manager. The Area Manager is expected to manage an average of nine Crown offices and 120 sub-offices and is held responsible for the efficient operation of all Crown offices in the area. Each CPO is managed by a Branch Manager These different layers of management are the responsibility centres. Decision-making is decentralised because decisions on controlling resources have been delegated to Area Managers and Branch Managers and each are in control of responsibility centres.

At present the CPOs are run as discretionary expense centres and given appropriate budgets, because whilst costs are known, income cannot be accredited to individual offices on a regular basis. There are a number of technical reasons for this lack of accreditation but these are being dealt with. Errors in the reporting by offices of transaction volumes were a problem but there are improvements planned in checking procedures. Present pricing arrangements with clients presents a difficulty because charges are made for the use of a network on a non-differentiated basis. It is hoped that in the longer-term this information will become available and ultimately that CPOs will become profit centres. However, in the absence of a measure of revenue and the profit measure other means have to be found of measuring CPOs efficiency and effectiveness.

Area Managers and Branch Managers are responsible for managerial control but to different degrees. Area Managers set targets on quality and resource reduction for Branch Managers to meet. Branch Managers are also responsible for operational control. The Branch Manager has overall responsibility for the effective management of an office and is required to supervise all aspects of client and customer service, buildings and security, expenditure and accounting and staff scheduling. Most of the non-counters duties entail routine form processing and stock accounting by staff, with the Branch Manager being responsible for the control of expenditure and the final cash balancing. Most importantly, the Branch Manager is responsible for matching in the short-term the numbers of staff assigned to the counter against the numbers of customers arriving. The extent to which the Branch Manager succeeds in achieving a match will determine the quality of service experienced by the customers and in part the productivity achieved by staff. If Branch Managers are responsible for operational control then it is Area Managers that are primarily responsible for managerial control (though it will not always necessarily be the case that their roles are so clearly defined, Branch Managers may well be more active managers in some areas).

Area managers are the agents of managerial control, their goal is the efficient operation of Crown offices and this can be done by meeting the objectives of reducing costs and maximising the quality of service. These objectives are to be attained by requiring Branch Managers to meet targets within an agreed budget on staff costs, staff hours and a range of quality indicators. In 1986 the Post Office reached a flexibility agreement with the Communications Managers Association (CMA) 'Enhancing Managerial Effectiveness'. It related Branch Managers bonuses to performance against waiting time targets in individual offices. So there is a key feature of an output-based control system that employees are rewarded on the basis of their performance.

The standard quality of service performance of CPOs is set nationally. The sampled results on customer waiting times are used to provide a variety of analyses to assist the Area Manager to diagnose what action is required at a particular Crown office to adjust quality of service performance to the standard. They are also given guidance on the accuracy of each statistic so that management can avoid taking precipitate decisions. If the quality of service performance of an office is far short of the national target, the District Manager can authorise the Area Manager to monitor performance against an interim target. Any use of such localised targets must be notified to Counters headquarters. The Area Manager also enforces targets on staff hours and costs and will take remedial action if these are not met

It would seem then that Counters control system corresponds to the output-based traditional model. There is a programming phase instituted by headquarters. The objectives they define are translated into quantitative terms in the budgeting phase. Budgets for CPOs are negotiated by Managers and targets set for future cost reduction. Measurement then takes place and the Area Manager decides whether CPOs have met their targets and need action to be taken. Managers are rewarded with bonuses if their targets have been met.

4.6.2 What Problems do Post Office Counters have in Achieving Control?

There are some problems with the way Post Office Counters is exercising management control. The Branch Manager has to balance the need to minimise the use of staff against the requirement to maximise quality of service. To meet targets on staff-hours the Branch Manager has to match staff levels to predicted and unpredicted traffic fluctuations. There can be a number of predictable variations in traffic (eg. issue of vehicle licenses, payment of pensions and telephone bills etc.). The Branch Manager uses their local knowledge of these work peaks to adjust staffing to the fullest daily and hourly pattern of demand. The success of this system is dependent upon their good judgement in matching staff levels against these predictable traffic variations. Even so, the information available to Branch Managers in this case is limited. There is an urgent need to improve short-term customer traffic forecasting and utilise computer-aided staff scheduling. In addition, there are no standard methods of fine-tuning staff to traffic from day to day. Counters has acknowledged this, it recognises the need to identify best practices and build these into the training courses offered to Managers

To accommodate unexpected fluctuations in the number of customer arrivals Branch Managers are able to adjust the staff schedules at short notice. They can do this by making use of the Area Managers reserve staff, using part-time staff and adjusting the timing of back-office work etc. However, the quality of service monitor, for which the data collection is the Branch Managers responsibility, does not provide output frequently enough to support the Branch Managers attempts to exercise on-line control of staff. Counters says that it was never intended to do so but the Monopolies and Mergers Commission Report (1988) on Counters suggests that this is necessary. It proposes a supplementary monitoring process it calls the 'Q-clerk ratio'. This uses a ratio of customers waiting for service to the number of counter clerks serving. They claim that this ratio is easier to record than customer waiting time and lends itself more easily to an on-line control measure.

In addition to queuing data the Branch Manager also receives regular reports of office traffic, measured in BTHs. They can also monitor the number of customer arrivals. But, it was not until the budget year 1988/89 that Branch Managers were given budget targets for staff costs and staff hours against which to receive reductions. This can only have limited the management responsibility and control of Area and Branch Managers.

Another problem with the control process is the competence of Branch Managers to achieve management control. In October 1986 Counters appointed 150 Area Managers and some 500 new Branch and Assistant Managers. But, it inherited a number of Branch Managers who traditionally had not been encouraged to regard their role as one of active management. Statistics produced by Counters show that in the 1987 staff appraisal review, 48 Branch Managers were reported as showing 'significant weakness in performance' and 289 were regarded as 'generally acceptable' (which is the category below 'fully acceptable'). Both figures excluded cases where similar assessments were regarded as being due to inexperience. Thus 22% of Counters Branch offices are not well-managed. Counters has tried to solve this problem through early retirement where possible. Also, giving guidance and development to equip people to do their jobs as a precursor to progressively increasing their responsibilities. The issue of budget targets is seen as part of that process.

Although Counters control strategy has been that of the traditional output-control model there have been problems in ensuring its success and that objectives are met. The objective that is a proxy for quality of output is not taken frequently enough for on-line decisions to be taken on staffing levels. In addition there is no forecasting system for short-term traffic fluctuations or ways of fine-tuning staff to traffic levels. This can only mean that quality suffers or that too much labour is used. Budgetary targets and targets for staff-hours have only been recently introduced so control has been weak for a long time. The quality of Managers has been such as to affect the control process. A large number of managers are of low quality because they are part of the old ethos which did not require real management or they are recently appointed and inexperienced. So, it can be concluded then that due to its relatively short existence the control system is in a rudimentary state, but Counters is aware of most of the problems and is beginning to take some steps to correct them.

4.6.3 How is the DEA Performance Measure being used as part of a Control System in Post Office Counters?

DEA is a very recent introduction to the Counters control strategy. Each Area Manager has a simple DEA package they can use to examine the on average, 9 CPOs in their area. The key aspect of this application of DEA, is how it is used to improve performance. Area Managers can identify those CPOs who form a reference group, by comparing those in their area to ones that have a similar input/output mix. Once it is established which CPOs are similar and are best-practice the Area Manager must find out what organisational practices they are using that give them high technical efficiency. There are a number of ways this has been taking place:

a. Arranging visits
b. Talking directly to the relevant Area and Branch Managers
c. Sending out questionnaires.

This yields information that Area Managers can use to improve technical efficiency, it has been found that practical advice about the following aspects of organisation of staff are most useful. Comparisons have been made with other standard staff scheduling procedures, that is, the measurement of traffic and waiting times. It is possible to utilise staff more effectively through the skillful use of casuals and part-timers. Certain times of the week or month are busier than others due to certain transactions having to take place on a specific day. These may be used more skilfully in some areas than others and the information on how this is done can be useful.

Specialist positions, that is a counter that only deals with one type of transaction can have an impact on waiting times. Specialist positions are effective in this way if they separate out people who want to make transactions that take a particularly small or large amount of time. Less people are then kept waiting who only want to make a quick transaction. Another use for them is for transactions that peak at a certain time of the month, for instance, car taxation. These can be dealt with more quickly by a counter clerk if they are just dealing with one type of transaction. Therefore, specific information on how other managers use specialist positions is useful to Area Managers.

Asking other Area and Branch Managers of CPOs identified by DEA can help Area Managers to improve their efficiency. But, DEA can also help to define how much their inputs could be reduced or outputs could be increased and specify performance targets. As transactions are not controllable and are defined by demand then the only output that can be affected is the average waiting time. There is only one input and that is staff hours. At present, targets on staff hours and waiting times are set nationally. The waiting times that are acceptable are set with reference to customer preferences. Budgets are set annually between Districts and Headquarters. Post Office counters is about to embark on a programme of cost reduction, it will use DEA to set the performance targets to do this. Targets will be set on the basis of DEA examinations of the 4 territories and co-ordinated at headquarters level.

4.6 Conclusion.

In conclusion, DEA would make a good performance measure in an output-based management control strategy. In the respects necessary to be a relative performance measure in the public sector it has similar qualities as the profit measure. DEA satisfies the criterion necessary for a performance measure, that of goal congruence, perceived fairness and low computational cost. When applied to Post Office Counters DEA still maintains these characteristics.

POC has an existing strategy to minimise resources and increase quality. Additionally it has started to use DEA at the area level. The criteria set by DTF to use DEA as a performance measure, are the identification of inefficient units, using peer groups to exchange information on achieving efficiency and setting targets to achieve best-practice. It would seem then that on a theoretical basis Counters is using DEA correctly and to its full potential. Though it is still too early to provide information on its success the fact that it is being implemented correctly is encouraging.

Footnotes.

1. The term goals will be used to mean the broad, overall fairly timeless statements of an organisations aims, and the term objectives for the more specific statements of planned accomplishments within a specific time period. Goals are developed in the strategic planning process and objectives are used in the management control process.

Chapter 5.
Applying DEA to Post Office Counters Data.

5.1 Introduction.

In the previous chapter it was established that DEA could be used to improve the performance of an organisation by:

a. Identifying inefficient DMUs.
b. Setting targets for inputs and outputs to achieve technical efficiency for the inefficient DMUs.
c. Identifying peer groups whose management practices can be observed to achieve best-practice.

Dyson et al. (1987) suggest that as DMUs identified as efficient are so important to this process they should be examined further to test their robustness. This is because efficiency can be achieved on a small subset of inputs and outputs. Ways of doing this are examined in chapter 6.

This chapter will demonstrate on a real production process how the three stages above would take place and the difficulties of use and interpretation. Firstly, it will examine the overall efficiency of Counters. It will establish how inefficient it is, how much labour inputs have to be reduced and quality of service increased to achieve best-practice. This will also be done on a regional basis to yield more information on efficiency and help identify reasons for inefficiency. The technique of clustering will also be examined in the Counters context to establish its usefullness in better achieving the three steps outlined above. This involves re-estimating frontiers for the regions examined from the results for the single all-inclusive DEA analysis.

Section 1 outlines the specification of the DEA model. That is, which variables are chosen, which are deemed to be controllable and the assumptions made about the parameters to be used in the programme. In Section 2, the results are examined for the overall performance of Post Office Counters. The degree of technical efficiency (TE) is revealed and targets outlined for the amount labour inputs would have to be increased and average waiting times reduced to achieve relative technical efficiency across the whole of Counters. Also, the level of scale efficiency and its significance to Counters will be assesses. Section 3 examines the results further by dividing them up on a regional basis. This can reveal more about relative efficiency. Aspects of the results are explained and the possible reasons for relative and absolute inefficiency examined. Section 4 examines the effect on the results of clustering the CPOs into their regions and re-estimating a frontier for each one. This is done to ensure that all CPOs face the same set of operating conditions. Clustering by a common factor can exclude the effect of a variable, factors that affect efficiency will have similar values within regions and so be taken into account. The effect of clustering on efficiency ratings, peer groups and targets is examined and compared to the non-clustered regional results.

5.2 Construction of the DEA model for Post Office Counters.

5.2.1 The Post Office Counters Data Set.

As already described in a previous chapter the DMUs chosen for analysis by DEA are Crown Post Offices belonging to Post Office Counters Plc. The data set made available to me by Counters was for 1301 Crown Post Offices. The Offices cover England (including the London area), Wales, Scotland and Northern Ireland and the data covered a 13 week period from September-November 1989. This was a fairly representative period of time and excludes the Christmas period which could distort the results. Two observations had to be discarded due to data errors and also eighteen of the offices had provided no quality data that quarter and so were also discarded. Quality data was not collected due to situations where it was not possible to conduct proper measurement, for instance, office refurbishment. This means that 1281 offices can be analysed. Due to reasons of commercial confidentiality none of the Crown Post Offices can be named individually. Each office is thus identified by a number and they have simply been numbered from 1 to 1281. The data provided can be divided into regions, though these ceased to be used officially after the reorganisation of Counters in 1986. These regions can be numbered as follows, London 1-325, Eastern 326-420, Midlands 421-550, North-East 551-657, North-West 658-779, South-West 780-904, Wales and the Marches 905-986, Northern Ireland 987-1010, Scotland 1011-1169 and South-East 1170-1281.

5.2.2 Selecting Variables For Use in DEA: Specification of the model.

The variables chosen to be run by DEA are for inputs and outputs into the counters transaction process (these are displayed in Table 5.2.1). They correspond to the set of factors identified at the end of chapter 3 but are limited by the fact that the data used was defined largely by what was available. Due to the simplicity of the defined model it was not thought necessary to examine its veracity using regression or DEA based analysis.

Counters inputs are the amount of work done by counter clerks measured in hours. The figure takes in hours worked in overtime and also those worked by part-time and temporary workers. This is because it is the total amount of time spent on the counter that is important in this study. Branch Managers hours are also recorded and are available but they are not relevant, the amount of hours Branch Managers work has no direct impact on counter transactions.

The outputs of counters are the number of transactions performed. These are divided into nine categories and measured in BTHs and although they have drawbacks (see Chapter 3 for the problems in using BTHs), they can be considered in general an adequate proxy for the volume of transactions. Each of the transactions can be thought of as a different output because they take different amounts of counter transaction time and hence need different amounts of labour inputs. DEA has the advantage of being able to take into account these multiple outputs, whereas using regression techniques the outputs would have to be aggregated.

A quality variable is also included and is classed as an output, this is the average waiting time of a customer in seconds. Average waiting time has to be inverted because outputs must be positively correlated with inputs. The more labour inputs there are the less time people have to wait. To invert average waiting time it can either be made negative or subtracted from the highest AWT which is 289 seconds. The latter course was chosen because it would mean that the targets would be positive. The quality variable is now a measure of how long each customer doesn't have to wait in seconds out of a maximum possible amount of time it is possible to wait.

5.2.3 Assumptions about Post Office Counters Objectives.

There are a number of choices to be made within the DEA programme used about Post Office Counters objectives and production technology. The assumption that most suits Counters as an objective is that of input minimisation rather than output maximisation. The rationale for this is as follows. As a public sector organisation in the current political climate it is committed to produce efficiently by minimising its inputs with the outputs it has at present. An efficiency target has been set for the Post Office as a whole by the Department of Trade And Industry. It is set in terms of a percentage reduction in annual real unit costs (RUC). This is calculated by taking the total costs of the Counters business (exclusive of interest paid or received) and divided by the total volume of saleable output measured in BTHs.Though simple, this method does have the drawback

Table 5.2.1.

Variables used in a DEA of Post Office Counters.

Input Variables (Controllable)

a. Counter Clerk Serving Hours.

Output Variables (Controllable)

Quality of Service

a. Average waiting time (seconds).

Output Variables (Non-Controllable)

Traffic (in BTHs)

a. Stamps

b. Vehicle Licenses

c. Visitors Passports (Family and Single)

d. Child Benefit

e. Unemployment Benefit and Income Support.

f. Pensions

g. Giro Deposits.

h. Giro Withdrawals.

i. All other transactions.

---

that by using BTHs back-office support is not taken into account and the extent of this support varies with product mix, making different years incomparable if the mix changes. The RUC target is set for three years at a time, for the period 1986-7 to 1988-89 the target reduction was set at 3.7%.

The cost of Labour inputs to Counters has to be reduced and these are 30% (1987-88) of total costs. This may not seem to be a very significant proportion but 40.7% of Counters costs are for payments to Sub-Postmasters, virtually all of which will be used to pay for labour costs. As the costs of labour inputs have to be reduced then implicitly labour will have to be used more efficiently and shed if necessary. This is why input minimisation has to be the assumption that Counters must be subject to.

In general, in the public sector it is easier not to use output maximisation as an assumption. This is because outputs are notoriously difficult to measure, whereas inputs are quite easily quantifiable. Another reason is that in many public sector organisations the demand for their outputs is exogenous. In both cases input adjustments to improve efficiency are more preferable. In Post Office Counters the reason why output maximisation is not a feasible assumption is because of the latter. Counters does not have control of its output due to the public service nature of its business. Whatever demand arises has to be met at the prevailing prices. As each individual office has no control of its output the only possible action available is to adjust input requirements to traffic. This lack of control of output is confirmed by the fact that despite recent growth, Counters forecasts a longer-term decline in a number its core markets. This resulting primarily from an erosion of benefit and bill payment by competition from banks and building societies. Existing business is expected to decline by about 10 to 15 per cent over the next ten years. Therefore, an assumption of output maximisation is rejected, input minimisation must be assumed.

5.2.4 The DEA program.

As was explained in chapter 2 the Cubbin programme can assume the production technology to be that of constant returns to scale (CRS) or varying returns to scale (VRS). In the VRS case the production surface may take on increasing, constant and decreasing returns as appropriate. A priori it would be thought, just looking at the production process in terms of labour inputs, the technology in Counters case would be that of constant returns to scale. Extra units of input assuming they are homogenous will produce increases in work performed in the same proportion. But given that labour productivity varies this is unlikely for all offices. Thus, the more flexible VRS program will be used.

In the literature a distinction is made between controllable and non-controllable input and output variables (Banker and Morey [1986], Smullen [1989]). The Cubbin programme requires them to be classified as such. Non-controllable variables are those which, although important to DMU performance, cannot be affected by managers. Often an attempt is made to adjust for the impact of factors for which management does not have immediate control. These would normally include inputs and outputs denoted as being environmental variables. Only variables that can be affected by management are important in the sense that if they are controllable then they can be given target changes to enable them to achieve best-practice. In the Post Office Counters case the variables that are not controllable are the counter transactions. The branch manager has no control of these variables because they are exogenously determined by demand. He cannot affect demand for the services provided by a CPO in any way. The controllable variables are the labour inputs and the inverted AWT output. The Branch manager can obviously affect the former by varying the number of staff working on the Counter or influencing the competence and efficiency of the staff and changing institutional arrangements so that transactions are dealt with more quickly. Average waiting time can be influenced by providing the correct number of staff for the traffic taking place at the time. Thus, CPOs càn only influence their DEA rating by minimising their labour inputs and/or the average waiting time that customers have to wait. Both of these courses of action are commensurate with the objectives of Counters.

5.3 DEA results: Counters overall efficiency.

5.3.1 Results and conclusions about overall technical efficiency.

DEA was performed on the 1281 DMUs with the assumption of input minimisation and a varying returns to scale technology. The VRS results will be used throughout the study because the aspect of efficiency we are concerned with is pure technical efficiency (PTE). This is not to say that scale efficiency (SE) is of no significance and the importance of this to POC will be examined shortly.

Due to the large size of the dataset, just displaying the VRS results with the technical efficiency numbers will not be very meaningful. So, initially the results are displayed here in the form of a histogram, following the example of Deprins, Simar and Tulkens (1984). Figure 5.3.1 shows the histogram for the VRS production technology.

Conclusions about the overall technical efficiency of Post Office Counters can be drawn by looking at the shape of the histogram in Figure 5.3.1. It is apparent that there is a very wide variation in efficiency over the sample. There are a number of DMUs that are technically efficient and just below this level there are a very few DMUs who are not quite. Then, the majority of DMUs have efficiency which is neither very good nor very bad. It should be noted that the histogram is bimodal, and this is often the case. Obviously a peak will be created by the inefficient DMUs and it might be expected that the remaining efficiency scores will be normally distributed through the remainder of the CPOs (for further reading see Lovell, Walter and Wood [1989]).

Table 5.3.1.

Histogram of DEA technical efficiency numbers with variable returns to scale.

Col. 1. is the number of post offices in the class.
Col. 2. is the relative frequency within the class.
Col. 3. is the mid-points of the efficiency classes.
Each * represents 5 observations.

INPUT MINIMISATION. VARIABLE RETURNS TO SCALE.

1	2	3	
2	0.002	0.40	*
4	0.003	0.43	*
4	0.003	0.46	*
12	0.009	0.49	***
10	0.008	0.52	**
15	0.012	0.55	***
29	0.023	0.58	******
38	0.030	0.61	********
38	0.030	0.64	********
61	0.048	0.67	*************
83	0.065	0.70	*****************
81	0.063	0.73	*****************
114	0.089	0.76	***********************
99	0.077	0.79	********************
106	0.083	0.82	**********************
100	0.078	0.85	********************
95	0.074	0.88	*******************
83	0.065	0.91	*****************
56	0.044	0.94	************
54	0.042	0.97	***********
197	0.154	1.00	**

Statistics relating to DEA technical efficiency ratings under variable returns to Scale.

No.of DMUs	:	1281
No. technically efficient	:	177
Mean Efficiency	:	0.814
Median	:	0.693
Stan. Dev.	:	0.132

By examination of the results it was found that of the 1281 DMUs, 177 were technically efficient, 13.82 per cent of the total amount. Even given the size of the enteprise this does not seem very favourable. Whilst extreme cases of inefficiency are present, they are not that many in number. Taking a completely arbitrary definition of what an extremely inefficient DMU, the number of DMUs with a TE number of less than 0.5 (half the efficiency of best-practice) is 17 which is only 1.33 per cent of the sample. Having said that, the lowest TE of all is 0.386, a CPO whose efficiency is very low indeed. The problem truly lies in the fact that though there are few extremely inefficient DMUs there are a large number of heavily inefficient DMUs. There are relatively few DMUs which are almost efficient. This can be quantified by looking at approximately where the peak on the histogram is. It can be seen that between the TE scores of 0.655 and 0.955 lie 69% of the DMUs. So, although there are a few extremely inefficient DMUs and a significant number that are, or are very nearly technically efficient, there are a substantial number that are very run of the mill and could improve their performance.

This overall mediocrity is reinforced by looking at the mean of the TE numbers over the sample which was 0.814. It is a better measure of location than the median which is 0.693, because the ratings are clustered at two points towards the efficient end of the distribution. The implications of the fact that efficiency is only 81.4% of what is possible is that substantial changes in the use of resources have to be made. The numbers of transactions processed is exogenous so changes in labour inputs and the quality output are needed. The resource targets needed to achieve best-practice will be discussed in part 5.2.3.

5.3.2 Scale efficiency in Post Office Counters.

As was described in Chapter 3, other financial institutions have had their efficiency evaluated using DEA.

Of those studies, three of them, Rangen et al. (1988), Field (1990) and Drake and Weyman-Jones (1991), decomposed technical efficiency into pure technical efficiency and scale efficiency. The rationale for undertaking this is described in Banker (1984).

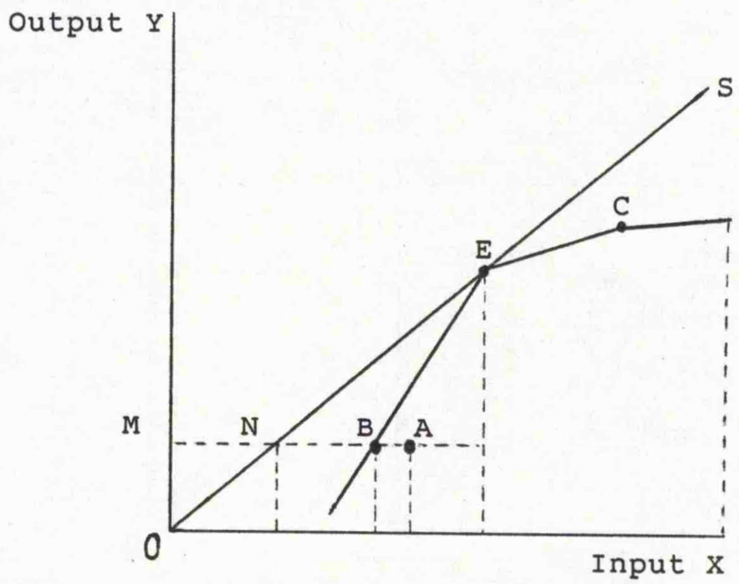
Before the results of measuring scale efficiency are presented and conclusions reached about their importance its method of determination is explained. Figure 5.3.1 illustrates the concepts of technical and scale efficiencies. The point A represents the DMU being evaluated. The choice of reference set and thus the efficiency of a DMU will depend upon whether we assume input minimisation or output maximisation. It was explained in an earlier section that input minimisation (minimising inputs with given outputs) is to be assumed, so efficiency is measured horizontally in relation to the input axis. If output maximisation was to be assumed then efficiency would be measured vertically in relation to the output axis.

Using the VRS program the production frontier generated will be BEC. The pure (input) technical efficiency of A is measured by the ratio MB/MA by comparing it with the point B on the efficient production frontier with the same scale size as A. Overall scale and technical efficiency is measured in relation to a CRS frontier 0S by the ratio MN/MA, comparing point A to the point N which reflects the average productivity attainable at the most productive scale size represented by the point E. Finally, the (input) scale efficiency of A is measured by the ratio MB/MA, so that the overall scale and technical efficiency MN/MA is equal to the product of the technical efficiency MB/MA and the scale efficiency MN/MB. It is apparent from Figure 5.3.1 that the aggregate scale and technical efficiency measure MN/MA is less than the pure (input) technical efficiency measure MB/MA. The relationship between the two efficiency measures holds also for the general case of multiple inputs and outputs. The measures are summarised below Figure 5.3.1.

Therefore to calculate the degree of overall scale in

Figure 5.3.1.

Technical and scale efficiency.



A represents the DMU being evaluated.
B represents a technically efficient reference point with the same (output) scale size.
E represents a technically and scale efficient reference point at the most productive scale size.

Technical and scale efficiency = MN/MA
Pure technical efficiency = MB/MA
Pure scale efficiency = MN/MB

POC a CRS version of DEA has to be performed on the data. The overall scale efficiency can be calculated thus

$$\text{Overall scale efficiency} = \frac{\text{Mean efficiency under CRS}}{\text{Mean efficiency under VRS}}$$

A CRS DEA was performed on the data and the results are summarised in a histogram in Figure 5.3.2. This represents the combined technical and scale efficiency of the Counters network. The mean technical and scale efficiency is 0.785, a low figure, but not very much lower than the PTE figure. 85 of the CPOs were deemed to be technically and scale efficient, or only 6.64% of the total sample.

The means of the CRS and VRS analysis, or the means of scale and technical efficiency and pure technical efficiency are now known. Thus the scale efficiency of the Counters network is

$$\frac{0.785}{0.814} = 0.964$$

A scale efficiency of 0.964 is very high and would seem to be satisfactory, adjusting the scale of operation to correct the loss of output would only yield an increase of 3.6%. The point has to be made that this conceals what is happening on an individual basis. Examining individual CPOs, only 169 or 13.19% are scale efficient. Only a small proportion of the CPOs are completely scale efficient, but, the high scale efficiency rating would indicate that most of the CPOs are operating at, or close to, their optimal scale. Given that the mean PTE is relatively low (see Table 5.3.3 for summary of efficiency ratings) we can be reassured that it was the correct decision to concentrate on pure technical

Table 5.3.2.

Histogram of DEA technical efficiency numbers with constant returns to scale.

Col. 1. is the number of post offices in the class.
Col. 2. is the relative frequency within the class.
Col. 3. is the mid-points of the efficiency classes.
Each * represents 5 observations.

INPUT MINIMISATION. CONSTANT RETURNS TO SCALE.

1	2	3	
3	0.002	0.40	*
4	0.003	0.43	*
4	0.003	0.46	*
13	0.010	0.49	***
14	0.011	0.52	***
22	0.017	0.55	*****
32	0.025	0.58	*******
47	0.037	0.61	**********
49	0.038	0.64	**********
85	0.066	0.67	*****************
93	0.073	0.70	*******************
94	0.073	0.73	*******************
135	0.105	0.76	***************************
104	0.081	0.79	*********************
122	0.095	0.82	*************************
100	0.078	0.85	********************
100	0.078	0.88	********************
76	0.059	0.91	****************
57	0.044	0.94	************
30	0.023	0.97	******
97	0.076	1.00	********************

Statistics relating to DEA technical efficiency ratings under constant returns to scale.

No.of DMUs	: 1281
No. technically efficient	: 85
Mean Efficiency	: 0.785
Median	: 0.693
Stan. Dev.	: 0.132

Table 5.3.4.

Histogram of DEA technical efficiency numbers with non-increasing returns to scale.

Col. 1. is the number of post offices in the class.
Col. 2. is the relative frequency within the class.
Col. 3. is the mid-points of the efficiency classes.
Each * represents 5 observations.

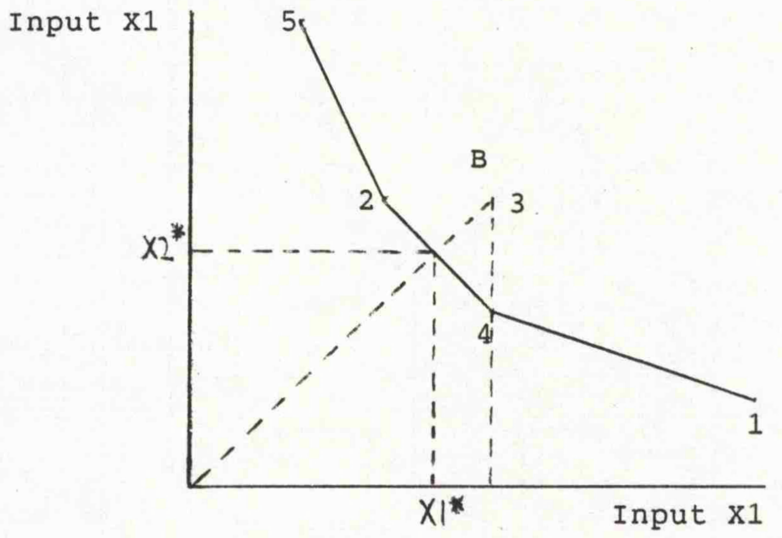
INPUT MINIMISATION. NON-INCREASING RETURNS TO SCALE.

1	2	3	
3	0.002	0.40	*
4	0.003	0.43	*
3	0.002	0.46	*
13	0.010	0.49	***
10	0.008	0.52	**
16	0.012	0.55	****
34	0.027	0.58	*******
32	0.025	0.61	*******
42	0.033	0.64	*********
72	0.056	0.67	***************
81	0.063	0.70	*****************
80	0.062	0.73	****************
111	0.087	0.76	***********************
96	0.075	0.79	********************
111	0.087	0.82	***********************
104	0.081	0.85	*********************
90	0.070	0.88	******************
84	0.066	0.91	*****************
53	0.041	0.94	***********
49	0.038	0.97	**********
193	0.151	1.00	***************************************

Statistics relating to DEA technical efficiency ratings under non-increasing returns to scale.

No.of DMUs	:	1281
No. technically efficient	:	175
Mean Efficiency	:	0.811
Median	:	0.693
Stan. Dev.	:	0.134

Figure 5.3.2.
Constructing a target for a Crown Post Office.



where the contents of the brackets are the scalar efficiency score and OB is the vector reflecting the CPOs current usage of inputs. The efficiency score implies an equi-proportionate contraction in each input of

5.3.2 (1-OA/OB)
to X*c, c=1,2, in Figure 5.3.2.

As an example to find out the level of adjustment in labour inputs implied by the efficiency scores, one CPO in the Northern Ireland region will be examined. This region being chosen because it is the smallest and the easiest to display. Table 5.3.5 shows the efficiency scores of CPOs in the Northern Ireland region. There are 24 CPOs in the region and these are codenamed 987-1010. Due to the fact that the efficiency score of efficient best-practice CPOs is unity and its peer group is simply itself, such CPOs are excluded from the table, leaving 18 inefficient ones. The CPO chosen was number 999 because it had the lowest efficiency score in the region of 0.638. On the basis of the formula for efficient production in formula 5.3.2 this implies a lower target level of the use of labour inputs of

$$OA^* = (0.638)2236 = 1426.57$$

1415.4 is the target level of labour inputs which would put CPO 999 on the best-practice isoquant. DEA is predicting that CPO 999 can support existing levels of traffic and quality with a reduction of (1-0.638) = 36.2 per cent in its current workforce. If the slack variables are known then further adjustments could be made but, this information is not available to me. Table 5.3.6 shows the targets and savings in labour inputs for CPOs with DEA relative efficiency less than unity for the whole of Northern Ireland.

There is not the space here to display the targets and efficiency savings for all of the Crown Post Offices and they are not necessarily that meaningful due to the large number of DMUs, it is better to examine the targets on an aggregated basis, and this will be done later. As a whole then, across the 1281 CPOs the mean efficiency is 0.814, this would suggest that labour inputs need to be adjusted by 18.6%. The average percentage saving in labour inputs conceals wide disparities in the savings available to individual DMUs. The most inefficient DMU which is DMU number 77 with a technical efficiency number of 0.386 has a target saving of labour inputs of 61.4%. But even so, the mean efficiency rating implies that there could be a large target reduction of labour inputs over the sample as a whole. This indicates that a large number of CPOs are overstaffed, they are not performing their functions efficiently or are not being utilised efficiently. Examining the slack variables shows that increases in output, in terms of the quality variable can also be made. Across the whole sample this output could be increased by 6.7%. As this is the inverted average waiting time this implies that AWT could be decreased by 6.7%. A not unsubstantial target but not very large in the context of the actual amount of time it consists of, a matter of a fraction of a minute. It must be remembered of course that this is an average, much more

Table 5.3.5.
Summary measures of Crown Post Office technical efficiency with varying returns to scale for the Northern Ireland region.

CPO	TE Number	Peer Group							
987	0.971	43	288	471	707	959	997	1010	
988	0.979	43	261	641	707	716	1108		
990	0.898	64	707	959	993	1010	1079		
991	0.819	43	64	462	473	1123			
992	0.942	471	627	639	673	1010	1026	1075	
995	0.860	43	471	497	499	707	959		
996	0.873	43	931	933	1010	1026	1263		
998	0.855	43	473	933	997	1010	1079	1123	
999	0.638	43	473	931	933	1010	1026	1123	
1000	0.894	471	473	716	1010	1026			
1001	0.917	43	641	716	933	1010	1079	1108	
1003	0.887	43	641	707	959	1002	1100	1104	
1004	0.803	43	288	641	707	930	959	1010	1079
1005	0.944	43	396	471	641	716	933	1010	
1006	0.928	43	64	707	716	959	1010	1079	1102
1007	0.851	43	64	497	707	716	933	1079	1102
1008	0.962	497	530	641	707	1010	1106		
1009	0.957	43	530	641	707	1010	1106	1108	

Table 5.3.6.
Targets and savings in labour inputs for Crown Post Offices with DEA relative efficiency less than unity.

CPO	Actual performance (Labour inputs in hours)	Target performance (Labour inputs in hours)	Saving	Saving %
987	2484.00	2411.96	72.04	2.90
988	2542.00	2488.62	53.38	2.10
990	3874.00	3478.85	395.15	10.20
991	1666.00	1364.45	301.55	18.10
992	1628.00	1533.58	94.42	5.80
995	3151.00	2709.86	441.14	14.00
996	1717.00	1498.94	218.06	12.70
998	1862.00	1592.01	269.99	14.50
999	2236.00	1426.57	809.43	36.20
1000	2110.00	1886.34	223.66	10.60
1001	2252.00	2065.08	186.92	8.30
1003	4021.00	3566.63	454.37	11.30
1004	3177.00	2551.13	625.87	19.70
1005	2281.00	2153.26	127.74	5.60
1006	3000.00	2784.00	216.00	7.20
1007	2824.00	2403.22	420.78	14.90
1008	2146.00	2064.45	81.55	3.80
1009	2216.00	2120.71	95.29	4.30

substantial reductions are available in individual cases.

To identify targets and savings is one thing to be able to implement and realise them in practice may not be so feasible. For instance, there will be an interaction between the numbers of staff-hours and the AWT. They will be inversely related so that improvements in one will cause a decline in the other. The solution is to have a more precise matching of staff to traffic, through more accurate forecasting of traffic and better scheduling. A possibility to help improve the performance of an inefficient Crown Post Office would be to examine its peer group. That is, those CPOs that form its reference best-practice group in relation to which it is inefficient. For instance, CPO 999 in Table 4.5.2 has for its peers CPOs 43, 473, 931, 933, 1010, 1026 and 1123, which in linear combination define its target performance. This reference group are likely to be implementing more efficient management procedures which the inefficient DMU can learn from. As was explained in Chapter 4 DEA can play a major part in the improvement of CPO performance. But, this has to be followed up with an examination of peer group methods and the introduction of new procedures.

The variation in efficiency noted above shows that there are wide disparities in efficiency overall. That is, relative to best-practice a large number of DMUs have a low level of performance. On casual inspection of the results it seems that the South-East region has a low number of technically efficient CPOs. It might seem possible then that the efficiency variation could be accounted for by regional efficiency variation. The results will be looked at more closely to establish that this is the case and then reasons why efficiency should vary on a regional basis will be examined.

5.4 DEA results: Counters regional efficiency.

5.4.1 Analysis of the Results on a Regional Basis.

The results can be divided up on a regional basis, they are the ten regions of the United Kingdom described earlier. Although the regions were re-organised and divided into four much larger territories in 1986, there are good reasons for classifying the data into these regions.[1] Greater disaggregation will yield more informative results. But, the main reason is that for historical reasons data collection is still being conducted on the basis of these regions and the data was supplied in these groups. A key reason for using them for the purpose of this study is that because the regions were administered and organised on this basis for a considerable period of time they are likely to be more homogenous in efficiency terms and have a similar management ethos.

To find out if efficiency varies on a regional basis it might be useful to examine the proportion of each regions CPOs that are technically efficient. As an aid to finding out how significant these proportions are, the percentage of technically efficient CPOs was calculated for all 1281 CPOs as being 13.82%. There are five regions whose number of technically efficient CPOs fall below this average and they are the Eastern, London, South-East, South-West and Wales regions (See Table 5.4.1), though Wales is only marginally so. The North-East and Northern Ireland regions are way out in front with 28.04 and 25% of their CPOs being technically efficient.

Just looking at the proportion of technically efficient CPOs in each region is not necessarily that useful, it does not show what is happening to the other offices within the region. A large number of them may be very close to being efficient. To illustrate this and show the distribution of TE numbers in each region a histogram is used to summarise each one. Each histogram is divided up into class intervals of 0.03. A measure of location to establish how different

Table 5.4.1.

Complete statistics relating to the regional breakdown of the results.

		Number Of Obs.	Number Tech. Eff.	% Tech. Eff.
1	London	325	24	7.38
2	Eastern	95	7	7.37
3	Midlands	130	26	20.00
4	North-East	107	30	28.04
5	North-West	122	25	20.49
6	South-West	125	13	10.40
7	Wales & the Marches	82	11	13.42
8	Northern Ireland	24	6	25.00
9	Scotland	159	29	18.24
10	South-East	112	6	5.36

		Mean TE	Stan Dev.	Lowest TE No.	Median
1	London	0.705	0.144	0.386	0.693
2	Eastern	0.813	0.097	0.628	0.814
3	Midlands	0.876	0.010	0.612	0.806
4	North-East	0.886	0.106	0.609	0.805
5	North-West	0.867	0.097	0.619	0.810
6	South-West	0.828	0.101	0.609	0.805
7	Wales & the Marches	0.854	0.105	0.661	0.831
8	Northern Ireland	0.916	0.086	0.638	0.819
9	Scotland	0.875	0.095	0.612	0.806
10	South-East	0.783	0.100	0.583	0.792

efficiency is between regions would be to take the median of the DEA numbers. Though with a fixed upper value of 1 and the lower values clustered around 0.6 this is not very revealing. The exception is London whose lowest TE number is much lower than any others region and so has a significantly lower median of 0.693. A better measure of location to allow differentiation between regions would be to take the arithmetic mean of each regions CPOs' TE numbers. Average technical efficiency for all 1281 CPOs is 0.814 so regions who have a TE of less than this are below average and regions above this level have greater than average technical efficiency.[2] Regions that have below average technical efficiency are London, Eastern and the South-East. The region with the highest technical efficiency is Northern Ireland with an average of 0.916.

The ranking of the regions arising from ordering the arithmetic means is virtually the same as the ranking from ordering the proportions of each regions TE numbers that are technically efficient. This is because a higher number of DMUs with a technical efficiency of 1 will contribute to a higher mean. An exception to this is London who whilst having the third lowest proportion of its CPOs being technically efficient, has the lowest average technical efficiency. This would seem to indicate that it has a larger number of heavily inefficient DMUs. This is confirmed by casual observation of the histogram of Londons results (see Table 5.4.2), also, it has the lowest median and its standard deviation is the largest. The North-West also has the same problem, it has the third highest ATE but only the fifth highest ranking when looking at the percentage of efficient DMUs in each region. The reason why Northern Ireland and the North-East reverse their rankings on these two criteria is that the North-East has a very wide spread of efficiency ratings, as confirmed by the fact that it has the joint third lowest median and the second highest standard deviation.

To get a clearer picture of exactly what the scale of the inefficiency of Londons CPOs is, the bottom 5% of the

Table 5.4.2.

Histograms of DEA technical efficiency numbers for the ten Counters regions.

Col. 1. is the number of post offices in the class
Col. 2. is the mid-points of the efficiency classes
Each * represents 1 observation

```
1. LONDON

 2    0.40   **
 4    0.43   ****
 4    0.46   ****
12    0.49   ************
10    0.52   **********
15    0.55   ***************
27    0.58   ***************************
30    0.61   ******************************
24    0.64   ************************
30    0.67   ******************************
31    0.70   *******************************
29    0.73   *****************************
23    0.76   ***********************
15    0.79   ***************
10    0.82   **********
 9    0.85   *********
 4    0.88   ****
10    0.91   **********
 6    0.94   ******
 4    0.97   ****
26       1   **************************


2. EASTERN

 3    0.64   ***
 4    0.67   ****
 7    0.70   *******
10    0.73   **********
13    0.76   *************
13    0.79   *************
10    0.82   **********
 7    0.85   *******
 8    0.88   ********
 7    0.91   *******
 3    0.94   ***
 1    0.97   *
 9       1   *********


3. MIDLANDS

 1    0.61   *
 2    0.64   **
 2    0.67   **
 4    0.70   ****
 7    0.73   *******
 6    0.76   ******
11    0.79   ***********
 9    0.82   *********
14    0.85   **************
13    0.88   *************
14    0.91   **************
12    0.94   ************
 6    0.97   ******
29       1   *****************************
```

4. NORTH-EAST

```
 1    0.61  *
 0    0.64
 1    0.67  *
 4    0.70  ****
 3    0.73  ***
14    0.76  **************
 8    0.79  ********
 6    0.82  ******
 7    0.85  *******
10    0.88  **********
 5    0.91  *****
 6    0.94  ******
 8    0.97  ********
34       1  **********************************
```

5. NORTH-WEST

```
 2    0.61  **
 0    0.64
 1    0.67  *
 3    0.70  ***
 3    0.73  ***
15    0.76  ***************
10    0.79  **********
15    0.82  ***************
16    0.85  ****************
11    0.88  ***********
 8    0.91  ********
 5    0.94  *****
 7    0.97  *******
26       1  **************************
```

6. SOUTH-WEST

```
 1    0.61  *
 1    0.64  *
 9    0.67  *********
12    0.70  ************
 7    0.73  *******
 7    0.76  *******
12    0.79  ************
20    0.82  ********************
12    0.85  ************
11    0.88  ***********
11    0.91  ***********
 2    0.94  **
 5    0.97  *****
15       1  ***************
```

7. WALES AND THE MARCHES

```
 5    0.67  *****
 5    0.70  *****
 4    0.73  ****
 8    0.76  ********
 6    0.79  ******
 6    0.82  ******
11    0.85  ***********
 8    0.88  ********
 4    0.91  ****
 3    0.94  ***
 9    0.97  *********
13       1  *************
```

8. NORTHERN IRELAND

```
 1    0.64   *
 0    0.67
 0    0.70
 0    0.73
 0    0.76
 1    0.79   *
 1    0.82   *
 3    0.85   * * *
 3    0.88   * * *
 2    0.91   * *
 3    0.94   * * *
 4    0.97   * * * *
 6       1   * * * * * *
```

9. SCOTLAND

```
 1    0.61   *
 1    0.64   *
 3    0.67   * * *
 6    0.70   * * * * * *
 1    0.73   *
12    0.76   * * * * * * * * * * * *
13    0.79   * * * * * * * * * * * * *
18    0.82   * * * * * * * * * * * * * * * * * *
13    0.85   * * * * * * * * * * * * *
21    0.88   * * * * * * * * * * * * * * * * * * * * *
16    0.91   * * * * * * * * * * * * * * * *
14    0.94   * * * * * * * * * * * * * *
 8    0.97   * * * * * * * *
32       1   * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * *
```

10. SOUTH-EAST

```
 2    0.58   * *
 2    0.61   * *
 6    0.64   * * * * * *
 6    0.67   * * * * * *
11     0.7   * * * * * * * * * * *
17    0.73   * * * * * * * * * * * * * * * * *
16    0.76   * * * * * * * * * * * * * * * *
10    0.79   * * * * * * * * * *
11    0.82   * * * * * * * * * * *
 8    0.85   * * * * * * * *
 6    0.88   * * * * * *
 6    0.91   * * * * * *
 2    0.94   * *
 2    0.97   * *
 7       1   * * * * * * *
```

Table 5.4.3.
Percentiles of the Efficiency Ratings and their regional composition.

Bottom Percen-tile	No. in Percen-tile	No. From London (%)	No. From SE (%)	No. From Other Regions (%)
5%	64	64 (100.00)	0 (0.00)	0 (0.00)
10%	128	114 (89.06)	7 (5.47	7 (5.47)
15%	192	148 (77.08)	13 (6.77)	31 (16.15)
20%	256	177 (69.14)	22 (8.59)	57 (22.27)
25%	320	198 (61.88)	33 (10.31)	89 (27.81)
50%	641	260 (40.56)	76 (11.86)	305 (47.58)
100%	1281	325 (25.37)	112 (8.74)	844 (65.87)

ranking of DMUs was examined (See Table 5.4.3). It transpired that of the 64 CPOs that existed in the bottom 5% all of them were in the London region. When the bottom 10% is examined, 114 of the technically inefficient CPOs are from London, 7 from the South-East and 7 from other areas. The South-East would be expected to have a large number of heavily inefficient CPOs because it has the second lowest mean. It might be that London's prevalence in the ranking could have something to do with the fact that it has the largest number of CPOs in Counters (325), just over a quarter of all CPOs are in London. But, in the bottom 50% of CPOs 40.56% of them are from the London region whereas in the top 50% only 5.07% are from London. Thus it is quite clear that a large number of the London regions CPOs are heavily inefficient.

Target changes in resources can also be revealed at the regional level and these are displayed in Table 5.4.4 the mean efficiencies imply the overall targets for labour inputs and so of course London has the largest possible target reduction of 29.48% and Northern Ireland the lowest

Table 5.4.4.

Potential gain from reductions in input and increases in output, due to meeting targets in each region.

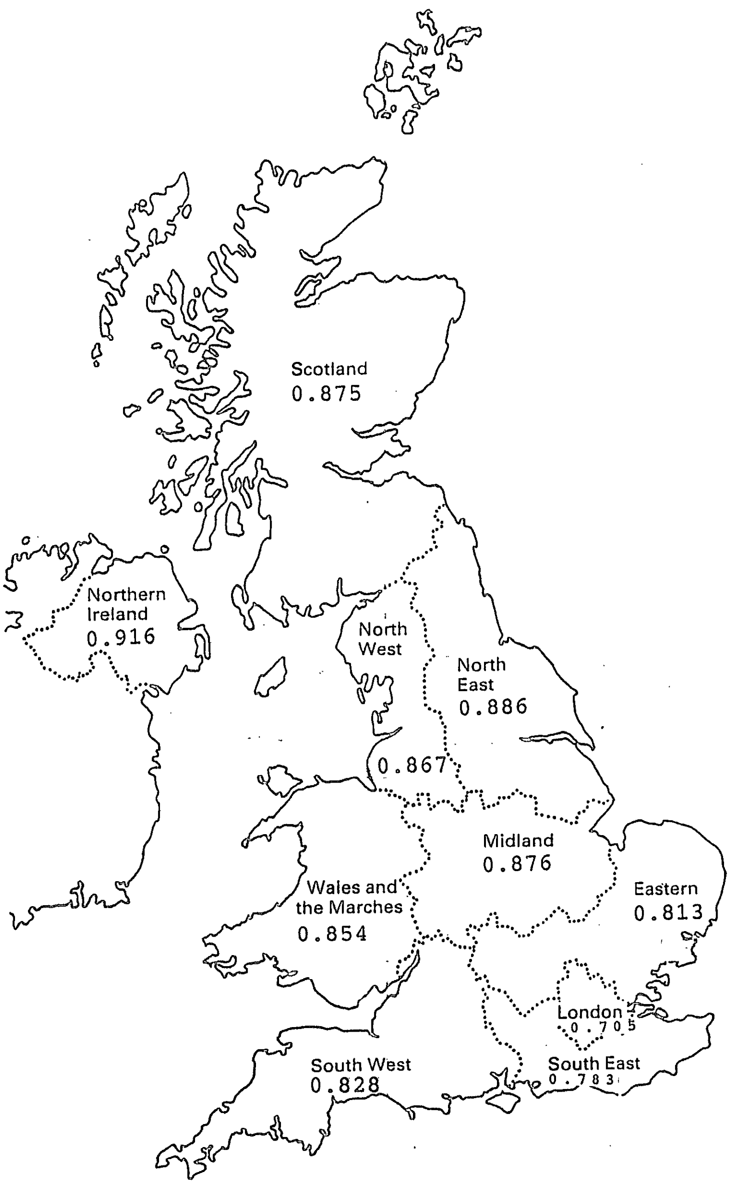
		LABOUR INPUT (HOURS)			QUALITY OUTPUT (INVERTED AWT)		
		ACTUAL	TARGET	POTENTIAL GAIN (%)	ACTUAL	TARGET	POTENTIAL GAIN (%)
1	London	1058217.00	746233.46	29.48	58370.00	65837.20	12.79
2	Eastern	252761.00	205373.37	18.75	17580.00	18385.75	4.58
3	Midlands	362350.00	317244.67	12.45	23921.00	25195.44	5.33
4	North-East	319249.00	282787.57	11.42	21739.00	22251.20	2.36
5	North-West	333299.00	288926.90	13.31	23552.00	24719.67	4.96
6	South-West	301473.00	249505.08	17.24	24006.00	25016.50	4.21
7	Wales and the Marches	181522.00	155028.86	14.60	15970.00	16530.62	3.51
8	Northern Ireland	66890.00	61254.52	8.43	4778.00	5011.84	4.89
9	Scotland	337654.00	295504.65	12.48	30230.00	32070.10	6.09
10	South-East	299520.00	234548.12	12.69	20819.00	22129.31	6.29
AVERAGE POTENTIAL GAIN (%)				18.58			6.72

with 8.43%. Much more interesting are the output targets for inverted AWT deduced through the slack variables. London again needs the largest reduction and could reduce its AWT by 12.79%, the smallest reduction is possible in the North-East with 2.36%. The regions with the lower labour targets are not necessarily the ones with the lowest quality targets (with the exception of London), implying that at some kind of level there must be a trade-off. The AWT targets, given that they refer to very small amounts of time, are very similar. This is indicative of the fact that Counters has much more freedom of action to control this output, are committed to reducing average waiting times and monitor them closely.

It has been established that efficiency varies across regions. Whilst it varies, it also seems that certain regions which are adjacent have similar average technical efficiencies (See Figure 5.4.1). London, South-East, Eastern and South-West regions, all in the South of England have the lowest ATEs. The North-East, North-West, Midlands and Scotland have similar statistical characteristics. With some adjacent regions having similar ATEs it must mean that through their proximity they share certain characteristics that determine efficiency. So far, differing levels of efficiency between regions have been shown, the reasons why efficiency might vary due to geographical location will be examined in the next section. This is important because it is necessary to identify other determinants of efficiency, to know why regions and hence individual CPOs are inefficient. Such differences would seem to go beyond differences in management practices. Also as DEA measures relative efficiency not absolute efficiency, Post Office Counters may be structurally ineffient. There may be factors common to all CPOs which decrease their efficiency. This possibility will be examined in a subsequent section.

Figure 5.4.1.

Map of the ten Counters regions and their ATEs.



5.4.2 What factors contribute to relative efficiency differences between regions?

There are a number of reasons why efficiency varies between Crown Post Offices and regions and they are to do with organisational practice and the utilisation of labour. The number of workers employed varies between regions so that overmanning can account for differences in efficiency. For, example London's CPOs employ an average of 7.9 Postal Officers per office, as against 4.8 per office elsewhere. Another aspect of this is that the number of hours worked varies between regions because of the amount of overtime worked. The amount worked is quite substantial but it is very much higher in London than elsewhere. For instance, overtime has ranged, in the period April to November 1987, from 3.11 per cent of total hours in one district to 14.56 in London East District. In addition, London's six districts worked more overtime than anywhere else.

A major determinant of the efficiency of a CPO and hence relative efficiency between CPOs and regions is the quality of management. There are two factors that play a part in this in Counters. Statistics produced by Counters show that in the 1987 staff appraisal review, 48 Branch Managers were reported as showing 'significant weakness in performance'. 289 were regarded as 'generally acceptable' (which is the category below 'fully acceptable'). Also, 314 staff were holding Branch Manager's jobs whose grades were below the level of the post holder. Thus, 22 per cent of Counter's branch offices are not well managed, and about 20 per cent are managed by staff in higher grades than their posts warrant. Counters explains the situation whereby staff held Branch Manager jobs graded below the level of the post-holder as arising as a result of reorganisation. Also, employees not reaching the required standard are understood to have been promoted to their grades under former promotion procedures which placed undue weight on seniority rather than on merit. The problem is that they cannot be absorbed elsewhere into the district structure, and under existing policies cannot be made redundant. They cannot be given voluntary early retirement without financial penalty or being stigmatised as inefficient. Staff who have performed inefficiently over a period of time through lack of ability are declared to be 'not culpably inefficient', and if over fifty given full benefits. Counters has tried to improve this by improving performance through guidance and development. Where individuals do not respond, the question of early retirement would have to be applied as appropriate. The problem is such that the Post Office is currently reviewing its corporate policy on any retirement and voluntary severance. Even given the fact that within Counters bad managers have been identified, the scope for improvement and differences between managers will continue to exist and account for variation in efficiency.

Whilst this former point accounts for variation in efficiency between CPOs it does not really account for regional variations in efficiency. Another factor is the experience of the Counters staff themselves. How long a member of staff has worked for Counters affects their competence and efficiency in dealing with transactions. This is a variable because the turnover of staff varies between regions, the higher the turnover, the lower the collective experience of the staff is. The problem of high staff turnover is a particular problem in London. Relatively low wages, high living costs and the easy availability of other forms of work due to low unemployment means that turnover is high. This arguement could also be true of the other regions in the South and East of England, hence explaining their low efficiency.

Another aspect of this problem is that sometimes staff need to be shed but cannot be. Counter's natural wastage rate is about eight per cent, which in general provides a margin within which jobs can be reduced without resorting to compulsory termination of employment. However, in some regions turnover is so low that there is little flexibility within which to reduce staff and overmanning results. This could be due to high local levels of unemployment. This could be a contributory factor in explaining why the

North-East, North-West and Wales regions ATE is below the mean for all CPOs of 0.871.

5.4.3 What special factors account for inefficiency in the London region?

It should be obvious from the DEA results that the London region has the lowest average technical efficiency. Its efficiency level is such that examination of its particular problems is justified. These are in part due to the overmanning and high overtime working already described and in part due to inefficient working practices which have become entrenched. This has not been easy to alleviate because of the strength and inflexibility of trade unions in London. There are no closed shops but unionisation is high.

Conclusion of the RCWSA agreement in London was followed by discussion between the General Manager, London Counters Territory, and the UCW London District Council. The discussions ended with a local agreement which, in effect, deferred the introduction of part-time and casual staff in London. The use of part-time and casual staff is crucial to increased flexibility in scheduling and hence the efficient use of labour. This agreement was in exchange for a commitment to secure budget savings of 468,000 hours per annum from desk-top staffing reviews and the eventual revision of working practices relating to:

a. preparation for duty time, balancing, and end of day procedures,
b. personal stock transfer arrangement,
c. security arrangements,
d. early day attendance and meal relief procedures.

It has become the case that time allotted to a task is not necessarily warranted by the nature of the various tasks or the time required. For example, half an hour is allowed for 'end of day' procedures at London offices compared with about fifteen minutes in a typical provincial office.

Secondly, when stocks are balanced and transferred between counter staff in London offices, this handover is made on a personal basis. This reflects historical problems with security, but in most provincial offices the more efficient arrangement of third party witnessing works satisfactorily and could be applied to London offices. Thirdly, security staffing and procedure is very complex in London offices, and results in up to four staff staying behind at the close of business to secure an office. This is compared to two, typically, outside London. Security arrangements could be simplified, and the number of keys and level of security reduced. Lastly, because counter duties in London typically extend to 2.30 p.m. on 'early day', a full (one hour) meal break applies under the terms and conditions of employment for Postal Officers. An earlier duty end time would reduce the meal break to half an hour, without reducing customer serving time.

In addition to these factors there are a number of other contributors to inefficiency in London. They are talked about in more detail in other parts of this chapter but are listed briefly here. London has the highest rate of sickness absence. A high proportion of PA work in London is done by POs at higher cost, its districts employ only 27 out of 278 PAs in all districts. Staff turnover is very high in London, resulting in lower levels of experience and therefore efficiency than elsewhere.

5.4.4 Additional variables that could have been incorporated into the DEA analysis.

There are some factors which it would have been desirable to incorporate into the DEA analysis of Counter data, and would have given a much clearer picture of the efficiency position. Though it was either not possible to gain access to the information at the time, or after analysis and discussion of the results it has became apparent that additional factors are important in determining efficiency.

Overtime is quantified in hours in the DEA analysis so that it contributes to the total labour input of an office. But, the cost to Counters of overtime is higher than it appears as wage payment is at a higher rate for overtime working. Two offices with similar labour inputs might have very different labour costs because their mixture of normal and overtime hours worked is different. Unfortunately, cost data could not be made available to me, so it was not possible to incorporate this into a more accurate DEA analysis.

Another factor which can only be incorporated if labour inputs are measured in terms of their cost, is that of the inefficient use of resources due to work being undertaken by workers of an inappropriate grade. Postal Assistants do routine clerical tasks of a simple nature. The Postal Officer grade is the main clerical grade and includes all counter clerks. Counter clerks constitute some eighty percent of the PO grade in the Counters business. In London, a great deal of PA work is being undertaken by POs, whose earnings in 1988 were sixty pounds per week higher than those of PAs. One example of this is the employment of POs on PA work in remittance units. These practices mean that the work is being done at much higher cost than it should be. As the DEA analysis uses the number of hours worked as an input this aspect of working practice cannot be taken into account. If the figures were available for the cost of staffing in each office the result would have given a much clearer indication of the true efficiency position. The problem of finding the true cost of the labour inputs is not just a problem with overtime and workers doing jobs below their grades, it is also a problem with other labour costs like sickness absence.

A useful piece of information to incorporate into the DEA analysis would have been the effect of staff turnover on efficiency. That is, as explained earlier, high turnover leads to low levels of experience and efficiency amongst staff and low turnover means that it is difficult to reduce the numbers of staff if this is desired. Incorporating these factors into the DEA analysis is problematic. The average resignation rate (ARR) could be used as a proxy for staff experience but a high turnover and a low turnover are both contributors to inefficiency. It might be possible to incorporate this aspect, bearing in mind that the closer a rate is to the average, the better (though of course this may not always be the case as the ARR could be high or low). Treating it as an environmental input it will have to be positively correlated to output, so that higher values result in higher output. Taking the average resignation rate as the ideal rate this can be achieved as follows. To find a value for improvements away from low turnover, subtract the lowest resignation rate from the ARR to give a range of below average resignation rates. Subtracting any value below the ARR will now give a value that can be used in a DEA. This process can be repeated for the range of above average resignation rates as well. Now resignation rates that are above and below the average rate, as they move towards it, make the calculated values become higher, thus they are positively correlated to output and are contributing to efficiency.

A problem for Counters is the high cost of maintaining a high street presence. This is in terms of rental costs and the opportunity cost of capital of owning valuable high street sites, when a move to a less prominent site in a standard shop unit would still enable an adequate service to be provided. Also, the use of premises that are too large or are really unsuitable for the purposes for which they are used. DEA cannot solve the problem of identifying those CPOs which would benefit from the move to a lower rent premises. This is because rent per square metre of floor space varies throughout the country. There will even be wide variation within regions so a clustered regional analysis would not be useful. Counters has already identified the problem of the unsuitability of using prestige prime sites. It owns high street sites within a total freehold and leasehold property valued at £133m in 1988. It has started a programme of property disposals. At that time it predicted that disposals could yield £25m per year for the following three years.

DEA can identify those CPOs who are occupying offices that are too large. If the information had been made available to me I would have liked to have included the size of branch by using square metres of branch floor space as an input. This would identify over large offices relative to all other CPOs and suggest the optimum branch size. Of course branch space itself has certain weaknesses because it does not reflect aspects such as floor plan, quality of installation etc. Nevertheless, this would still be the best proxy for branch size.

If it had been possible to incorporate the additional information described above I do not think that it would affect the efficiency ranking of the regions very much, though of course the CPO ratings would change. It would affect the DEA ratings because most of the factors would have a detrimental affect on some of the CPOs. The result would be that the gap between the most and least efficient would be much wider. London, for instance, would be shown to be even more inefficient than it is at the moment.

5.4.5 Does structural efficiency exist within Post Office Counters?

As has been seen, there are factors that contribute to variation in efficiency between regions and CPOs, and this accounts for different DEA ratings (relative efficiency differences). But, there are factors that cause inefficiency in all CPOs, they are integral to the way that Counters is managed and operated. These factors may be worse in some offices than others, but they are common to all.

The inefficient working practices referred to when describing the specific problems of London are not unique to the area. Although they are most deeply entrenched in London, the problem exists throughout the branch office structure. Though the worst of these have been removed in the RCWSA and EME agreements, and also there is increased flexibility due to the removal of the inhibition on the use of part-time and casual staff.

As well as the inefficient working practices identified by the agreements, the MMC report 1988 identified some others. Balancing time , the weekly balancing of cash and stock, is left to local discretion. It varies from one hour ten minutes to two hours, according to location. As there is discretion, and there are no standard times, structural inefficiency must exist. In addition, the MMC found that work appropriate to PAs in branch offices is being undertaken by POs and in remittance units POs do the work of PAs thus raising the cost.

As mentioned earlier, overtime working is too high throughout Counters and is used for the wrong purposes. Most of it is worked on rest days and Saturdays and overtime has become routine in covering sick leave and other absences. It is also implicated in carrying out some of the tasks described as inefficient working practices. Another inefficient practice which is related to overtime working, is that in the pre-Christmas period, the granting of overtime has become not so much a matter of responding to workload as a matter for prior negotiation with the unions. Pre-Christmas working is negotiated on the basis of previous experience and knowledge of levels of demand. The facility to work predicted levels of overtime at particular times is ultimately discussed at branch level between the manager and workplace representatives. The Branch Manager decides if overtime needs to be increased or decreased as a result of the day-to-day working of that office. This practice is clearly not efficient, a block of overtime is negotiated and worked even though there may not be sufficient work to do. To maximise efficiency labour should be more flexible (though of course fairly compensated for this flexibility) so that overtime is worked when necessary or cost-effective.

Sickness absence by Counters staff as a whole is too high. In 1986/87 absences were 3.18 per cent of total hours worked, with overtime worked on account of absence equating to 2.69 per cent of gross staff hours paid. In the period

April to August 1987 the corresponding figures were 2.8 per cent and 3.53 per cent. A survey undertaken in spring 1987 by the Confederation of British Industry showed the average rate of sickness absence for non-manual workers is 2.2 per cent.

Counters is keen to ensure that there is an adequate supply of trained staff available at the right time to each office. In deciding how many staff are needed in each location, allowance is made to reflect sickness absence and staff on holiday. Some districts have a pool of employees earmarked to cover leave throughout the district. In others, 'leave reserves' are in small groups, covering two or three offices. Whilst this is not inefficient in itself, and ensures adequate staffing levels, it is a potential source of inefficiency through overstaffing of the leave reserve. This occurs if the leave reserve is staffed for summer leave levels which are high, rather than winter ones which are low. This has been a source of inefficiency, though appropriate guidance is being produced for districts on this.

Post Office Counters is a largely paper-driven organisation. Unlike other high street organisations there has been no computerisation of transactions and it has no automatic teller machines (ATMs). This has to be a major source of inefficiency, much labour could be saved by computerising those operations currently done by hand. Fortunately, Counters is about to remedy this problem. A pilot scheme involving the installation of computer terminals in 250 post offices in the Thames Valley region was completed in April 1990. Each post office is connected to a network linked to a central computer. This is linked to three POC clients, the DVLA, National Savings and Girobank. Girobank customers can withdraw cash using ATM cards, the card is passed through the computer terminal by a counter clerk and the customer keys in their personal identification number on a counter pad. Not really the same concept as providing external twenty-four hour a day cash machines, more an unwieldy method of withdrawing money, but a step in the right direction (Girobank customers can still of course withdraw money from their own cashpoints). The scheme has been deemed to be a success and on this basis Counters plans to install 6,300 Unisys counter terminals in Crown and sub post offices. These will significantly increase the efficiency of counter clerks and the technical efficiency of Counters as a whole, reductions in waiting times should also be possible.

Thus, there is some structural ineffiency in Post Office Counters. Technical efficiency as measured by DEA, is measured relative to best-practice. As all CPOs suffer from structural efficiency a CPO with unit efficiency as measured by DEA is not necessarily truly technically efficient. Having said that, the position should improve, Counters seems to be actively seeking to eliminate some of the causes of structural inefficiency.

5.5 The effect of clustering by geographical location and conducting a DEA.

5.5.1 The necessity of clustering the CPOs by geographical location.

The previous sections looked at the interpretation of the results of a DEA run on Crown Post Offices. Conclusions were drawn about the overall relative technical efficiency and some more insight was gleaned by looking at the results on a regional basis. Reasons were identified why there was a regional variation in efficiency. Accurate measurement can be taken a step further by disaggregating the data into regions and running the DEA programme once for each region. This clustering of DMUs amounts to varying the number of constraints in the DEA programme. Conclusions can also be drawn about the sensitivity of DEA to the number of DMUs used in the DEA run and the effect on efficiency ratings, peer groups and efficiency targets.

For a relative measure of efficiency it is very important that DMUs exist in similar operating environments. Dissimilarities may affect inputs or outputs adversely. It might be productive then to conduct some kind of cluster analysis. Clustering by a common factor can exclude the effect of a variable that cannot be included directly by making it common to all. Factors that affect efficiency will have similar values within regions and so be taken into account. Deciding the basis of why DMUs are similar and organising them into homogenous groups is usually quite difficult in the public sector as no two agencies will be operating in exactly similar circumstances. Even if it is possible, there can never be a homogenous set of units, differences will always exist in the way units are managed because they are led by different decision-makers. But, as this is the factor in efficiency determination we are trying to isolate it is not a problem.

For the Post Office CPOs they could be divided up into similar groups on the basis of output-mix. However, the reasons for the dissimilar performance of CPOs have already been identified, the turnover of staff (too high or too low) and the militancy of unions (conducting strikes or impeding staff reductions and more flexible working arrangements in areas where they are strong). As this varies regionally the best basis upon which to choose clusters are the regions already identified. A separate DEA run can then be conducted on each cluster to reveal more about efficiency within the region. The only drawback to this approach is that as the clusters will not be of equal size, the smaller the cluster, the higher will be the efficiency scores and the less discrimination there will be. This means that efficiency scores cannot be compared between clusters. But, as has already been pointed out, there are factors which are largely endogenous that determine that efficiency between regions will vary anyway, so this point may not be a cause for concern. What can be compared are efficiency disparities within regions.

This clustering approach is supported by Golany and Roll (1989) who say that in general 'the larger the number of units in the analysed set, the lower the homogeneity within the set'. Therefore they suggest that 'another direction in the analysis of efficiency outcomes is partitioning the group of DMUs into categories, according to some characteristic which was not entered into the model as a factor determining input/output relationships. The purpose of such categorisation is twofold: one is to gain a better relative assessment of efficiency, by comparing performance within sub-groups of units operating under similar conditions (eg. the same geographical region). The other is a comparison between categories, such as in the case where a category signifies a programme which a sub-group of DMUs operate'. In this case the reason is the former, thereby producing more information about the relative efficiency of Counters branches.

The clustering idea is not without criticism. There may be a problem in defining the cluster. Even though clustered DMUs share a similar geographical location, they may still be very different and have dissimilar characteristics. Also, why is clustering necessary when certain variables that vary with geographical location could just be included in the first place as background variables? Whilst this is true, identification of variables may be a problem and in the Counters case the data was not available anyway.

The clustering of DMUs into different groups has already been given some attention in the recent literature. DMUs are split into rural and urban classes for Sengupta and Sfeir's (1986) study of high school districts and groups of 'not-for-profit' and public hospitals in California are compared in Grosskopf and Valdamis (1987) evaluation of the issue of ownership in hospital performance. Ganley (1989) divides English local education authorities into three regional categories to draw conclusions about changes in efficiency ratings.

Ganley does not just analyse the change in efficiency he also examines the effect of clustering on different aspects of the discriminating power of DEA. That is the changes in rankings, peer groups and targets for untied DMUs. As we are clustering Post Office DMUs to gain more information about regional efficiency the findings will be used to assess Ganleys results later in this chapter.

There are more sophisticated ways of clustering DMUs than dividing them up on a regional basis and these should be mentioned briefly. Sexton et al. (1986) cluster DMUs so that apparent similarities amongst sub-populations of diverse DMUs can be checked. They point out that a DMUs optimal weights reflect the evaluation scheme under which the DMU is as efficient as possible. It follows then that DMUs which select similar weighting patterns are likely to use similar production processes and that the weights themselves can be used to form clusters of similar DMUs. The weights can be normalised to make them comparable. This approach though appearing to have value, will not be explored, its examination will require research at some time in the future.

A separate DEA run was conducted and a new frontier estimated for each of the ten clusters of CPOs based on the ten Post Office regions. Greater internal homogeneity now exists because factors that could not be included as variables, like trade union strength and turnover of staff that vary across the country, will be quite similar in the clustered regions. Although performing DEA on these clusters may produce more meaningful rankings it is necessary to find out what happens to efficiency ratings, peer groups and implied targets after clustering. This is important because each of these things is necessary to provide CPOs with the means to improve performance.

5.5.2 The effect of clustering on efficiency ratings.

In a previous section for a DEA run across all DMUs, 177 or 13.8% were identified as best-practice, that is, having a technical efficiency of 1. Examining the number of efficient DMUs after the CPOs have been clustered shows that a much greater number of them are efficient (histograms of the new efficiency ratings are displayed in Table 5.5.2). There are now 597 or 46.6% of the CPOs that are technically efficient. When comparing the number of technically efficient DMUs within each region from the clustered analysis (See Table 5.5.1), to the number from the single analysis (See Table 5.4.1), it seems that the number of best-practice ratings have risen in all regions. The average technical efficiency has also risen in all regions. Across the sample as a whole, average technical efficiency has risen from 0.814 to 0.930.

What does this result actually mean? As the DMUs are clustered within regions they are more homogenous. It would be expected then that DEA efficiency ratings within each region would be more similar. They are, as confirmed by the fact that if the standard deviation of the results in each region (the standard deviation shows how much variation around the mean there is) are compared to the standard deviation of results in the clustered analysis, in each case they are lower. As the variance around the mean is higher and the mean itself is higher, this confirms the fact displayed in the histograms that the DMUs ratings in each region are more similar to each other than they are to all CPOs as a whole. In general it seems that those regions that had the lowest numbers of efficient CPOs have increased their number of efficient DMUs by the highest factor. That is, the Eastern, South-East, South-West, Wales and London regions respectively, have increased their number of efficient DMUs the most (Eastern and South-East region by a factor of 8). But, whilst it may first appear that the incidence of best-practice in the regions was more prevalent than first appeared, this is not actually the case. It just shows that the gap between the best and worst CPOs is less within regions than throughout the Post Office as a whole. That is, as efficiency is measured relatively, relative to the new frontier more DMUs are efficient. In the cases where the mean technical efficiency and number of technically efficient DMUs is still low (eg. London) it shows that the range of CPOs with differing efficiency is still relatively very wide.

Table 5.5.1.

Complete statistics relating to the clustered results.

		Number Of Obs.	Number Tech. Eff.	% Tech. Eff.	Factor by which best-practice increases. after clustering.
1	London	325	84	25.85	3.5
2	Eastern	95	56	58.95	8.0
3	Midlands	130	72	55.38	2.8
4	North-East	107	64	59.81	2.1
5	North-West	122	58	47.54	2.3
6	South-West	125	63	50.40	4.8
7	Wales & the Marches	82	46	56.10	4.2
8	Northern Ireland	24	20	83.33	3.3
9	Scotland	159	86	54.09	3.0
10	South-East	112	48	42.86	8.0

		Mean TE	Stan Dev.	Lowest TE	Median No.
1	London	0.857	0.132	0.468	0.734
2	Eastern	0.970	0.054	0.764	0.882
3	Midlands	0.951	0.072	0.693	0.847
4	North-East	0.962	0.062	0.771	0.886
5	North-West	0.957	0.059	0.737	0.869
6	South-West	0.952	0.064	0.757	0.879
7	Wales & the Marches	0.950	0.075	0.738	0.869
8	Northern Ireland	0.981	0.053	0.770	0.885
9	Scotland	0.956	0.062	0.756	0.878
10	South-East	0.941	0.068	0.767	0.884

Table 5.5.2.

Histograms of DEA technical efficiency numbers for the ten clustered Counters regions.

Col. 1. is the number of post offices in the class
Col. 2. is the mid-points of the efficiency classes
Each * represents 2 observations unless otherwise stated

1. LONDON

1	0.46	*
1	0.49	*
2	0.52	*
7	0.55	****
5	0.58	***
5	0.61	***
8	0.64	****
6	0.67	***
14	0.70	*******
18	0.73	*********
21	0.76	***********
19	0.79	**********
24	0.82	************
24	0.85	************
22	0.88	***********
28	0.91	**************
9	0.94	*****
19	0.97	**********
92	1	**

2. EASTERN

1	0.76	*
1	0.79	*
2	0.82	*
3	0.85	**
4	0.88	**
8	0.91	****
7	0.94	****
12	0.97	******
57	1	*****************************

3. MIDLANDS

1	0.70	*
0	0.73	
2	0.76	*
5	0.79	***
3	0.82	**
12	0.85	******
5	0.88	***
8	0.91	****
12	0.94	******
5	0.97	***
77	1	***************************************

4. NORTH-EAST

1	0.76	*
4	0.79	**
1	0.82	*
7	0.85	****
4	0.88	**
8	0.91	****
7	0.94	****
6	0.97	***
69	1	***********************************

5. SOUTH-EAST

2	0.73	*
0	0.76	
1	0.79	*
1	0.82	*
6	0.85	***
9	0.88	*****
12	0.91	******
15	0.94	********
8	0.97	****
68	1	**********************************

6. SOUTH-WEST

1	0.76	*
1	0.79	*
8	0.82	****
8	0.85	****
9	0.88	*****
7	0.91	****
12	0.94	******
13	0.97	*******
66	1	*********************************

7. WALES AND THE MARCHES

1	0.73	*
1	0.76	*
4	0.79	**
4	0.82	**
3	0.85	**
7	0.88	****
3	0.91	**
6	0.94	***
2	0.97	*
51	1	**************************

8. NORTHERN IRELAND *=1

1	0.76	*
0	0.79	
0	0.82	
0	0.85	
1	0.88	*
0	0.91	
2	0.94	**
0	0.97	
20	1	********************

9. SCOTLAND

2	0.76	*
3	0.79	**
3	0.82	**
11	0.85	******
10	0.88	*****
14	0.91	*******
14	0.94	*******
14	0.97	*******
88	1	**

10. SOUTH-EAST

1	0.76	*
3	0.79	**
5	0.82	***
11	0.85	******
11	0.88	******
11	0.91	******
9	0.94	*****
7	0.97	****
54	1	***************************

There are two complicating factors in the interpretation of the results after clustering. As Bowlin (1987), Ganley (1989) and others have pointed out, the lower the number of DMUs in a DEA analysis the higher the efficiency ratings will be. The results would seem to bear this out. In general, the smaller the cross-section, the higher is the number of technically efficient DMUs and the mean technical efficiency. For instance, Northern Ireland the smallest cross-section, although already having the second highest proportion of technically efficient CPOs (25%) manages to increase this to the highest proportion after clustering (83.3%). Although very efficient overall to start with, because of its very small cross-section it manages to increase its number of technically efficient CPOs by a very large amount. This means that as the rankings are dependent upon the size of the cross-section it is not possible to compare DMUs between clusters.

In addition, the clustered DMUs do not tell us much about technical efficiency because they are not being compared to the other CPOs in Counters. All the clustered results are, is an index of similarity of efficiency within a region. The DMUs that are efficient in one cluster may not be as efficient as DMUs in other clusters. Relative to these they may be inefficient. Therefore disaggregating the data will cause rises in measured DEA ratings, how much will vary depending upon how wide the variation in efficiency is within each region.

These two factors mean that even though clustering takes into account variables that could not be included in the analysis, any gains in accuracy of efficiency rankings through increased DMU homogeneity is lost. It would seem that only the regional results of the single analysis can truly reveal the efficiency position within a region.

5.5.3 The effect of clustering on targets.

The targets for the labour input and quality output suggested by the clustered results are displayed in Table 5.5.3. It would be expected a priori that the targets would be lower than under the regional analysis and this is indeed found to be the case. This is because the CPOs are relatively more efficient than under the single analysis. The comparative change in regional targets from the all-inclusive to the clustered analysis is revealed in Table 5.5.4. The overall target across all DMUs has fallen from an 18.58% reduction in labour inputs to 7.8%, and the quality target has fallen from 6.72% to 3.95%.

Examining the targets by region, London still needs the largest labour input reduction, a target of 15.88% and Northern Ireland the least with 1.63%. London has the largest quality target of 6.55% but strangely the Eastern region now has the smallest target of 1.12%. An anomaly in the results is that the North-East actually has a higher quality target under the clustered analysis than under the single analysis (2.36% to 4.35%). Indicating, that relatively it must have a wide variation within the region of this output than it had when compared to all other DMUs under the single analysis, but it would seem that this isexceptional.

Clustering seems to have reduced the targets to level where they are not that significant. Indeed, excluding London, the average potential gain across all regions is 4.13% for the input and 3.01% for the output. This would indicate that the regions are reasonably homogenous in terms of their internal efficiency and are not being challenged by the reference set.

Of course targets for individual CPOs may still be quite large because some will remain relatively very inefficient. But, in general this will not be the case, nearly half the CPOs are now deemed to be efficient and do not need targets anyway (the issue of improving the efficiency of DMUs deemed to have achieved TE is dealt with elsewhere). Ganley (1989) concluded on the change in targets after clustering that 'the target may be subject to significant and unpredictable

Table 5.5.3.

Potential gain from reductions in input and increases in output, due to meeting targets in each clustered region.

		LABOUR INPUT (HOURS)			QUALITY OUTPUT (INVERTED AWT)		
		ACTUAL	TARGET	POTENTIAL GAIN (%)	ACTUAL	TARGET	POTENTIAL GAIN (%)
1	London	1058217.00	890201.69	15.88	58370.00	62195.29	6.55
2	Eastern	252761.00	243849.53	3.53	17850.00	17777.28	1.12
3	Midlands	362350.00	345843.47	4.56	23921.00	24686.51	3.20
4	North-East	319249.00	307684.28	3.62	21739.00	22685.28	4.35
5	North-West	333299.00	319125.39	4.25	23552.00	24317.60	3.25
6	South-West	301473.00	288133.14	4.42	24006.00	24751.20	3.10
7	Wales and the Marches	181522.00	172232.85	5.12	15970.00	16509.99	3.38
8	Northern Ireland	66890.00	65798.61	1.63	4778.00	4832.98	1.15
9	Scotland	337654.00	323331.85	4.24	30230.00	31495.77	4.19
10	South-East	299520.00	282597.14	5.65	20819.00	21518.93	3.36
AVERAGE POTENTIAL GAIN (%)				7.80			3.95

Table 5.5.4
Comparison of resource targets for regions and clustered regions.

		LABOUR INPUT TARGET REDUCTION (%)		QUALITY OUTPUT TARGET INCREASE (%)	
		R	C	R	C
1	London	29.48	15.88	12.79	6.55
2	Eastern	18.75	3.53	4.58	1.12
3	Midlands	12.45	4.56	5.33	3.20
4	North-East	11.42	3.62	2.36	4.35
5	North-West	13.31	4.25	4.96	3.25
6	South-West	17.24	4.42	4.21	3.10
7	Wales and the Marches	14.60	5.12	3.51	3.38
8	Northern Ireland	8.43	1.63	4.89	1.15
9	Scotland	12.48	4.24	6.09	4.19
10	South-East	12.69	5.65	6.29	3.36
AVERAGE POTENTIAL GAIN (%)		18.58	7.80	6.72	3.95

change' and 'the scale of adjustments required to improve efficiency have become unclear'. This is quite clearly the case, but in general targets have become smaller or disappeared. Whilst the changes are significant it is possible to predict that the change will be to much smaller targets. Perhaps in the clustering context the scale of adjustment is not as important as identifying those areas in which DMUs need to concentrate. Whilst their quantification is now difficult, problematic areas are at least still identified.

The problem then is that DMUs under the clustered analysis are deemed to be relatively more efficient. This is not altogether very satisfactory because this is probably not just a product of taking into account background variables but also reflects the decrease in the number of constraints. Disentangling these two factors is too problematic for true resource targets to be identified and so where it is at all possible background variables should be incorporated directly.

5.5.4 The effect of clustering on peer groups.

The peer group of an inefficient DMU are those DMUs which form that part of the frontier relative to which its efficiency is measured. The peer group is useful for improving the efficiency of an inefficient DMU because by examining their operating and management practices, new methods can be instituted to achieve best-practice (Dyson et al. [1987], Vassioglou and Giokas [1990]).

It would be thought that the clustering of CPOs into regions would identify more plausible and useful peer groups than would an all-inclusive analysis. This is because with 1281 DMUs spread over ten regions the idea of the peer group can lose its significance due to differences in operating conditions. For instance, looking at Table 5.5.5 which identifies the peer group of inefficient CPOs in Northern Ireland, it is clear that the peer group consists

Table 5.5.5.

Peer groups for Northern Ireland under a single and clustered DEA analysis.

CPO No.	TE Single Analysis	TE Clustered Analysis	Peer groups for single analysis	Peer groups for clustered analysis
987	0.971	1.000	43 288 471 707 959 997 1010	
988	0.979	1.000	43 261 641 707 716 1108	
990	0.898	1.000	64 707 959 993 1010 1079	
991	0.819	1.000	43 64 462 473 1123	
992	0.942	1.000	471 627 639 673 1010 1026 1075	
995	0.860	1.000	43 471 497 499 707 959	
996	0.873	1.000	43 931 933 1010 1026 1263	
998	0.855	0.973	43 473 933 997 1010 1079 1123	987 991 996 1010
999	0.638	0.752	43 473 931 933 1010 1026 1123	991 992 1010
1000	0.894	0.974	471 473 716 1010 1026	988 991 997 1010
1001	0.917	1.000	43 641 716 933 1010 1079 1108	
1003	0.887	1.000	43 641 707 959 1002 1100 1104	
1004	0.803	0.915	43 288 641 707 930 959 1010 1079	987 989 1002 1010
1005	0.944	1.000	43 396 471 641 716 933 1010	
1006	0.928	1.000	43 64 707 716 959 1010 1079 1102	
1007	0.851	1.000	43 64 497 707 716 933 1079 1102	
1008	0.962	1.000	497 530 641 707 1010 1106	
1009	0.957	1.000	43 530 641 707 1010 1106 1108	

of CPOs from a scattering of different regions including London. A region one would have thought that has nothing in common with Northern Ireland. On the other hand, considering Northern Ireland is such a small region so that its CPOs are being compared to hundreds of others, its peer groups do contain Northern Ireland peers. Each inefficient DMU has an average of 6 or 7 best-practice DMUs as its peer group, in general, one of them is a Northern Ireland DMU.

After clustering in Northern Ireland the peer groups are likely to be smaller because the cluster is so very much smaller. Only four of the CPOs are inefficient (See Table 5.5.5), their peer group includes 9 of the efficient DMUs. The most cited peer in the single analysis was CPO 1010, it is still the most popular and is cited by all inefficient DMUs. The peer identified by clustering is likely to be more useful than the peer identified through the single analysis. But, the actual clustering itself has presented an impediment to providing better peer groups for all inefficient DMUs. The increase in efficiency ratings and the decrease in discrimination identified in section 5.4.2. means that only four of the CPOs are identified as inefficient. Therefore these four are the only CPOs to acquire new peer groups to improve their performance. In reality, there are many more inefficient DMUs in the region than this because they were identified in the all-inclusive analysis.

The new information is still useful though. If a better peer group can be identified for just a few of the inefficient DMUs, and they are likely to be the most inefficient anyway, then this will be helpful. The problem then, is which CPOs to use as examples of best-practice for the DMUs identified as inefficient under the single analysis but efficient in the clustered analysis. An inefficient CPO in Northern Ireland for example may have as a peer, a CPO with a similar input/output mix in another region, but, even if its reasons for being inefficient were the same, it is much more practical for a Branch Manager to give close examination to a peer CPO within their own region.

Fortunately, in the single analysis, there is CPO No. 1010 from Northern Ireland which seems to be a peer for most of the inefficient DMUs in the region. As well as this CPO there are three others which are peers from the region (these are underlined in Table 5.5.5). Only 3 out of the 18 inefficient DMUs in Northern Ireland identified under the all-inclusive analysis has no peers at all in the region.

Looking at the other regional results under the single all-inclusive analysis would seem to suggest that this result is largely the same. For DMUs that are inefficient under the single analysis and efficient under the clustered analysis, so that the peer identified initially has to be used, there will in most cases for each CPO be one peer from within the region. There are 423 CPOs (33.02% of the total inefficient CPOs) that are inefficient under the single analysis and efficient under the clustered analysis, and of those 262 have no peer under the single analysis from within their own region. This means that 23.73% of all inefficient CPOs are not presented with a new peer group after clustering and have no peers from their own region under the single analysis. This does not mean that the peers that they have are completely unsuitable, they are just not as plausible as would be desirable. The fact still remains that of the initial inefficient DMUs 76.27% have a peer within their own region or are given new peer groups from within their own region after clustering.

Clustering, whilst identifying more plausible peers within a region for an inefficient CPO will not be able to do so for all (685 or 61.96% of the inefficient CPOs in this case). Under the clustered analysis they no longer need peer groups because they have become efficient. More DMUs are efficient because of the reduction in the size of the cross-section.

5.6 Conclusion.

The variables used in the DEA were defined very much by the data available but fortunately conformed to the variables thought to be most desirable in Chapter 3. The variables used were labour (in staff hours) as an input, a quality variable as an output which was average customer waiting time (in seconds), which had to be inverted for the calculation and ten transactions categories as the other outputs. These variables define the essential dimensions of the Counters transaction process, though additional information could have been yielded if data on costs of labour inputs, office space (per square metre) and the average resignation rate (as a proxy for staff turnover) were available. It was decided that the objective that most suited Counters to be incorporated into the DEA programme was one of input minimisation rather than output maximisation.

The productive efficiency of an organisation consists of pure technical efficiency (which is measured with a variable returns to scale technology) and scale efficiency (which when a constant returns to scale technology is used is amalgamated with PTE). The aspect of technical efficiency we are most interested in is PTE and it was shown that scale efficiency is not of great significance to Counters. Scale efficiency was calculated to be very high with an overall rating of 0.964. To establish which type of scale inefficiency exists for the small amount that there was, an NIRS frontier was used to establish that decreasing returns to scale is most common (72.39% of the scale inefficient having DRS). Those CPOs that have DRS are operating at below their optimal scale because the branch is not large enough. Even so, the amounts involved are so small as to make scale inefficiency an unimportant factor. A variable returns to scale technology to measure PTE was thus used throughout the study.

The overall results for pure technical efficiency at POC were not very favourable. Only 13.82% of the CPOs were efficient and the mean technical efficiency across the sample was 0.814, the worst CPO having a rating of 0.386. The targets defined by this level of efficiency that would have to be implemented to achieve best-practice across all DMUs were an 18.6% reduction in labour inputs and a 6.7% increase in the quality output. To find out why such a level of inefficiency exists, it was thought necessary to examine Counters on a regional basis. Efficiency between regions is so diverse as to imply that some unknown factors must account for the inefficiency and hence inefficiency between individual DMUs.

The London region's inefficiency was most marked and helped gain insight into the problem of regional inefficiency. One reason for inefficiency is variation in managerial quality though it is not possible to establish whether this varies on a regional basis. London has high unionisation which seems to impede the process of establishing the correct quantity of staff hours needed to be worked. In addition, there are a large number of entrenched inefficient working practices. This will then vary from region to region. Another factor in the variation of efficiency is the turnover of staff. Too high and the competence and efficiency of staff must be in question as they will have little experience, too low and flexibility of staffing is reduced. It must also be stated that absolute efficiency of those CPOs declared to be efficient is not without question as DEA is a relative measure. A degree of structural inefficiency exists due to universal inefficient working practices, higher sick leave than normal, high leave reserves, lack of computerisation etc., though it has to be said that Counters is tackling all these areas. A possible method to take into account variations in regional efficiency and hence define better efficiency ratings, resource targets and peer groups is clustering.

The CPOs were clustered by geographical area and a new efficiency frontier estimated for each region. How useful was the clustering process in providing more information to improve performance? Clustering is useful in theory, in that by grouping the CPOs by some common factor, variables that could not be incorporated because the information is not available can be taken into account. Efficiency rankings should thus be more accurate. In the case of the Post Office the variables identified were the strength of unions and the turnover of staff, clustering the CPOs by geographical location means that they are corrected for because they appear to vary from region to region (this is not to say that there are other variables which have not been identified). However, the clustered rankings really only reveal how similar the CPOs are in a region. Only a single all-inclusive analysis can reveal the true efficiency position relative to the most efficient CPOs in the UK. The situation is complicated by the fact that decreasing the size of the cross-section increases ratings. Better regional rankings may be established but as the clusters are very dissimilar in size it is not possible to compare ratings between clusters.

Peer groups identified through clustering should be more suitable because they come from the same region. But, because of the reduction in the number of inefficient DMUs due to the decrease in the size of the cross-section, only a few CPOs can be given new regional peer groups. In reality there are a much larger number of inefficient DMUs as they were identified in the all-inclusive analysis. These cannot be given regional peer groups because they are efficient in a clustered analysis. Fortunately, in most cases these DMUs will have at least one peer from the same region. Thus, the only DMUs that are helped by clustering to better identify their peer groups, and improve performance by examining best-practice, are those that remain inefficient. They may be very few in number as was shown in the case of the Northern Ireland cluster. But, it is fortunate that such DMUs are able to be allocated better peer groups because they are the most inefficient of the sample.

Clustering provides the same problems for targets as it does with efficiency ratings. The smaller the cross-section, the higher the efficiency ratings, the less resource adjustments have to be made to achieve best-practice. The targets are therefore better set under an accurate all-inclusive analysis.

Therefore, clustering to incorporate background variables does not necessarily enable better definition of ratings and targets. The decrease in the number of constraints raises efficiency levels and decreases targets, so that increased accuracy through clustering is lost. Every effort should be made to include background variables directly to produce accurate ratings and targets under a single all-inclusive analysis. This does not preclude the use of clustering for the identification of better peer groups. It will produce a peer group within a region that is more plausible and more conveniently examined. However, this is only the case for those DMUs that remain inefficient after clustering. In this study it was found that of those CPOs that were inefficient under the all-inclusive analysis, 61.96% could be given new regional peer groups.

Footnotes.

1. It can be seen that the four territories correspond quite closely to the ten regions. The London territory is the London region, the Eastern territory is the Eastern and South-East regions, the Western territory is the Midlands, Wales and South-West regions and the Northern territory is the North-East, North-West, Northern Ireland and Scotland regions.

2. As a matter of interest the ATE of the four new territories is, London = 0.705, Eastern = 0.797, Western = 0.853 and Northern = 0.878.

Chapter 6.
Sensitivity analysis: Establishing the robustness of the results.

6.1 Introduction.

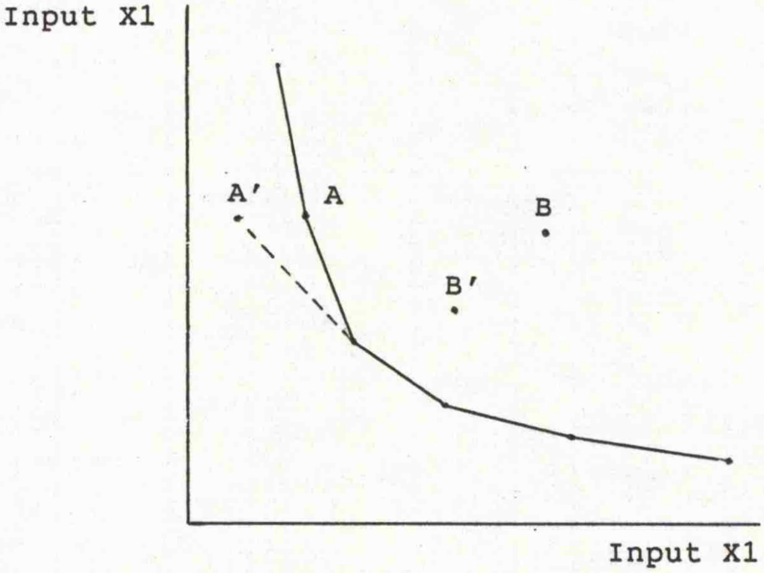
A major weakness as well as strength of DEA is the non-parametric construction of the efficiency frontier. The key feature of DEA is that the efficient frontier is formed by the outer frontiers of all DMUs actual achievements. This is in contrast to many existing techniques, such as regression analysis which seeks to average out stochastic error terms in order to estimate a pre-specified functional form for the production frontier. Unlike regression analysis, there may be no direct way under DEA of assessing whether a DMUs deviation from the frontier is statistically significant. In the absence of hypothesis testing, some other means has to be established of checking the validity of efficient DMUs. In order to establish the reliability and robustness of the position of the efficiency frontier under DEA, and implied deviations from it, it therefore becomes desirable to carry out an analysis of the sensitivity of the results of DEA to a number of underlying factors.

The first factor which necessitates the use of sensitivity analysis, is caused by the fact that DEA interpolates from DMUs' actual achievements. The use of interpolation of actual attainments makes the DEA results potentially very sensitive to the misspecification of 'extreme' DMUs' achievements and resource use. The consequences of this are important because they will form part of the efficiency frontier for assessing the efficiency of other units. The most extreme example of this is if a DMU concentrates on one particular output to the exclusion of other outputs, and is the only unit to do so, it will automatically be deemed to be efficient. The DMU will form part of the efficiency frontier on its own. As there are no comparable DMUs on either side of it with an exceptional mix of resources or peculiar choice of outputs, so the interpolation assumption of DEA is likely to yield high efficiency ratings for such DMUs. But, their efficiency has been over-estimated, DEA is over generous in that inefficient DMUs are quite conclusively inefficient, but, efficient DMUs are not necessarily so.

It is very important that the validity of the efficient DMUs ratings that define the frontier can be relied upon. Firstly, because those DMUs that define the frontier form the peer group and determine the ratings of the inefficient DMUs. Secondly, because they will be used to improve the efficiency of the inefficient DMUs. As explained in the previous chapter this can be done in two complementary ways. DEA can show an inefficient organisation its best-practice peer group of efficient DMUs, it can then by examination and adaptation of their best-practice managerial procedures improve its own performance. In addition, DEA can improve performance by quantifying for an inefficient unit the reduction in inputs or increase in outputs that is theoretically possible relative to other DMUs. Thus, examining efficient DMUs for their best-practice procedures and using them to set targets means that their validity is quite crucial.

It should be pointed out that a similar problem arises from the interpolation characteristic of DEA when there are errors in data such as misreporting or miscoding. Since it is an extremal method there is particular sensitivity to misspecification in certain situations. Consider Figure 6.1.1 which plots the normalised outputs of a number of comparable DMUs. In this simple case, each DMU uses two variable inputs to produce a single output. Errors in data affect the results of the DEA technique in one of two ways. In an isolated fashion when the error occurs in an observation off the frontier or in an interactive fashion when the error occurs in data pertaining to an efficient DMU (in contrast to regression techniques, where all errors offset the estimated regression line and hence each and every observation). Interactive errors in DEA are the most troublesome, since they alter the shape of the isoquant, which can affect efficiency scores. To illustrate assume the observation A in Figure 6.1.1 is either misreported or miscoded as A′. The

Figure 6.1.1.
Illustration of two types of computer error resulting from errors in data.



resulting isoquant may be very different from the true isoquant, and will cause many DMUs to be misclassified as less efficient than they actually are. In contrast, if an inefficient DMU is misreported as B′ rather than B, the resulting error will be local and will not disrupt other efficiency scores. Whilst the resulting problems are exactly the same, it is extreme DMUs we will be concerned with in the following sections, it is assumed that there are no interactive data errors. Thus, misspecification can have severe consequences for DEA. Unlike mean-based methodologies, where on average data errors will have broad though small quantitative impacts, using extremal techniques such as DEA the impact is magnified.

A second consequence of the extremal construction of the DEA frontier is that it will always perfectly explain the relationships between the variables. As DEA is non-parametric and has no specified functional form, there is no way of assessing the relative strengths of different model specifications. Application of DEA provides no information regarding the structural definition of the problem. It fails to generate any measures either of statistical or of causal relationships between the inputs and outputs of the observed DMUs. Therefore, the selection of variables to be included in a DEA model is not at all straightforward. The process of selecting variables and the problems involved were looked at in Chapter 3 and will not be discussed here.

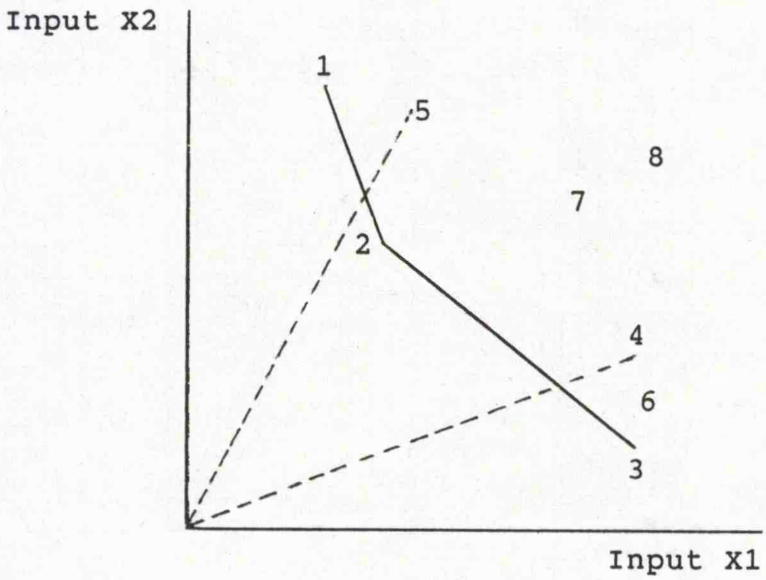
Therefore, due to the fact that the DEA frontier is based on extremal observations rather than on average ones the frontier is easily distorted, through accident or design, by outliers. There needs to be some form of sensitivity analysis to establish the reliability of the position of the efficiency frontier under DEA and establishing the sensitivity of the results of DEA to efficient DMUs that are outliers. This can be done in a number of ways and this chapter examines them in turn in the Post Office Counters context. They include examining efficient DMUs' peer group citations (PGCs), how well-rounded their performance is and through sensitivity analysis.

6.2 Examining the efficient DMUs by looking at their peer group citations.

With respect to establishing the reliability of the frontier, given that extreme DMUs may distort it, efficient DMUs may have to be examined to confirm that their rating has not been over-estimated. Mayston and Smith (1987) suggest that 'an important supplementary measure in assessing the robustness of the result is the number of inefficient authorities for which the authority forms the efficient frontier'. So that, 'if this number is high, the authority is genuinely efficient with respect to a large number of authorities'. The number of peer group citations by inefficient DMUs attributed to an efficient DMU that it forms part of their frontier would thus seem to be a good indicator of the 'genuineness' of a technically efficient DMUs rating.

Cubbin and Ganley (1987) suggest that the number of peer group citations that a DMU receives is only useful as a test of robustness if inefficient DMUs are distributed evenly throughout the feasible production space. In Figure 6.2.1 a hypothetical DEA isoquant is constructed for two inputs and one output. DMU performance is deliberately bunched in the

Figure 6.2.1
Relating a DMUs position on the isoquant to its number of citations.



South-West of the feasible set. It shows that DMUs 1, 2 and 3 are best-practice with efficiency scores equal to unity. DMUs 4, 5, 6, 7 and 8 are inefficient relative to 1, 2 and 3 and will have targets set for them for the resources X1 and X2. The reference or peer DMUs for most of the inefficient producers (other than DMU 5) are DMUs 2 and 3 (See Table 6.2.1). For example the target for DMU 4 is a weighted average of 2 and 3. DMU1 has relatively unusual input proportions and forms part of the frontier on its own. It is not cited frequently as part of the peer group of any of the DMUs except DMUs which also have unusual input proportions such that its target is an interpolation of DMUs 1 and 2. As can be seen in Table 6.2.1 DMU1 is not

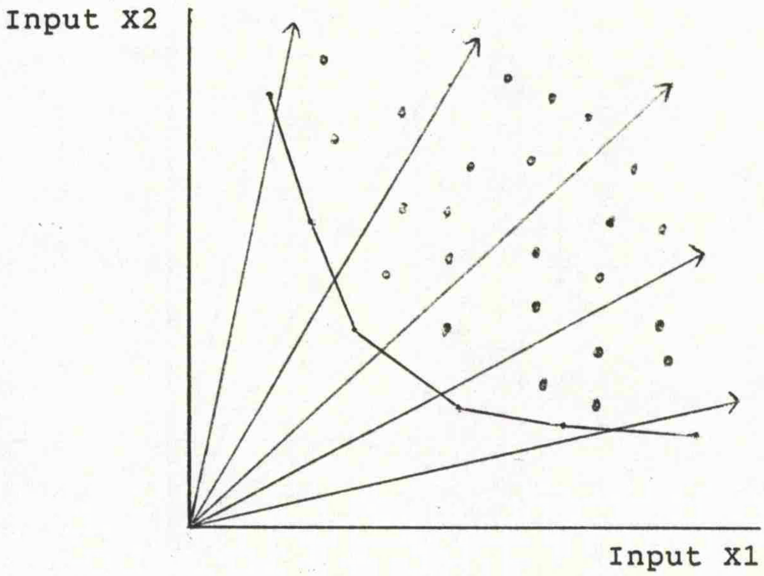
Table 6.2.1
Peer groups and citations for DMUs in Figure 6.2.1.

DMU	Peer Group	Citations in Peer Group
1	1	1
2	2	5
3	3	4
4	2,3	0
5	1,2	0
6	2,3	0
7	2,3	0
8	2,3	0

cited very often in the peer group citation index. Cubbin and Ganley say that this is because there is a smaller probability of finding inefficient DMUs. Their premise is that as each best-practice authority has an equal probability of citation, if inefficient DMUs are spread evenly through the feasible production space a high number of citations only implies comparability with a large number of inefficient DMUs. They conclude that if there are a high number of citations 'it is not clear that this also conveys information on the intrinsic quality (or "genuineness" - as argued in ibid) of different examples of best-practice'.

Cubbin and Ganleys re-interpretation of peer group citations for efficient DMUs merely as an indicator of comparability, rather than the intrinsic quality of an efficient DMU, is not altogether accurate. In a statistical sense, we can have much greater confidence in those segments of the frontier which have large peer groups. Consider Figure 6.2.2 which plots the normalised outputs of a number of comparable DMUs. Note that each extreme technology (the maximum and minimum X1/X2) lies on the isoquant and that it is thus efficient. If the quadrant is divided into cones formed by rays from the origin, at say, successive fifteen-degree intervals, all DMUs in the sample or on the isoquant, will lie within one of the cones. Is there any statistical difference between a segment of the isoquant that lies within a cone containing few observations (extreme

Figure 6.2.2.
Ineffficient DMUs related to different facets of the frontier.



technologies) and a segment that lies within a cone containing many observations (common technologies). How certain can we be of those segments of an isoquant constructed with very limited information? Statistical confidence depends on sample size, the larger is the sample in a particular area, the closer is the sample frontier likely to be to the true frontier. In a statistical sense then, a DMU that forms part of the frontier that contains a large number of observations should be more robust than one with only a few observations. The question then, is how many peer group citations would constitute a high enough number to be sure that the sample frontier approximates the true frontier. Unfortunately, the answer is unclear.

Examining the peer group citations ranking (See Table 6.2.2) it is apparent that a large proportion of the 177 DMUs have a very low number of citations. 46 or (25.99%) of the efficient DMUs have only one citation, which is themselves. They are comparable to no other DMU and they are so unusual that they form part of the frontier on their own. There are a large number of DMUs with low PGCs, 96 (54.24%) of them have between 1 and 10 citations. The highest number of PGCs is 784 gained by DMU 43, this is way out in front as the next lowest PGC is 379. 26 of the DMUs (26.9%) have PGCs in the hundreds. If we are to have statistical confidence in any of these DMUs it must at least be in these because they form part of the frontier for such a large number of DMUs. Though this is quite arbitrary because it is not known whether this is a big enough sample, and as will be seen later it is not very easy to be sure of even these.

6.3 Well-rounded performance.

A further way of validating the efficiency of DMUs identified in the previous section as needing further scrutiny is evaluating whether they have 'well-rounded performance'. That is, a unit whose efficiency is based fairly evenly on all its inputs and outputs. An efficient unit with well-rounded performance is relatively efficient when all aspects of its performance are taken into account rather than just a small subset of them. That is, that a DMU is not 'extreme' and has achieved its efficiency by devoting its resources almost exclusively to one output rather than by performing all relevant operations efficiently.

The idea of using the breadth of an efficient DMUs performance to establish the robustness of its rating, and the means of establishing this, were suggested by Dyson, Foster and Thanassoulis and discussed in Chapter 4. They suggest examining the virtual inputs/outputs (these are the

Table 6.2.2.

Ranking of Peer Group Citations.

NO	PGCs	NO	PGCs	NO	PGCs	NO	PGCs
102	1	1167	1	1129	8	621	58
185	1	1169	1	848	8	935	59
186	1	1228	1	989	8	396	61
265	1	188	2	599	8	189	63
269	1	1136	2	867	10	275	68
320	1	608	2	62	10	716	68
326	1	628	2	1075	11	930	69
365	1	1124	2	993	11	1123	73
408	1	657	2	1098	11	462	75
411	1	742	2	655	12	296	84
432	1	1044	2	638	12	288	84
435	1	731	2	1087	12	372	88
457	1	1065	2	1196	13	673	89
480	1	300	3	283	13	614	90
500	1	951	3	267	14	677	91
528	1	780	3	160	15	690	96
564	1	704	3	1043	16	813	103
584	1	504	3	703	16	529	108
585	1	31	3	637	16	471	110
590	1	467	3	807	18	707	111
615	1	287	4	1100	18	1108	119
644	1	870	4	880	18	1263	126
654	1	1127	4	262	19	433	142
656	1	495	4	1225	20	1010	155
689	1	1002	4	1209	21	261	168
694	1	756	4	361	22	64	168
708	1	536	4	475	22	478	175
734	1	1089	4	429	22	1079	203
741	1	578	5	571	23	678	203
759	1	494	5	1145	23	959	211
765	1	747	5	650	23	427	217
878	1	636	5	1104	24	1026	238
879	1	688	5	1106	28	641	257
899	1	997	6	1102	31	593	278
910	1	586	6	639	31	933	285
914	1	822	6	811	31	530	295
963	1	1222	6	905	32	497	303
972	1	800	6	629	32	473	305
994	1	279	6	527	32	757	305
1024	1	579	7	491	38	627	358
1058	1	718	7	751	39	931	379
1088	1	187	7	730	41	43	784
1115	1	587	7	466	51		
1135	1	1039	8	604	54		
1152	1	667	8	499	55		

product of the weights and the inputs/outputs) of an efficient DMU as these can be used to establish whether it has well-rounded performance. This can be done because they show what proportion of a DMUs efficiency rating is contributed by each input/output. A wide range of virtual inputs/outputs with no single one dominating, shows that efficiency is due to a wide range of factors and efficiency is not attributable to the influence of an extreme value in one dimension. Well-rounded performance is not only good for this reason, a DMU which only derives its efficiency from one dimension may not be that efficient in that dimension anyway. A relatively efficient unit with a high particular virtual input or output may not be alone in performing that operation efficiently. Units with well-rounded performance may be performing those operations efficiently. Indeed, units with well-rounded performance may be performing the operations relating to the inputs and outputs in question more efficiently than units whose performance rating is concentrated on them, because concentration of resources to achieve limited outputs is less likely.

There is a problem in deciding what constitutes well-rounded performance in any specific DEA run. This will obviously depend on the number of inputs/outputs. It would seem to me to be commonsense that this should at least mean, as a minimum, that the largest virtual input/output should not contribute to more than 50 per cent of the efficiency rating and that at least half the virtual inputs or outputs contribute to the efficiency rating in the possible dimensions. Of course a lot depends on the context, as factors like the relative importance of inputs/outputs might vary.

With respect to the Post Office Counters run, virtual inputs/outputs were calculated for the 177 efficient CPOs identified in the previous section and that are perhaps not truly efficient. Since there was a single input, Labour, in the DEA run, total virtual input for each unit is always attributable to that unit, will always be 100 per cent and therefore is of no significance. Unfortunately, because the input minimisation, variable returns specification is being used, the virtual outputs do not add up to 1 and so it is not possible to ascertain exactly how much of a units efficiency is attributable to that output.

Nevertheless, well-rounded performance can still be established by examining a DMUs weights. Something like this is suggested but not explored by Ganley (1989), as a means of ordering untied DMUs. Examining the weights does not establish what contribution each input/output makes to efficiency per se, because the weights are scale dependent. DMUs using lower levels of inputs and outputs, the smaller DMUs, will generally have larger weights than larger DMUs. Therefore, two DMUs operating with the same production process but on different scales will place positive weights on the same inputs and outputs, but they will select appropriately scaled values that differ substantially. This is not a problem as the weights can be just as good an indicator of how well-rounded a DMU is. This is because it is the number of output dimensions that efficiency can be attributed to rather than the size of the contribution in each dimension that is important. Thus, although the size of the weights does not tell us anything, the number of weights shows how well-rounded the performance of a DMU is and hence that its efficiency is not derived from a narrow range of activities.

Given the rule-of-thumb suggested earlier, in the CPO context, well-rounded performance can be taken to mean that a DMU has weights in the majority of its output dimensions (the weights in the input dimensions are of no consequence because there is only one input and therefore always one weight). This means that an efficient DMU must have weights in at least five of the ten outputs.

In Table 6.3.1 the 177 technically efficient DMUs are ranked according to how many weights they have in the output dimensions. It is clear, that even on the criteria defined above, that not that many DMUs can be deemed to have well-rounded performance. Only 38 CPOs have five or more weights, which means that 78.53% have not. The highest number of weights gained by a DMU, (and it is the only one to reach that many), is 7 by DMU 1228. As the highest number of weights attainable is 10, perhaps it would be fairer to define those that don't have well-rounded performance as those DMUs that have less than the average number of weights. The average number of weights is 3.28, the number of DMUs with weights less then this is 96 (54.24%).

An aspect of well-roundedness that has not been discussed so far is whether it is more important to have output weights in some dimensions rather then others. This is of course against one of the precepts of DEA in that outputs should be of equal importance. Nevertheless, in the Counters context it is very important to achieve high quality of service and this should have prominence above other outputs. Its importance is underlined by the fact that it is controllable, and is also in a direct trade-off between itself and the labour input. It is interesting to note that of the 24 efficient DMUs that weight this output, most of them are also quite well-rounded. The average number of weights for those DMUs that weight quality is 4.79. This is substantially more well-rounded than the overall average, but not completely surprising because a DMU with a low AWT that is also efficient, must be dealing with all of its transactions efficiently. These observations are useful because it reveals the fact that those DMUs that weight efficiency are more robust and are therefore better peers.

A possible reason why the average number of weights is not very high (only 3.04 for those that don't weight quality) is that the level of disaggregation of output transactions may be too high. Perhaps smaller product groups should have been chosen. Though it is still the case that even at the present level of disaggregation the number of weights still provides an indicator of well-roundedness of performance.

Table 6.3.1.
Number and type of output weights for efficient DMUs.

1 Average Waiting time (inverted).
2 Stamps.
3 Vehicle Licenses.
4 Visitors Passports.
5 Child Benefit.
6 Unemployment Benefit and Income Support.
7 Pensions.
8 Giro's Deposits.
9 Giro's Withdrawals.
10 All Other Transactions

CODE NO	WEIGHTS 12345678910	NO	CODE NO	WEIGHTS 12345678910	NO
31	0000011111	5	471	0000000100	1
43	0100000000	1	473	0000000100	1
62	0000000010	1	475	0000010010	2
64	0000000001	1	478	0010000000	1
102	0000000010	1	480	0011101110	6
160	0000000110	2	491	0011011001	5
185	0010001001	3	494	1001101101	6
186	0010101011	6	495	0110011000	4
187	0110001010	4	497	0010001001	3
188	0110101010	5	499	0001001100	3
189	0000001011	3	500	0011000101	4
261	0000100110	3	504	1010001011	5
262	0000000011	2	527	0010001001	3
265	1110011010	6	528	0001000111	4
267	0100100010	3	529	0011100000	3
269	0100011010	4	530	0010000000	1
275	0000000010	1	536	0011010000	3
279	0000010010	2	564	0110101001	5
283	0000011010	3	571	0101001001	4
287	0000010010	2	578	0001111000	4
288	0000001010	2	579	0101000111	5
296	0000000110	2	584	0111011000	5
300	1010001000	3	585	1100000111	5
320	0110001010	4	586	0111100001	5
326	1010101011	6	587	0010001001	3
361	0010001001	3	590	1100100111	6
365	0001100100	3	593	0001000000	1
372	0011000010	3	599	0011001001	4
396	0000100000	1	604	1101000010	4
408	0000100011	3	608	0111001001	5
411	0111101011	7	614	0011001101	5
427	0011000000	2	615	0111010000	4
429	0110001001	4	621	0100001001	3
432	0000101011	4	627	1010011000	4
433	0010101001	4	628	1111100010	6
435	0001000010	2	629	0110101000	4
457	1001000110	4	636	0111000000	3
462	0000000001	1	637	0010000000	1
466	0000011010	3	638	0010101000	3
467	0001001100	3	639	0011000000	2

Table Continued.

641	0000110000	2
644	1101101000	5
650	0001000000	1
654	0111100101	6
655	0010001011	4
656	0110010001	4
657	0111001101	6
667	0011001100	4
673	0010001000	2
677	0001000100	2
678	0001001010	3
688	0001000010	2
689	0101000100	3
690	0001100110	4
694	1110000011	5
703	0110010100	4
704	0000000010	1
707	0000010100	2
708	0110010000	3
716	0000000100	1
718	0000100000	1
730	0010000000	1
731	0101101001	5
734	1100100101	5
741	0100011001	4
742	0110011011	6
747	0101000011	4
751	0000111010	4
756	1110000011	5
757	0000001001	2
759	0101001001	4
765	0100101001	4
780	0001101010	4
800	1010000010	3
807	0010000010	2
811	0001001011	4
813	0010001001	3
822	0010001000	2
848	1000000100	2
867	0100001001	3
870	0110001000	3
878	0110001010	4
879	1011010011	6
880	1000001011	4
899	1110100010	5
905	0010000100	2
910	0110001001	4
914	0101101110	6
930	0000101001	3
931	0000001000	1
933	0000001001	2
935	0001100001	3
951	0010000100	2
959	0000001000	1
963	1100001001	4
972	0000001001	2
989	0000000100	1
993	0010101101	5
994	1100010001	4
997	0000011101	4
1002	0000010000	1
1010	0000011000	2
1024	0100000001	2
1026	0000000000	0
1039	1001011001	5
1043	0100001000	2
1044	0110011001	5
1058	0100001101	
1065	0000011001	3
1075	0010011000	3
1079	0000010001	2
1087	0000011011	4
1088	0001001011	4
1089	0000011000	2
1098	0000011000	2
1100	0000011000	2
1102	0000001001	2
1104	0001101001	4
1106	0000111001	4
1108	0000010000	1
1115	0100101000	3
1123	0000000001	1
1124	0000010011	3
1127	0011001101	5
1129	0000011001	3
1135	0011100001	4
1136	0010001101	4
1145	0001000010	2
1152	0001000101	3
1167	0010001101	4
1169	0111010001	5
1196	0011000001	3
1209	0000001000	1
1222	0110001100	4
1225	0111100011	6
1228	1110101011	7
1263	0100001010	3

6.4 Comparison between peer group citations and well-rounded performance.

So far the results have not been very conclusive. We have been concerned to establish the robustness of efficient DMUs so that they can be used to set standards of best-practice. The extremal nature of the frontier construction means that DMUs can be deemed efficient on the basis of one extreme observation, whether this is an accident, or through concentration of resources on one output. The 48 out of the 177 efficient DMUs that have only one peer group citation form part of the frontier on their own, and are not used to define another DMUs targets, therefore we needn't be to concerned about them. The most worrying are those DMUs with low PGCs, they are under suspicion but form the peer group of some DMUs. At the other end of the scale those DMUs that have large numbers of citations and are highly comparable may have some validity in a statistical sense, the problem being in deciding how many is enough for them to be valid. Examining the weights of the efficient DMUs has not been very enlightening because on any criteria over half of them do not have well-rounded performance. Perhaps more light could be thrown on the situation by comparing the ranking determined by PGC's with the number of weights that each PGC has.

Comparison of the peer group citations and the number of output weights for each CPO are shown in Table 6.4.1 Through casual observation of the result it is immediately obvious that efficient DMUs with well-rounded performance have low numbers of peer group citations and vice-versa. The reasons for this will now be discussed.

Normally, if DMUs are spread throughout the feasible production space, well-rounded performance would result in a large number of peer group citations. This is because the more dimensions that are dominated, the more DMUs fall under its portion of the frontier. But, DMUs appear to be clustered near the input/output frontier of efficient DMUs that are producing a lot of one or two outputs. The inefficient DMUs that are contributing to the high peer group citations must be using the same production process but not as efficiently. Even though this is a common technology it involves what we have described earlier as 'extreme' resource use. That is, a large number of DMUs concentrate on one or two outputs. This may mean that they produce these outputs efficiently, but if all input resources are concentrated on a few outputs it will be to the detriment of the production of all others, so the DMU may not be truly efficient. The frontier is being dominated by a DMU with an output-mix that is extreme in some dimension, the fact that it forms the peer group for a large number of inefficient DMUs just means that others are exhibiting similarly bad behaviour, in a cluster in relation to the frontier. Thus, DMUs that do not have well-rounded performance have high peer group citations.

It could be said that even though well-rounded performance of a DMU is important, it is not that relevant in the POC context. When there are a large number of non-controllable output variables in a DEA and there is no scope for determining how much of each output is produced by managers. Concentration of inputs on a particular output (for whatever reason) cannot take place. A lot of one output being produced just reflects a high demand for that output, or in the case of Counters, a large number of transactions of one type being processed. Nevertheless, even though there is no volition on the part of managers, the results are still being distorted. There are two possible solutions to the problem of 'extreme' DMUs in this context. The transactions could be amalgamated into similar products, thus alleviating some of the problem, it is too late at this stage because the analysis has taken place. Alternatively, the problem DMU, or one of its outputs could be excluded from the analysis. These possibilities are examined in the next section.

Table 6.4.1.

Comparison of peer group citation ranking and numbers of virtual dimensions.

CODE NO	PGCs	WEIGHTS	CODE NO	PGCs	WEIGHTS
102	1	1	731	2	5
185	1	3	742	2	6
186	1	6	1044	2	5
265	1	6	1065	2	3
269	1	4	1124	2	3
320	1	4	1136	2	4
326	1	6	31	3	5
365	1	3	300	3	3
408	1	3	467	3	3
411	1	7	504	3	5
432	1	4	704	3	1
435	1	2	780	3	4
457	1	4	951	3	2
480	1	6	287	4	2
500	1	4	495	4	4
528	1	4	536	4	3
564	1	5	756	4	5
584	1	5	870	4	3
585	1	5	1002	4	1
590	1	6	1089	4	2
615	1	4	1127	4	5
644	1	5	494	5	6
654	1	6	578	5	4
656	1	4	636	5	3
689	1	3	688	5	2
694	1	5	747	5	4
708	1	3	279	6	2
734	1	5	586	6	5
741	1	4	800	6	3
759	1	4	822	6	2
765	1	5	997	6	4
878	1	4	1222	6	4
879	1	6	187	7	4
899	1	5	579	7	5
910	1	4	587	7	3
914	1	6	718	7	1
963	1	4	599	8	4
972	1	2	667	8	4
994	1	4	848	8	2
1024	1	2	989	8	1
1058	1	4	1039	8	5
1088	1	4	1129	8	3
1115	1	3	62	10	1
1135	1	4	867	10	3
1152	1	3	993	11	5
1167	1	4	1075	11	3
1169	1	5	1098	11	2
1228	1	7	638	12	3
188	2	5	655	12	4
608	2	5	1087	12	4
628	2	6	283	13	3
657	2	6	1196	13	3

Table 6.4.1 continued.

CODE NO	PGCs	WEIGHTS
267	14	3
160	15	2
637	16	1
703	16	4
1043	16	2
807	18	2
880	18	4
1100	18	2
262	19	2
1225	20	6
1209	21	1
361	22	3
429	22	4
475	22	2
571	23	4
650	23	1
1145	23	2
1104	24	4
1106	28	4
639	31	2
811	31	4
1102	31	2
527	32	3
629	32	4
905	32	2
491	38	5
751	39	4
730	41	1
466	51	3
604	54	4
499	55	3
621	58	3
935	59	3
396	61	1
189	63	3
275	68	1
716	68	1
930	69	3
1123	73	1
462	75	1
288	84	2
296	84	2
372	88	3
673	89	2
614	90	5
677	91	2
690	96	4
813	103	3
529	108	3
471	110	1
707	111	2
1108	119	1
1263	126	3
433	142	4
1010	155	2

CODE NO	PGCs	WEIGHTS
64	168	1
261	168	3
478	175	1
678	203	3
1079	203	2
959	211	1
427	217	2
1026	238	0
641	257	2
593	278	1
933	285	2
530	295	1
497	303	3
473	305	1
757	305	2
627	358	4
931	379	1
43	784	1

6.5 Sensitivity analysis.

Another way of testing the robustness of a DMUs apparent efficiency is through sensitivity analysis. It would seem necessary for this to take place as DEA results rely heavily on the figures from a small subset of the data. Mayston and Smith (1987) suggest that a robustness indicator of a DMUs apparent efficiency is the extent to which the omission of just one input or output would render the DMU inefficient. Sensitivity analysis can be conducted by:

a. leaving out an input from the analysis,
b. leaving out an output from the analysis,
c. leaving out an efficient authority from the analysis.

a. and b. would allow validation of individual DMUs whilst c. would gauge their impact on the results of all other DMUs.

Mayston and Smith conducted sensitivity analysis, illustratively, on an education dataset by the procedures outlined above, with results that gleaned more information about the DMUs. Some other work has already been carried out in this area. Charnes, Cooper, Lewin, Morey and Rousseau (1985) conducted a study to find out the affect of altering the outputs of one efficient DMU on the status of inefficient DMUs. Sexton et al. (1986) demonstrate the effect of a data error that increased the value of a single output. They use data based on the Program Follow Through sites reported by Charnes et al. (1981). This error will reduce the efficiency scores of some DMUs and it was found that on average, measured efficiency fell by 12%.

To determine whether highly efficient DMUs (defined by having high PGCs) are robust, sensitivity analysis must be conducted by excluding those output dimensions that define their efficiency, (inputs cannot be excluded in the POC case because there is only one). These outputs are identified by seeing whether they have been given weights within the DEA programme. For instance, DMU No.43 has only one weight and it is on output No.2, the sale of stamps. To test its robustness, this output should be excluded and the programme run again. If it is highly inefficient after the analysis, then it cannot be considered robust.

Sensitivity analysis was conducted in this way on those DMUs who had more than a hundred PGCs (an arbitrary cut-off point but includes, undeniably, DMUs with high PGCs). Not just a single output, but all outputs were excluded if there was a weight in that dimension. The outputs could have been excluded one at a time in each dimension, for each DMU. But, as these DMUs had only a few weights and did not have 'well-rounded performance', it was felt necessary to excluded all weighted outputs.

DMU No.43 was found to be highly inefficient after excluding its dominant output and its rating fell the most (from 1 to 0.440). The reason why this was the case is partly due to a fault in the way the data on stamps is aggregated. The size of this output is determined by the value of stamps sold, rather than the number of stamp transactions. This will distort the values of the output if the number of transactions is low, but the monetary value is high. In DMU No.43's case it is in an area of London that requires large bulk sales of stamps, hence inflating the figure. This is what makes DMU No.43 an outlier.

It was found that with the exclusion of 1 or more outputs, efficiency levels of most of the efficient DMUs fell below 1 (see Table 6.5.1). Of the 25 DMUs, only 6 were efficient after sensitivity analysis. The fall in efficiency does not seem that dependent on how many outputs were excluded. On the basis of their PGCs, these DMUs should be highly robust, but it is now clear that their efficiency rating is derived from just a few dimensions of their output. This does not necessarily invalidate the inclusion of the DMUs that become inefficient. As Mayston and Smith point out it is the extent by which they become efficient that is important. On this basis it is far more likely that DMU No.43 should be excluded from the analysis if after only

Table 6.5.1.
Change in efficiency rating after excluding weighted outputs for those DMUs with high PGCs.

CODE NO	PGCs	NO. OF WEIGHTS	NEW TE RATING
43	784	1	0.440
931	379	1	1
627	358	4	0.908
757	305	2	1
473	305	1	1
497	303	3	0.811
530	295	1	0.899
933	285	2	0.909
593	278	1	0.752
641	257	2	0.757
427	217	2	0.675
959	211	1	0.883
1079	203	2	0.843
678	203	3	0.679
478	175	1	0.874
261	168	3	0.678
64	168	1	0.448
1010	155	2	1
433	142	4	0.733
1263	126	3	0.856
1108	119	1	1
707	111	2	1
471	110	1	0.911
529	108	3	0.706
813	103	3	0.861

one output is left out its TE falls to 0.440, than DMU No.627 whose efficiency only falls 0.908 after four outputs are left out. This highlights the subjective element in modifying the DEA model. There are no hard and fast rules about the validation of certain DMUs after sensitivity analysis so there is scope for informed judgement to take place.

Leaving out a suspect DMU and examining the change in efficiency ratings is the other aspect of sensitivity analysis. Evaluating the extent of the affect this would have on the results is complex in a DEA analysis the size of the POC one. The DMUs which could possibly be affected are those that form the peer group for the excluded unit. The expected result would be that the efficiency of some peers would rise. This is because they have been measured unfairly against an 'extreme' DMU's input/output.

This form of sensitivity analysis was tried to establish its usefullness, with the Counters data, by just excluding DMU No.43 and seeing if there was any significant impact on the ratings of other DMUs. It was found that large numbers of DMUs had improved their efficiency ratings, some quite dramatically. Overall, 741 (56.33%) DMUs had measured efficiency improvements, though most of these changes were quite small. DMU No. 87's technical efficiency rose the most, from 0.542 to 0.846. This provides further evidence that an output of DMU No.43 should be excluded, or that it should be left out of the analysis completely because of its negative distortionary effect on the results of other DMUs. That this effect exists, confirmed by the fact that ATE over the whole sample increased from 0.814 to 0.822 when the DMU was excluded, highlights the sensitivity of the results to outliers. If the exclusion of just one DMU out of a sample of 1281 can cause such dramatic results then great care should be taken in the construction of the model, the selection of DMUs and inputs/outputs.

6.6 Conclusion.

The purpose of this chapter has been to establish the robustness of efficient DMUs so that we can be sure that they are reliable enough to be used to set standards of best-practice. Efficiency ratings of 1 may not be valid because an 'extreme' DMU may be concentrating its resources on only a few outputs. In this sense, the use of Counters data to explore this issue must largely be viewed as illustrative. Nine out of the ten outputs are non-controllable transactions, so the output-mix is not in danger of being distorted by managers concentrating resources on particular outputs. An unusual DMU in this context is just dealing with a lot of one particular type of transaction.

The results were examined on different criteria to determine their robustness. The number of peer group citations for efficient DMUs were examined and also how well-rounded the performance of each efficient DMU is. This did not produce clear results, and there is some debate about the interpretation of PGCs as an indicator of robustness. Also, well-rounded performance, defined by how many weights (rather than evenly distributed virtual outputs) exist in the output dimensions was difficult to establish. This is because the vast majority of CPOs had low numbers of weighted outputs. It could be the case that the transaction groups need to be brought down to a lower level of aggregation

Comparing the efficient DMUs PGCs to the number of weights each have was quite revealing for the Post Office data. It transpired that in general the higher the number of PGCs, the less well-rounded performance is. This is because the DMUs with high PGCs are forming the frontier for a large number of others because of their extreme resource use. In this context the PGC measure is shown to be very reliable as a robustness indicator.

Sensitivity analysis was used to validate those DMUs that appeared to be unusual and to assess their impact on the efficiency of other DMUs. Examining DMUs with high PGCs by excluding those output dimensions which define their efficiency produced interesting results. On the criterion that their robustness is determined by the degree by which they become efficient, a few could not be considered good peers and should be excluded. Though because it has to be taken into account how much their efficiency changes and how many outputs were excluded there is a certain amount of arbitrariness about which ones are considered to be robust. Excluding a DMU that could be considered to be non-robust to assess its impact produced convincing evidence of the impact of an outlier. DMU No.43 had a very large peer group and when it was excluded, and the programme rerun, most of them registered efficiency improvements. The rise in efficiency after this form of sensitivity analysis highlights the fact that DEA is very susceptible to distortion from data inaccuracies or 'unusual' DMUs. The far-reaching impact of an outlier shows that care should be taken in the construction of the original model. Thus, due to the extremal way in which it constructs the frontier, it has to be accepted that DEA is not an infallible performance measure, unusual DMUs will distort the frontier. There is a certain amount of scope for skilled informed judgement in the setting up of the model and the interpretation of the results. In addition, results have to be related to DMUs outside the DEA context.

Chapter 7.
Measuring the relative technical efficiency of the area electricity boards using data envelopment analysis. A dynamic approach.[1]

7.1 Introduction.

This chapter contains one of the first British studies using a dynamic or time-series approach to applying Data Envelopment Analysis to measure relative technical efficiency. The dynamic approach was suggested, though not applied, by Silkman et al. (1986). The technique will be applied to the twelve Area Electricity Boards of England and Wales over a twenty year period. The aim of this chapter is to draw conclusions about the level and distribution of the technical efficiency of the Area Boards, individually and as an industry. In addition, to evaluate different aspects of the technique.

The DEA model assumes that all observations relate to a single common time period on a cross-section of data. Silkman suggests that some efficiency evaluations require the inclusion of multiple time periods. The proposed approach is to pool all time periods and perform one DEA in which each decision-making unit is present T times, where T is the number of time periods. The efficient DMUs will then be some mixture of the DMUs from the different time periods. A possible motive for using this approach might be the analysis of the efficiency effects of a particular policy initiative or intervention. In this situation the effect of the policy initiative can be gauged by the prevalence of post-initiative DMUs amongst the efficient DMUs and by the average efficiency change between the two time periods. This obviously has relevance to privatisation of the Area Electricity Boards. There is scope for a future study to evaluate the effect of the change of ownership on their efficiency using this method. For now, this preliminary study will evaluate efficiency up until 1989 using the dynamic DEA approach.

A related approach to the dynamic approach is 'window' analysis. It is a particular kind of application of DEA and was introduced by Bowlin (1987) and Charnes, Clark, Cooper and Golany (1985). It also involves comparing DMUs through time by treating organisations in different time periods as separate DMUs and then applying DEA. The difference, is that subsets of the data are examined by a moving window, which is constructed in a way that provides overlaps and checks on DMU behaviour over a period of time. Trends and behavioural properties can be made clear and checks made for stability within the data.

DEA will be applied to the Area Boards to show how their efficiency has changed over time. To confirm the validity of any trends revealed, a separate DEA run will be performed on each Area Board over the twenty year period. Window analysis will be applied to the data set to check for stability within the data. Its limitations will be evaluated when compared to the 'open window' approach of DEA's dynamic application. In years in which Area Boards are efficient, examining virtual inputs and outputs shows how the efficiency of the unit is derived. Thus, it is possible to deduce the reasons why efficient DMUs are efficient. Examination of an Area Boards virtual inputs/outputs through time can also reveal the trend in efficiency derivation.

The chapter begins in section 7.2 with a description of the reasons why a comparative analysis of Area Boards' performance is useful. This is followed in section 7.3 by a discussion of the dataset: a sample of input and output relationships amongst Area Boards in England and Wales between 1969 and 1988. Section 7.4 of the paper presents the major results of the standard case, a production possibility set that associates 4 outputs with 3 inputs under variable returns to scale. Conclusions are presented in section 7.5.

7.2. Comparative performance and data envelopment analysis.

Comparative performance measurement amongst public utilities may be suggested for several reasons. Two of the most important concern, (a) incentive compatible regulatory mechanisms, and (b) the provision of a more complete information set to an asymmetrically informed regulator carrying out regulatory review of a public utility.

The first reason arises in the case where a set of public utilities is being regulated by some mechanism that relates the allowed prices for one utility to be at least partially based on the current performance of other similar utilities. The second reason emerges when a regulator is carrying out a review of the past actions and performance of a public utility. If the regulator is in a position of an asymmetrically informed principal in a principal-agent game with the utility as the agent with an information monopoly, then the regulator will seek to compare the utility's observed past performance with that of other similar utilities (see Rees [1985] for an extended discussion).

With this background, a variety of performance measures can be suggested. These range from measures of regulatory price effectiveness which investigate relative differences in output pricing procedures, to measures of relative productive efficiency. Amongst the latter the classic paper by Farrell (1957) identified:

1) allocative efficiency: the choice of cost minimising input mix;
2) technical efficiency: the relative distance of the utility from the best-practice frontier of the industry's production possibility set;
3) overall productive efficiency: the product of the first two factors.

There already exists a considerable literature developing the Farrell ideas. The emphasis in this chapter is to develop a pooled cross section - time series estimate of the production possibility set for electricity distribution in England and Wales over the period 1969/70-1988/89. This best-practice, observed production possibility set can be used to measure the relative technical efficiency of the different Area Electricity Boards over the period, allowing us to determine both trends in efficiency and the important characteristics of the relative performance of the different Area Boards.

7.3 Dataset and Area Electricity Boards 1969-88.

Determining the appropriate input and output set is important in any DEA study. As indicated earlier, the objective in this paper is to generate a pooled time series cross section sample from which to construct the production possibility set for electricity distribution. This permits each Area Boards' input-output mix in each year of the period 1969-70 to 1988-9, (financial years), to be treated as a separate point in the reference production possibility set. There are therefore 240 separate such points, each of which is treated as a separate decision making unit.

In an earlier study, Weyman-Jones (1991), Area Boards were assumed to deliver three outputs from the usage of two inputs. The inputs were manpower and capital services in physical units; respectively: numbers of employees, and the size of the distribution mains in circuit-kilometres. No separate energy input is used, since the Boards are modelled as using their capital stocks and manpower to add value to the flowthrough of electricity, (kilowatt hours [KWH]) from the generation end of the industry to final consumers. Electrical losses, which might be characterised as an input were assumed to be functionally related to the size of the distribution network, and therefore not to need separate measurement. Broadly the same approach is used in this paper, except that the input set is extended to include two types of capital service, as well as manpower. In addition to the size of the distribution network in circuit-kilometres, there is a supply of capital services from the total size of the transformer capacity, (megavoltamps [MVA]), available in each year.

On the output side the previous paper distinguished three outputs: electricity distribution services to three different types of market: domestic, (i.e. residential and small commercial premises), commercial and industrial. Assuming that distribution services are proportional to electricity distributed allows us to use KWH units supplied as the output variables.

It is important to distinguish these three markets separately, since although each market receives an apparently homogeneous product, (KWH of electricity), the characteristics of demand and the nature of the product in a wider sense are quite different. The differences include the timing of demand, and its peaks, the voltage at which load can be delivered, the reactive power facilities which can be provided by some types of customer, the geographical concentration of demand, and the technical aspects of the contracting and service arrangements which are implicit in each market category. It is commonplace now to treat electricity supply as a multiproduct industry, and the distinction between market categories is one manifestation of this. The same three outputs are used in this paper, with the addition of one more. This arises from an attempt to capture the fact that the timing of demand is especially critical in determining the organisation of the supply industry. Distributing a constant load of 1 kW every hour of the year is a completely different supply exercise from delivering 8760 kW for 1 hour on a single winters day. To capture this 'peakiness' aspect of demand, the additional variable is each Board's distributed load at system maximum demand in each year.

In summary, the production possibility set is constructed from 240 observed input-output bundles comprising, for each year from 1969-70 to 1988-9:

Outputs:

y1: maximum demand, (kW)
y2: domestic units supplied, (kWh)
y3: commercial units supplied, (kWh)

y4: industrial units supplied, (kWh)

Inputs:

x1: manpower employed, (number of employees)

x2: distribution mains, (circuit-kilometres)[2]

x3: transformer capacity, (MVA)

All of the data are from the published source of the UK Electricity Council (1989).

Running each DEA exercise therefore involved solving 240 linear programmes, (241 variables and 7 constraints), for the constant returns to scale case, and a further 240 programmes, (241 variables and 8 constraints) for the variable returns case. (In addition each programme contained lower bounds on the dual variables.)

The full results of running the programmes are given in tables 7.4.1 and 7.4.2. A broad summary of the relative numbers of efficient and inefficient DMUs is as follows

	constant returns	variable returns
number of efficient DMU s	35	45
number of inefficient DMU s	205	195
	---	---
totals	240	240

This clearly establishes that only 10 of the observed points below the production frontier under the assumption of constant returns to scale are measured as productively inefficient due to scale inefficiency alone. There are 195 out of 240 DMUs which are measured to show pure technical inefficiency.

We calculated the average technical efficiency (ATE) under constant returns to scale to be 0.902 and ATE under variable returns to scale as 0.923. The ratio of these two values, 0.977, gives us an average scale efficiency (ASE) for the whole database. As a result we can see that scale efficiency alone is, on average, not an important determinant of the average technical efficiency results.

This relative unimportance of scale inefficiency need not be surprising since the Area Boards were initially drawn up in a way that very roughly equalised their size and scale. It is also an indication that the previous study of the Boards did not greatly overestimate their inefficiency despite its assumption of constant returns to scale. We turn now to a detailed discussion of the results.

7.4 Results.

7.4.1 Analysis of efficiency change over time.

The DEA for twelve Area Boards over twenty years can reveal information about trends and the differences between Area Boards (see Table 7.4.1, the VRS program will be used for all analyses in this section). What is apparent through casual observation of the results, is that all of the Area Boards were efficient, or nearly so, at the beginning and end of the twenty years. There is also a period in the middle of this time when efficiency levels have risen. The sharp fall in technical efficiency in 1971/72 can be attributed to a large increase in the size of the mains circuit. In general though, changes in inputs do not seem to affect the results as significantly as changes in outputs. To show trends in efficiency, the overall technical efficiency for each year is determined by averaging across all area boards for the year and a mean figure is produced.

The main trends in technical efficiency seem closely linked to the state of the economy (see Figure 7.4.1 which shows ATE of all 12 boards for each year of the sample period). The stagnation of the mid-seventies following the oil shock of 1973/74 marks a trough in Area Electricity Board average technical efficiency. This is confirmed by the fact that in 1974/75 average technical efficiency peaks at 0.933 and in 1976/77 it is at a low point for the seventies of 0.882. After a recovery, average technical efficiency

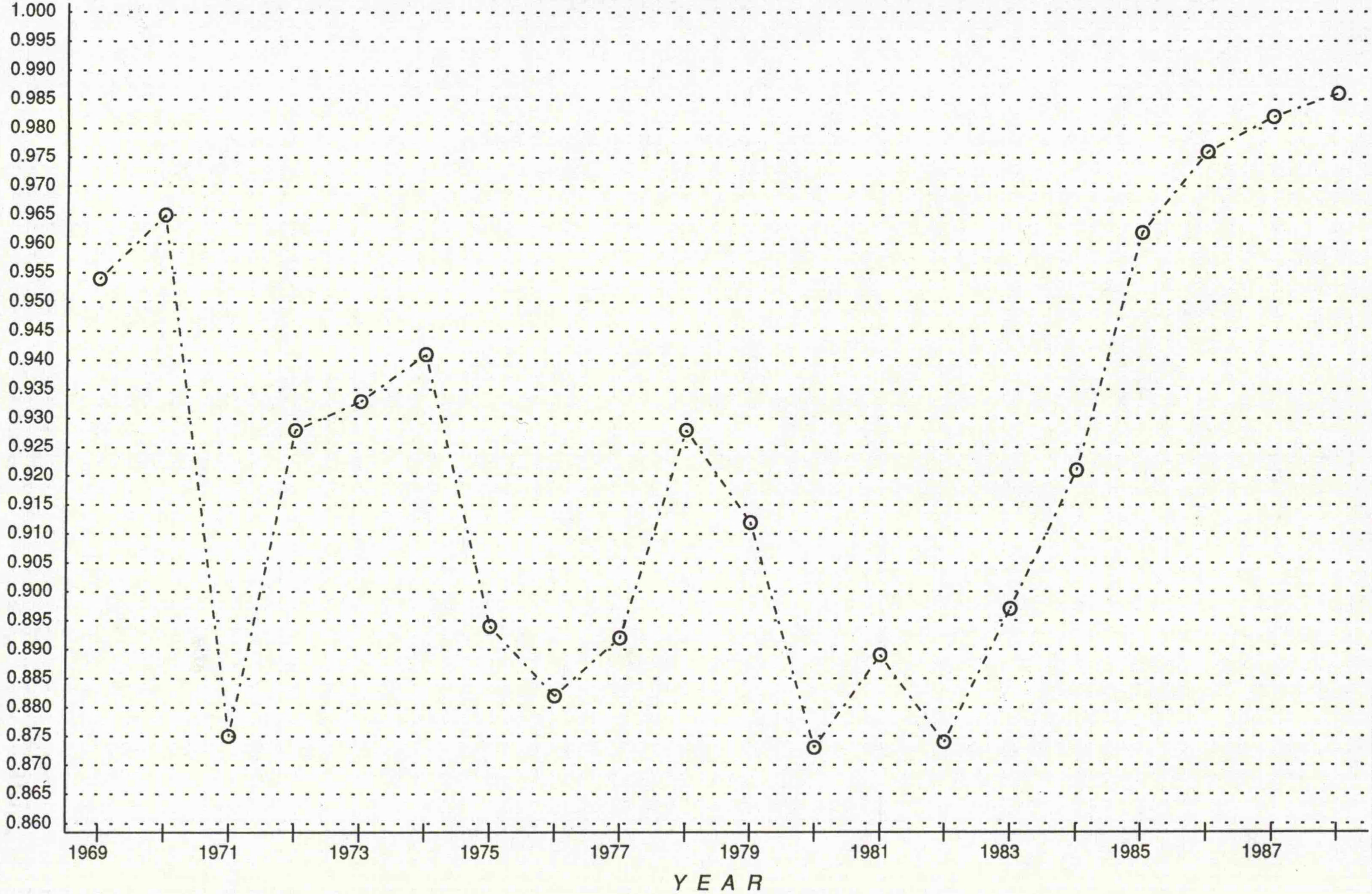
Table 7.4.1. Dynamic DEA analysis under variable returns to scale.

	69/70	70/71	71/72	72/73	73/74	74/75	75/76	76/77	77/78	78/79	79/80	80/81	81/82	82/83	83/84	84/85	85/86	86/87	87/88	88/89	AVER.
1.NORWEB	0.965	0.994	0.854	0.917	0.927	0.936	0.874	0.857	0.854	0.874	0.847	0.810	0.818	0.802	0.818	0.826	0.862	0.883	0.887	0.885	0.875
2.EASTERN	0.981	1	0.935	0.988	1	0.998	0.990	0.895	0.923	0.960	0.936	0.924	0.941	0.946	0.943	0.972	0.991	1	1	1	0.966
3.SEEBOARD	0.964	1	0.917	0.993	0.983	1	0.966	0.900	0.926	0.951	0.915	0.896	0.904	0.939	0.895	0.952	0.972	1	1	1	0.954
4.MEB	1	1	0.985	1	1	1	0.950	0.940	0.964	1	0.999	0.914	0.963	0.887	0.903	0.956	0.964	1	1	1	0.971
5.MANWEB	1	1	0.951	1	1	0.992	0.929	0.947	0.933	0.957	0.949	0.881	0.895	0.875	0.930	0.965	0.993	1	1	1	0.960
6.SOUTH WALES	1	1	0.860	0.905	0.882	0.882	0.802	0.875	0.870	0.928	0.920	0.839	0.845	0.846	0.913	0.891	0.953	0.971	1	1	0.909
7.SOUTHERN	0.932	0.956	0.843	0.890	0.925	0.927	0.892	0.86	0.873	0.925	0.909	0.894	0.921	0.885	0.930	0.945	0.981	1	1	1	0.924
8.LONDON	1	1	0.913	0.973	0.964	0.989	0.954	0.916	0.896	0.952	0.925	0.906	0.906	0.887	0.914	0.961	1	1	1	1	0.953
9.NORTH EASTERN	0.992	0.959	0.800	0.849	0.876	0.881	0.829	0.838	0.845	0.880	0.884	0.847	0.865	0.866	0.907	0.935	1	0.959	0.989	0.980	0.899
10.YORKSHIRE	1	1	0.916	0.976	0.987	1	0.945	0.970	0.972	1	0.977	0.920	0.932	0.914	0.922	0.910	0.974	0.974	0.991	1	0.964
11.SOUTH-WEST	0.780	0.808	0.734	0.803	0.799	0.816	0.774	0.766	0.806	0.829	0.819	0.807	0.835	0.816	0.848	0.876	0.941	0.980	0.944	0.987	0.838
12.EAST MIDLANDS	0.836	0.868	0.786	0.840	0.856	0.866	0.825	0.817	0.839	0.875	0.861	0.840	0.837	0.826	0.839	0.868	0.909	0.950	0.968	0.980	0.864
AVERAGE	0.954	0.965	0.875	0.928	0.933	0.941	0.894	0.882	0.892	0.928	0.912	0.873	0.889	0.874	0.897	0.921	0.962	0.976	0.982	0.986	

Table 7.4.2. Dynamic DEA analysis under constant returns to scale.

	69/70	70/71	71/72	72/73	73/74	74/75	75/76	76/77	77/78	78/79	79/80	80/81	81/82	82/83	83/84	84/85	85/86	86/87	87/88	88/89	AVER.
1.NORWEB	0.937	0.972	0.814	0.876	0.884	0.894	0.837	0.824	0.820	0.840	0.813	0.777	0.786	0.773	0.789	0.800	0.834	0.854	0.858	0.859	0.842
2.EASTERN	0.954	1	0.812	0.864	0.874	0.868	0.874	0.821	0.864	0.917	0.925	0.918	0.939	0.946	0.941	0.971	0.991	1	1	1	0.924
3.SEEBOARD	0.962	1	0.849	0.930	0.934	0.946	0.916	0.867	0.897	0.922	0.891	0.875	0.888	0.934	0.887	0.951	0.971	1	1	1	0.931
4.MEB	1	1	0.928	1	0.993	1	0.921	0.905	0.930	0.974	0.935	0.876	0.884	0.871	0.889	0.927	0.957	1	1	1	0.950
5.MANWEB	1	1	0.947	1	1	0.992	0.929	0.947	0.933	0.957	0.949	0.881	0.893	0.873	0.927	0.959	0.989	1	1	1	0.959
6.SOUTH WALES	0.981	1	0.858	0.904	0.873	0.877	0.802	0.874	0.869	0.924	0.915	0.838	0.830	0.843	0.912	0.891	0.946	0.964	0.991	1	0.905
7.SOUTHERN	0.883	0.931	0.757	0.793	0.843	0.846	0.819	0.793	0.804	0.853	0.842	0.831	0.876	0.853	0.910	0.934	0.963	1	1	1	0.877
8.LONDON	1	1	0.846	0.915	0.916	0.944	0.921	0.863	0.858	0.895	0.886	0.870	0.880	0.874	0.906	0.954	1	1	0.998	1	0.926
9.NORTH EASTERN	0.914	0.913	0.799	0.849	0.876	0.881	0.829	0.838	0.845	0.880	0.884	0.845	0.860	0.856	0.894	0.920	0.954	0.944	0.973	0.968	0.886
10.YORKSHIRE	1	1	0.899	0.961	0.982	0.996	0.936	0.944	0.948	0.985	0.953	0.909	0.917	0.895	0.913	0.894	0.970	0.970	0.990	1	0.953
11.SOUTH-WEST	0.734	0.779	0.732	0.799	0.786	0.811	0.773	0.765	0.805	0.828	0.818	0.805	0.815	0.796	0.823	0.844	0.898	0.925	0.905	0.913	0.818
12.EAST MIDLANDS	0.826	0.864	0.773	0.823	0.836	0.842	0.802	0.804	0.834	0.869	0.851	0.836	0.832	0.819	0.833	0.858	0.907	0.950	0.968	0.980	0.855
AVERAGE	0.933	0.955	0.835	0.893	0.900	0.908	0.863	0.854	0.867	0.904	0.889	0.855	0.867	0.861	0.885	0.909	0.948	0.967	0.974	0.977	

Figure 7.4.1. Area Electricity Boards average technical efficiency 1969-88.



peaks again in 1978/79 at 0.928. Another oil shock and the start of a recession coincides with average technical efficiency declining for the next two years. It then rises year on year, until it reaches its highest efficiency level ever in 1988/89 at 0.994, a year in which the economic boom of the late 1980s is at its peak. These trends could indicate that the technical efficiency of the electricity industry is closely related to the amount of electricity supplied, which in turn depends on the level of demand for electricity, which depends on the level of activity in the economy. An increase in economic activity increases technical efficiency and vice-versa. To confirm that this linkage actually exists, the direction of causation will be followed and explained.

To test the basic hypothesis we can investigate the relationship between the productive efficiency for a Board and the stage of the economic activity cycle in its corresponding economic region.

To do so, we must first note that we cannot safely assume that the measured efficiency indices are the result of random sampling from a known probability density function, e.g. the Normal distribution. Consequently, as in most studies of DEA efficiency, we need to adopt distribution-free statistical testing procedures. The well known cost of this choice is that such methods make inefficient use of sample information.

To measure the stage of the economic cycle in a region, we use the ratio of measured regional GDP at factor cost to its linear trend: (Y/Y^*). The distribution-free correlation measure used is Spearman's rank correlation coefficient, r_s, whose sampling distribution is known, Siegel (1956).

Table 7.4.3 gives the value of the sample rank correlation coefficient, $r_s(\theta, Y/Y^*)$, between each Board's measured efficiency index in a given year and the ratio of regional GDP to its long term trend. For sample sizes

Table 7.4.3

Correlation between measured efficiency and regional economic activity: 1970/71 to 1986/87.

Area Board	Regional GDP series for Y/Y*	rank correlation coefficient	t value
Norweb	Northwest	0.334	1.372
Eastern	East Anglia	0.546*	2.524
Seeboard	Rest of South East	0.513*	2.315
Midlands	West Midlands	0.517*	2.339
Manweb	Wales	0.541*	2.491
South Wales	Wales	0.640*	3.226
Southern	South West	0.562*	2.632
London	Greater London	0.723*	4.053
North Eastern	North	0.527*	2.402
Yorkshire	Yorkshire	0.255	1.021
South-West	South West	0.544*	2.511
East Midlands	East Midlands	0.555*	2.584

*: significant at 5% level;

critical values: $t_{0.025} = 2.131$, $t_{0.005} = 2.947$

exceeding 10, Siegel suggests testing the null hypothesis: H_0: $\rho^s = 0$, against the alternative H_1: $\rho_s \neq 0$, using the t distribution. Under the null, we have the result that:

$$r_s\sqrt{[(n-2)/(1-r_s^2)]} \sim t_{n-2}$$

The table shows the regional GDP series used to generate each Board's (Y/Y*) series, the sample r_s and the corresponding t value for the null hypothesis. Recall that the measured efficiency index used here relates each Board's efficiency to its own and all other Boards' performances over the sample period, (1970-71 to 1986-7 for this test).

The context of the results is this. A positive correlation would indicate that a given Board moves onto or away from the industry production frontier according to the relative strength of economic activity in its own local region. Such a positive correlation, if found, would be a critical constraint on the operation of yardstick performance measurement. For yardstick comparisons to be meaningful in these circumstances, the Boards being compared must be at synchronous positions in their own regional GDP trends, and regulatory reviews might need to be synchronous to take account of the role of regional GDP trends in determining different Boards' positions relative to the industry production frontier. Any suggestion of a positive correlation introduces these constraints on yardstick comparisons, but the degree of correlation found is quite surprisingly high.

All the sample correlation coefficients are positive, and for 10 of the 12 Area Boards the correlation is significant at the 5% level. Given the very approximate nature of the test, the incomplete matching of regions and Boards, and the loss of information in using a distribution-free test statistic, these are unexpectedly strong findings. They have major implications for the nature and timing of regulatory review procedures.

To confirm that the state of the economy affects the supply of electricity, it has to be examined what actually happened to the demand for electricity during the slump after 1973/74, and the recession after 1979/80. Looking at the raw data for electricity supplied to the domestic, commercial and industrial sectors reveals a rise in demand for electricity in total. There are fluctuations around peaks and troughs in economic activity. Though to reveal what is actually happening requires examination of the data for each sector that electricity is supplied to.

Commercial supply has risen continuously over the last twenty years, and it is the domestic and industrial sectors that respond most to economic events. There is a peak in electricity supply to both these sectors in 1974/5. This is followed by a fall in domestic supply for one year and industrial supply for two years before they start to rise again. The decline in electricity supply during this time coincides with a dip in economic activity. After another peak in electricity supply in 1978/79 there is a fall in the supply of electricity to both sectors for the following four years coinciding with the recession. The boom of the late 1980s occurred at a time when electricity demand is at its highest point ever.

The key question then is, why does the supply of electricity change the level of technical efficiency in the electricity industry? The answer is that technical efficiency is related to the amount of spare capacity, (or unused resources) that the industry has. There is always some spare capacity to allow for peaks in demand. But, if electricity supply has to be increased at a greater rate than previously due to a boom and a rise in demand then spare capacity is being utilised more fully, and so used more efficiently. During a slump there is unused capacity and technical efficiency falls.

In the light of this, several aspects of the results can be explained. Average technical efficiency is lower during the recession in the early 1980s than during the decline in economic activity in the mid-seventies. This is because the recession in the early eighties was that much more severe and produced a greater decline in the demand for electricity. In most cases even less electricity was supplied than in the previous trough. Area board resources were thus utilised less efficiently than they had been even previously. Another interesting feature is the increasing levels of technical efficiency in the late eighties to their highest levels ever. This can be explained by efficient utilisation of resources due to large increases in electricity demand during the economic recovery. A contributory factor was the continuing reduction in labour inputs, which took place at a much greater rate than previously in response to the more malleable labour relations climate generated by the recession.

The trends described above need validating in some way. As efficiency is measured relative to other Area Boards over time it could be that the trend in an individual Area Board is obscured by what is happening in the others. To find out what was happening to the Area Boards individually a DEA was conducted on each one over the twenty year period (see Table 7.4.4). Looking across the table, for each region the results are telling us how technical efficiency has varied over the twenty years for an Area Board relative to itself. The true trend for each Area Board is revealed. The results are not meaningful in determining which was the most efficient as they show only how much fluctuation there has been in technical efficiency and a high number of technically efficient DMUs only show that there was little change. Table 7.4.4 shows that the trends shown in Table 7.4.1 were correct. The change is slightly obscured by the large number of years in which technical efficiency was unity, but the movements are exactly the same.

7.4.2 Analysis of efficiency across each area.

Examining the results on a regional basis is useful and can reveal information about the industry as a whole. Eight

Table 7.4.4. Dynamic DEA analysis for each individual AEB under variable returns to scale.

	69/70	70/71	71/72	72/73	73/74	74/75	75/76	76/77	77/78	78/79	79/80	80/81	81/82	82/83	83/84	84/85	85/86	86/87	87/88	88/89
1.NORWEB	1	1	1	1	1	1	0.955	0.986	0.982	1	1	0.942	0.949	0.923	0.928	0.955	1	1	1	1
2.EASTERN	1	1	0.981	1	1	0.998	1	0.898	0.930	0.986	1	0.961	0.992	1	0.974	1	1	1	1	1
3.SEEBOARD	1	1	1	1	0.996	1	0.990	0.949	0.946	0.987	0.983	0.931	0.919	1	0.908	1	1	1	1	1
4.MEB	1	1	1	1	1	1	0.953	0.954	0.972	1	1	0.927	0.976	0.912	0.916	0.966	0.979	1	1	1
5.MANWEB	1	1	1	1	1	1	0.951	0.958	0.952	0.988	0.968	0.931	0.951	0.915	0.964	1	1	1	1	1
6.SOUTH WALES	1	1	0.863	0.927	0.969	0.980	0.925	0.931	0.931	0.980	0.964	0.927	0.946	0.929	0.919	0.946	0.979	1	1	1
7.SOUTHERN	1	1	0.983	1	0.998	1	0.960	0.936	0.958	1	1	0.975	1	0.953	1	0.983	0.985	1	1	1
8.LONDON	1	1	0.980	1	0.988	1	1	0.963	0.981	1	1	0.951	0.945	0.913	0.923	0.968	1	1	1	1
9.NORTH EASTERN	1	1	0.967	1	1	1	0.962	0.951	0.948	0.989	0.990	0.948	0.925	0.923	0.959	1	1	1	1	1
10.YORKSHIRE	1	1	1	1	1	1	0.949	0.970	0.972	1	0.979	0.928	0.932	0.919	0.927	0.936	1	1	1	1
11.SOUTH WEST	1	1	1	1	0.988	1	0.940	0.918	0.951	0.977	0.970	0.994	0.927	0.905	0.971	1	1	1	1	1
12.EAST MIDLANDS	1	1	1	1	1	0.990	0.940	0.936	1	1	0.992	0.941	0.942	0.933	0.940	0.968	0.974	1	1	1

regions were deemed to be efficient in 1988/89 and of those remaining, three were only marginally efficient. This result is unrevealing because the number of inputs and outputs is large relative to the number of Boards in a single year sample. Much more interesting information is gained by averaging the technical efficiency numbers over the twenty year period for each area. It is clear from this that MEB is most efficient with average technical efficiency of 0.971 and the South-West area board the least efficient with 0.838.

What is apparent from the results is that the area boards which have been most efficient over the last twenty years also supply predominantly to one sector. Table 7.4.5 shows the proportion of each area board's supply that has gone to the three sectors over the last twenty years). It will be confirmed later when examining the virtual outputs that sectors which receive the largest part of an Area Board's supply contribute most to the efficiency rating. It could thus be presumed that specialisation confers greater efficiency. MEB, the most efficient Area Board, supplied similar amounts to two sectors so it can only be presumed that its overall success was gained by being efficient in supply to more than one sector. This is not to say that specialisation in supply automatically confers greater technical efficiency, there may be factors which decrease technical efficiency on the input side. For instance, overmanning or a large mains circuit due to a high dispersion of customers.

Given that specialisation in supply is a factor in determining technical efficiency ratings, it might be in order to conduct a cluster analysis. That is, dividing the Area Boards into groups according to which sector they supply predominantly to and running a DEA on each group. (London Area Board would have to be put into the domestic supply group because it supplies most to the commercial sector and is alone in doing so). Such an analysis would be revealing because it would be comparing more homogenous Area Boards.

Table 7.4.5.

Proportion of electricity supplied to each sector.

	D	C	I
1.NORWEB	0.39	0.19	0.42
2.EASTERN	0.49	0.23	0.28
3.SEEBOARD	0.53	0.23	0.24
4.MEB	0.39	0.18	0.44
5.MANWEB	0.31	0.14	0.55
6.SOUTH WALES	0.27	0.13	0.60
7.SOUTHERN	0.48	0.23	0.29
8.LONDON	0.40	0.45	0.15
9.NORTH EASTERN	0.32	0.17	0.51
10.YORKSHIRE	0.32	0.13	0.55
11.SOUTH WEST	0.57	0.28	0.16
12.EAST MIDLANDS	0.37	0.17	0.46

7.4.3 Window analysis.

Window analysis is a related approach to applying DEA dynamically in that it also analyses technical efficiency over time. But, instead of examining DMUs in all time periods as in the approach already outlined, a moving window examines subsets of the data over time. Charnes et al. (1985) introduced the technique in an application to US Army recruiting districts. Bowlin (1987) also uses window analysis to measure and evaluate the operational efficiency of US Air Force organisations. Bowlin identifies three reasons for using window analysis: to increase the number of DMUs and hence the number of degrees of freedom, to obtain information trends and to assess stability and validity of ratings obtained through DEA cross-sections. Each of these reasons will be evaluated in the context of Bowlin's study, their applicability to an Area Electricity Board study and how they compare with the dynamic DEA approach.

Bowlin uses window analysis to increase the number of DMUs, he does this by disaggregating annual data into quarters. The reason why there are insufficient numbers of DMUs for the variables being used is that under the circumstances there is not an adequate number of degrees of freedom. This means that all DMUs might be rated 100% efficient. A DEA cross-section analysis of the air base in Bowlin's study using annual data produced very little discrimination between the bases, five of them were rated efficient and the other two were only marginally inefficient. A heuristic minimum number of DMUs per Input/Output, suggested by Charnes et al. after empirical testing is two. On this basis Bowlin's annual data is insufficient because there are not enough DMUs. Disaggregation to five quarters increased the number of DMUs and allowed greater discrimination to take place between efficient and inefficient DMUs and to evaluate more clearly the degree of inefficiency.

Looking at the Area Electricity Board data in the light of this information it is clear that conducting a cross-

Table 7.4.6.

DEA cross-section of area electricity boards for 1988/89.

1. NORWEB	0.902
2. EASTERN	1
3. SEEBOARD	1
4. MEB	1
5. MANWEB	1
6. SOUTH WALES	1
7. SOUTHERN	1
8. LONDON	1
9. NORTH EASTERN	1
10. YORKSHIRE	1
11. SOUTH WEST	1
12. EAST MIDLANDS	0.981

section on the data for a single year would just not quite produce a credible result. There are twelve DMUs and seven inputs/outputs, so there should be at least two more DMUs or one less input/output. The result in Table 7.4.6 for 1988/89 proves this, there is almost no discrimination. In Weyman-Jones (1991) original study of Area Electricity Boards on a cross-section of 1987/88 data there was clearly a similar problem. Even though there were twelve DMUs for five inputs/outputs, giving enough degrees of freedom, the level of discrimination was not high. In this situation disaggregating the annual data would yield more information, but this is not possible. Fortunately, our aim is to evaluate the twelve DMUs over a twenty year period anyway, so there are enough degrees of freedom and this is not a problem. Disaggregation to increase the number of degrees of freedom when there is not enough information to produce a useful result is a good idea in itself, and not specific to the use of window analysis. Applying DEA to all disaggregated data instead of windows of the data is just as feasible.

To reveal trends over time, Bowlin applied DEA to a three quarters window of the five quarters, that is, three fifths of the data. In doing this it is not clear why he did not also do a dynamic DEA run over the five quarters. Trends would have been revealed much more clearly because a single DEA number would exist in each time period for an organisation. Of course the three fifths coverage window analysis would not relate directly to this 'open' window analysis because in general the more DMUs there are the lower are the DEA ratings. Thus, to reveal trends a window analysis is not necessary, a dynamic DEA run including all time periods is much simpler and more revealing. This can be confirmed by comparing the DEA results for area electricity boards in Table 7.4.1 to the window analysis conducted on the same data (see Table 7.4.7).

Another reason for conducting a window analysis according to Bowlin is to assess the stability and validity of DEA ratings prior to disaggregation. Though, as the

Table 7.4.7. Window analysis of area electricity boards from 1969 to 1988.

	69/70	70/71	71/72	72/73	73/74	74/75	75/76	76/77	77/78	78/79	79/80	80/81	81/82	82/83	83/84	84/85	85/86	86/87	87/88	88/89
1.NORWEB																				
An 1	0.965	0.994	0.874	0.933	1	0.959	0.899	0.887	0.886	0.914	0.899	0.858	0.862	0.852	0.876					
An 2		0.994	0.874	0.931	1	0.956	0.895	0.88	0.879	0.906	0.888	0.846	0.850	0.839	0.862	0.868				
An 3			0.919	0.977	0.981	0.987	0.917	0.901	0.898	0.918	0.894	0.856	0.858	0.837	0.846	0.855	0.893			
An 4				0.977	0.981	0.987	0.915	0.889	0.885	0.908	0.881	0.841	0.840	0.818	0.834	0.839	0.878	0.897		
An 5					0.981	0.987	0.915	0.889	0.885	0.908	0.880	0.841	0.840	0.818	0.832	0.837	0.873	0.893	0.898	
An 6						0.984	0.912	0.889	0.885	0.906	0.880	0.841	0.840	0.818	0.831	0.836	0.872	0.891	0.894	0.888
2.EASTERN																				
An 1	0.982	1	0.958	0.993	1	1	1	0.932	0.962	1	1	0.987	1	1	1					
An 2		1	0.958	1	1	1	1	0.932	0.962	1	0.998	0.974	0.983	0.993	0.980	1				
An 3			0.955	0.993	1	1	0.994	0.925	0.949	1	0.989	0.951	0.963	0.968	0.956	1	1			
An 4				0.991	1	1	0.990	0.905	0.932	0.967	0.939	0.925	0.947	0.958	0.952	0.976	1	1		
An 5					1	1	0.990	0.904	0.932	0.967	0.939	0.925	0.941	0.946	0.943	0.973	0.991	1	1	
An 6						1	0.992	0.902	0.929	0.965	0.938	0.924	0.941	0.946	0.943	0.972	0.991	1	1	1
3.SEEBOARD																				
An 1	0.988	1	0.961	1	0.996	1	0.993	0.952	0.986	1	0.973	0.963	0.995	1	1					
An 2		1	0.959	1	0.994	1	0.984	0.934	0.962	0.987	0.954	1	0.950	1	0.950	1				
An 3			0.917	0.993	0.985	1	0.971	0.913	0.943	0.969	0.932	0.917	0.931	0.964	0.926	1	1			
An 4				0.993	0.983	1	0.967	0.900	0.927	0.951	0.916	0.897	0.905	0.939	0.895	0.952	0.976	1		
An 5					0.983	1	0.967	0.900	0.927	0.951	0.916	0.897	0.904	0.939	0.895	0.952	0.975	1	1	
An 6						1	0.967	0.900	0.927	0.951	0.917	0.897	0.904	0.939	0.895	0.952	0.975	1	1	1
4.MEB																				
An 1	1	1	1	1	1	1	0.953	0.941	0.965	1	1	0.939	0.996	0.947	0.983					
An 2		1	1	1	1	1	0.953	0.941	0.965	1	1	0.935	0.98	0.935	0.967	1				
An 3			0.986	1	1	1	0.951	0.941	0.965	1	1	0.924	0.973	0.908	0.934	0.993	1			
An 4				1	1	1	0.950	0.941	0.965	1	0.999	0.922	0.963	0.898	0.917	0.956	0.974	1		
An 5					1	1	0.954	0.940	0.965	1	0.999	0.917	0.963	0.888	0.904	0.956	0.965	1	1	
An 6						1	0.954	0.940	0.965	1	0.999	0.914	0.963	0.887	0.903	0.956	0.964	1	1	1

Table 7.4.7 continued.

5.MANWEB

	69/70	70/71	71/72	72/73	73/74	74/75	75/76	76/77	77/78	78/79	79/80	80/81	81/82	82/83	83/84	84/85	85/86	86/87	87/88	88/89
An 1	1	1	0.979	1	1	1	0.966	0.992	0.977	1	1	0.936	0.988	0.951	1					
An 2		1	0.981	1	1	1	0.961	0.992	0.974	1	1	0.925	0.951	0.927	0.987	1				
An 3			0.968	1	1	1	0.949	0.986	0.963	0.992	1	0.902	0.919	0.898	0.958	1	1			
An 4				1	1	0.999	0.943	0.983	0.962	0.985	1	0.901	0.910	0.892	0.943	0.976	1	1		
An 5					1	0.997	0.937	0.969	0.949	0.974	0.983	0.892	0.897	0.880	0.934	0.968	0.999	1	1	
An 6						1	0.934	0.961	0.943	0.967	0.970	0.885	0.895	0.875	0.930	0.965	0.993	1	1	1

6.SOUTH WALES

	69/70	70/71	71/72	72/73	73/74	74/75	75/76	76/77	77/78	78/79	79/80	80/81	81/82	82/83	83/84	84/85	85/86	86/87	87/88	88/89
An 1	1	1	0.932	0.980	0.972	0.966	0.898	0.945	0.941	1	0.995	0.927	1	1	1					
An 2		1	0.938	0.981	0.974	0.969	0.892	0.945	0.940	1	0.996	0.916	0.965	0.970	1	1				
An 3			1	1	0.975	0.979	0.889	0.975	0.964	1	0.999	0.934	0.946	0.939	1	0.981	1			
An 4				1	0.975	0.979	0.886	0.970	0.957	1	0.998	0.930	0.938	0.931	1	0.947	1	1		
An 5					1	0.999	0.916	1	0.977	1	0.996	0.926	0.933	0.924	1	0.944	0.975	0.985	1	
An 6						1	0.943	1	0.975	1	0.992	0.919	0.919	0.909	1	0.928	0.967	0.971	1	1

7.SOUTHERN

	69/70	70/71	71/72	72/73	73/74	74/75	75/76	76/77	77/78	78/79	79/80	80/81	81/82	82/83	83/84	84/85	85/86	86/87	87/88	88/89
An 1	0.934	0.956	0.859	0.907	0.941	0.947	0.917	0.891	0.909	0.967	0.954	0.944	0.988	0.951	1					
An 2		0.956	0.859	0.907	0.941	0.947	0.917	0.889	0.908	0.962	0.949	0.937	0.971	0.936	0.986	1				
An 3			0.914	0.949	0.995	0.997	0.956	0.921	0.943	0.986	0.962	0.940	0.959	0.916	0.958	1	1			
An 4				0.949	0.995	0.997	0.955	0.918	0.940	0.980	0.951	0.930	0.939	0.897	0.938	0.947	0.983	1		
An 5					0.995	0.997	0.955	0.918	0.940	0.978	0.949	0.928	0.938	0.896	0.935	0.947	0.983	1	1	
An 6						0.992	0.950	0.911	0.933	0.971	0.943	0.922	0.933	0.890	0.931	0.947	0.982	1	1	1

8.LONDON

	69/70	70/71	71/72	72/73	73/74	74/75	75/76	76/77	77/78	78/79	79/80	80/81	81/82	82/83	83/84	84/85	85/86	86/87	87/88	88/89
An 1	1	1	0.968	1	1	1	1	0.962	0.949	1	0.992	0.979	0.998	0.986	1					
An 2		1	0.964	1	1	1	1	0.949	0.945	1	0.974	0.961	0.964	0.946	0.977	1				
An 3			1	1	0.988	1	1	1	0.947	1	1	0.969	0.973	1	1	1	1			
An 4				1	0.988	1	1	1	0.933	0.981	1	0.959	1	1	1	1	1	1		
An 5					1	1	1	1	0.934	0.981	1	0.964	0.962	1	1	1	1	1	1	
An 6						1	1	1	0.934	0.981	1	0.964	0.962	1	1	1	1	1	1	1

Table 7.4.7 continued.

	69/70	70/71	71/72	72/73	73/74	74/75	75/76	76/77	77/78	78/79	79/80	80/81	81/82	82/83	83/84	84/85	85/86	86/87	87/88	88/89
9.NORTH EASTERN																				
An 1	1	0.965	0.868	0.911	0.918	0.908	0.869	0.874	0.883	0.934	0.950	0.914	0.933	0.947	1					
An 2		0.970	0.866	0.907	0.916	0.905	0.854	0.867	0.878	0.921	0.945	0.892	0.913	0.924	0.980	1				
An 3			0.877	0.923	0.944	0.956	0.893	0.912	0.918	0.960	0.984	0.908	0.911	0.899	0.942	0.967	1			
An 4				0.923	0.944	0.956	0.893	0.912	0.918	0.960	0.984	0.907	0.906	0.894	0.934	0.965	1	0.994		
An 5					0.941	0.953	0.890	0.909	0.915	0.957	0.981	0.902	0.899	0.886	0.923	0.955	0.990	0.975	1	
An 6						0.944	0.882	0.899	0.905	0.947	0.970	0.893	0.888	0.874	0.911	0.941	0.975	0.959	0.991	0.989
10.YORKSHIRE																				
An 1	1	1	0.919	0.978	1	1	0.945	0.971	0.977	1	1	0.958	0.988	0.990	1					
An 2		1	0.918	0.978	1	1	0.945	0.971	0.977	1	0.999	0.956	0.986	0.984	1	0.984				
An 3			0.927	0.978	0.988	1	0.946	0.978	0.976	1	0.988	0.939	0.961	0.947	0.972	0.952	1			
An 4				0.978	0.987	1	0.946	0.978	0.976	1	0.986	0.931	0.958	0.935	0.956	0.933	1	1		
An 5					1	0.987	1	0.946	0.978	0.975	1	0.979	0.923	0.938	0.921	0.930	0.918	0.983	0.983	
An 6						1	0.946	0.978	0.975	1	0.977	0.920	0.932	0.914	0.922	0.910	0.974	0.974	0.991	1
11.SOUTH WEST																				
An 1	1	1	0.972	1	0.979	0.997	0.934	0.909	0.949	0.978	0.967	0.946	0.954	0.922	0.958					
An 2		1	0.971	1	0.979	0.996	0.932	0.906	0.945	0.974	0.963	0.941	0.949	0.915	0.948	0.975				
An 3			0.734	0.803	0.799	0.818	0.781	0.776	0.821	0.844	0.836	0.828	0.864	0.847	0.882	0.914	0.983			
An 4				0.803	0.799	0.816	0.775	0.768	0.809	0.832	0.822	0.810	0.840	0.821	0.855	0.883	0.949	0.989		
An 5					0.800	0.816	0.775	0.767	0.808	0.831	0.821	0.809	0.837	0.818	0.851	0.879	0.944	0.984	0.948	
An 6						0.816	0.774	0.766	0.806	0.829	0.819	0.807	0.835	0.816	0.848	0.876	0.941	0.980	0.944	0.987
12.EAST MIDLANDS																				
An 1	0.864	0.896	0.803	0.865	0.895	0.910	0.875	0.879	0.921	0.953	0.940	0.918	0.920	0.910	0.925					
An 2		0.882	0.798	0.861	0.888	0.903	0.867	0.867	0.892	0.936	0.915	0.895	0.889	0.879	0.895	0.922				
An 3			0.794	0.853	0.875	0.888	0.850	0.852	0.890	0.922	0.910	0.884	0.885	0.877	0.886	0.920	0.944			
An 4				0.845	0.863	0.875	0.837	0.834	0.856	0.900	0.882	0.860	0.854	0.842	0.860	0.884	0.930	0.969		
An 5					0.859	0.869	0.829	0.821	0.845	0.882	0.866	0.846	0.842	0.833	0.845	0.875	0.915	0.956	0.976	
An 6						0.866	0.825	0.817	0.839	0.875	0.861	0.840	0.837	0.826	0.839	0.868	0.909	0.950	0.968	0.980

original results were not deemed to be meaningful due to an insufficient number of degrees of freedom and a lack of discrimination, only analyses producing extremely inefficient results can be validated. Using window analysis to validate cross-sections is problematic anyway because the results can only be viewed in very general terms. The values for each analysis for a particular time period can vary quite widely. It can be seen a priori that as a window is only a partial look at a sequence of events any trend change will distort the ratings. This can be seen from Table 7.4.6 where the ratings in a specific time period decline (if they change at all) as they are read down vertically. The level of variation in DEA ratings, in a single run, is determined by the efficiency difference between the most efficient and the least efficient DMUs. There must be increasing variation in efficiency as each successive fifteen year window moves across. The results need to be clarified because in any single year the results of the different runs can vary as much as ten per cent. But, a mean cannot be taken of the different analyses in a specific time period because the number produced is not that useful. The average is across different DEA runs with different reference groups. A mean number is determined according to the average of a DEA number formed relative to these reference groups and so has no meaning.

Thus, conducting a dynamic DEA including DMUs in all time periods is superior to window analysis. Disaggregating data to increase the number of degrees of freedom is not exclusive to the application of window analysis. Validating the ratings produced prior to disaggregation is not really the point as there may be very low discrimination and can only really confirm some of the more inefficient results. The point then, is that window analysis can show trend changes in efficiency. But, it does not do so simply or clearly because of variations in the data and the meaninglessness of taking an average across the different analyses in a single time period. Doing a complete DEA on all time periods would produce much more clarity.

7.4.4 Virtual inputs and outputs.

A way of revealing information about the reasons for certain trends, is the examination of the virtual inputs and outputs of an efficient DMU. For an efficient unit, the virtual input/output attributable to a given input/output is the product of the input/output and its corresponding weight. Examining the optimal set of weights for a unit would reveal which of its inputs and outputs contribute to its efficiency rating but not their relative importance. A greater weight of an input or output does not necessarily mean that the unit utilises the corresponding input or produces the corresponding output more efficiently than inputs or outputs with lower weights. The size of each weight is dependent upon the scale of measurement of the relevant input/output. The virtual input/output, i.e. weight times variable, does show the proportion of the efficiency rating which is contributed by each input and output. Virtual inputs/outputs are therefore useful in determining how efficient units derive their efficiency. This is helpful information because when used in a cross- section it can be used to identify aspects of performance of relatively efficient units which are worthy of further investigation to identify good operating practices.

Examining the virtual inputs/outputs of the efficient Area Electricity Boards in different time periods can reveal why they were efficient, and to an extent, the reasons why they became efficient. It is revealed in Table 7.4.8, that of the nine area boards which were efficient at some point in time, seven were efficient near the beginning or end, or both ends of the twenty year period. The Southern and North-East areas were only efficient towards the end of the time period.

In this paper we concentrate attention on the virtual outputs since these are where the Boards have greater freedom to optimise. In input usage, a Board's choice of distribution mains length or transformer capacity may be regarded as approximately exogenous. In output choice, by contrast, the Boards can direct marketing and sales effort to different market sectors with differential load characteristics. Tariff policy may be a primary driving force in this context.

7.4.5 Virtual outputs.

The key piece of information to be gleaned from the virtual outputs is the confirmation of the earlier finding that area boards derive their efficiency from those sectors which they supply most to (they may not necessarily derive all their efficiency in these dimensions because it may also be derived from maximum demand capability). Comparing the virtual outputs in Table 7.4.8 to the proportion of output supplied in Table 7.4.5 quite categorically shows this. The most striking example is South Wales area board which supplied sixty per cent of its electricity overall to the industrial sector and in those years in which it was efficient, derived all of its efficiency rating from the industrial sector. This is also true of the Southern area which supplied fifty three per cent of its electricity overall to the domestic sector and derived its efficiency in that dimension alone. Supply to the commercial sector provides very little contribution to the efficiency rating derivation though its small contribution seems to have been made over the last few years. The exception being the London area which distributes a large proportion of its supply to the commercial sector. The association between the proportion of electricity supplied and efficiency rating derivation can be seen in varying degrees across all of the area boards. Thus it can be concluded that specialisation in supply to a sector means that an area board will derive its efficiency from that sector.

The level of capability of meeting maximum demand features a lot or not at all in contributing to the derivation of the efficiency rating. It would be expected that those who supply a high proportion of their supply to the domestic sector would also be efficient in achieving a

Table 7.4.8.
Virtual outputs for efficient area electricity boards.

		VIRTUAL OUTPUTS			
		MD	D	C	I
2.EASTERN	70/71	0.871	0.130	0.000	0.014
	73/74	0.000	0.997	0.000	0.000
	86/87	0.660	0.000	0.357	0.000
	87/88	0.000	0.989	0.007	0.000
	88/89	0.482	0.000	0.520	0.000
3.SEEBOARD	70/71	0.000	1.021	0.000	0.000
	74/75	0.000	0.981	0.000	0.000
	86/87	0.000	0.917	0.050	0.000
	87/88	0.000	0.984	0.006	0.000
	88/89	0.000	0.905	0.116	0.000
4.MEB	69/70	0.986	0.000	0.000	0.032
	70/71	1.018	0.000	0.000	0.000
	72/73	0.197	0.480	0.000	0.324
	73/74	0.000	0.986	0.000	0.000
	74/75	0.000	0.986	0.000	0.040
	78/79	0.940	0.000	0.000	0.049
	86/87	0.918	0.000	0.000	0.090
	87/88	0.719	0.000	0.082	0.189
	88/89	0.000	0.082	0.342	0.570
5.MANWEB	69/70	0.602	0.000	0.000	0.409
	70/71	0.619	0.000	0.000	0.384
	72/73	0.000	0.461	0.013	0.520
	73/74	0.000	0.449	0.000	0.551
	86/87	0.936	0.000	0.000	0.052
	87/88	0.545	0.000	0.000	0.462
	88/89	0.000	0.000	0.000	1.027
6.SOUTH WALES	69/70	0.000	0.000	0.000	1.013
	70/71	0.000	0.000	0.000	0.993
	87/88	0.000	0.000	0.000	1.024
	88/89	0.000	0.000	0.000	0.991
7.SOUTHERN	86/87	0.999	0.000	0.000	0.000
	87/88	0.314	0.498	0.186	0.000
	88/89	0.488	0.000	0.469	0.042
8.LONDON	69/70	0.999	0.000	0.000	0.000
	70/71	1.013	0.000	0.000	0.000
	85/86	0.000	0.000	0.976	0.000
	86/87	0.425	0.000	0.554	0.000
	87/88	0.000	1.002	0.000	0.000
	88/89	0.000	0.000	0.966	0.000
9.NORTHEAST	85/86	0.000	0.281	0.000	0.706
10.YORKSHIRE	69/70	0.804	0.000	0.000	0.178
	70/71	0.720	0.000	0.000	0.290
	74/75	0.000	0.469	0.000	0.529
	78/79	0.545	0.000	0.000	0.440
	88/89	0.491	0.000	0.045	0.483

high maximum demand capability. This is due to the peakiness of the demand from the domestic sector throughout the day which has to be met. There does seem to be a strong association between the efficiency derivation of maximum demand capability, and efficiency derivation of supply to the domestic sector. Those area boards whose efficiency rating over the years is derived from the maximum demand output also derive efficiency from the domestic sector. This need not necessarily be in the same years, but this makes sense, because those area boards which are good at supplying to the domestic sector have to be good at meeting peaks in demand.

7.5 Summary and conclusions.

In this paper, a new way of using DEA dynamically was implemented, pooling DMUs from a number of time periods to produce measurement of relative technical efficiency over time. It was applied to the Area Electricity Boards from 1969-1988. This time-series approach revealed which area boards were most efficient and showed trends in efficiency. More importantly, in doing so, it was shown to be superior to another form of the dynamic application of DEA, 'window analysis'. The Area Electricity Boards are a set of public utilities that has recently been privatised. The results produced by our application can yield information useful to a regulator.

The results revealed that there are fluctuations in average efficiency over the twenty year period. These fluctuations were tested in each area by correlating productive efficiency with regional GDP. There proved to be a strong correlation between them. The reason for this being so was deduced to be the fact that economic activity affects the demand and thus the supply of electricity. The inputs to electricity supply are fixed in the short-run so that spare capacity will arise if there is a slump in demand. These unused resources are a source of technical inefficiency. In a boom technical efficiency will rise because spare capacity is reduced.

Examining the efficiency of individual regions, there proved to be a large disparity between them. Measuring the technical efficiency, ATE, across each region for the twenty years showed that some were more efficient than others. For instance, MEB's ATE was 0.971, whilst in the South-West the ATE was 0.838.

The results produced using a dynamic analysis were judged to be superior in accuracy and clarity to window analysis. Our approach can also be used to validate an initial cross-section or produce results from disaggregated data. Window analysis can also show trends in efficiency but the efficiency rating in any one year cannot easily be determined. This is because the results for all analyses may be different in any one year because the peer group is changing. These shifting peer groups also mean that taking an average technical efficiency for an individual year is not meaningful.

So how can the information provided by this technique be useful to a regulator? In UK electricity distribution, the privatised utilities are to be regulated by a price-cap regime, subject to periodic reviews. It is at these periodic reviews that the regulator is required to compare the efficiency of the different distribution companies in order to set his separate price-cap incentives applying to each over the following years. The analysis carried out here could form a fundamentally important to such a comparative performance or yardstick exercise.

One piece of information that is yielded by DEA that might be useful to a regulator when examining the results is whether the Area Boards are producing at their optimal scale. Although they were chosen to be of similar size and scale DEA can confirm if this is still the case. If they were not similar then a comparison for regulatory purposes may not be a fair one.

To ascertain the position two different formulations of DEA were employed, one of constant returns to scale and one of variable returns to scale (see Banker, Charnes and Cooper [1984] and Fare, Grosskopf and Lovell [1985]). The ratio of the rating from the variable to the constant returns case reveals the degree of scale efficiency. Averaging across all efficiency points the ratio was very high indicating that scale inefficiency is not a major consideration.

Another result that this study has produced that an electricity regulator would need to know to compare Area Boards, is that there is very strong correlation between productive efficiency and regional GDP. Technical efficiency does not just depend on the behaviour of an Area Board but also unexpected changes in the demand for electricity. As changes in economic activity may occur at different times, with differing magnitudes in each region, a regulatory review would have to take this into account.

Our dynamic application of DEA would be a useful tool for measuring performance for a regulator to use, if they wished to relate prices for an Area Board to its performance compared with other similar utilities. Also, an analysis can take place even if there is asymmetric information in favour of the utility. This is because performance can be analysed on limited information. Greater insight can be gleaned from the results through the use of virtual inputs and outputs to identify the sources of efficiency.

FOOTNOTES.

1. This chapter is based on a PSERC University of Leicester working paper by Doble M. and Weyman-Jones T. (1991), 'Measuring productive efficiency in the area electricity boards of England and Wales using data envelopment analysis. A dynamic approach'.

2. One problem with using the size of the mains circuit as an input is that it discriminates against Area Boards which have a widely dispersed clientele. But, in this case the effect does not have wide significance. A DEA was conducted excluding this input and the rankings do not differ from the results of the main study.

Conclusion.

DEA represents a new and innovative way of measuring performance in the public sector. First suggested by Farrell in (1957), linear programming techniques have made it possible to measure technical efficiency through the estimation of non-parametric production frontiers. Given the multiplicity of objectives in public organisations, the one aspect of performance that is least in conflict with other aims is that of technical efficiency (the maximum output that can be achieved with given inputs, or the minimum inputs that can be used with given outputs). In the public sector, due to the absence of competition, market prices and costs may be lacking for some or all inputs or outputs. Profit can not then be calculated and is not necessarily the aim of the organisation anyway. DEA measures technical efficiency, it does not need information on prices or costs and produces a single efficiency criterion using data purely on measured volumes of inputs and outputs. These inputs and outputs can also be qualitative in nature which represents a useful step forward given the often intangible nature of public sector production.

I think DEA is superior to other attempts to measure the performance of public sector organisations. It has been traditional to use partial productivity ratios to measure efficiency and these have recently become popular in the public sector as 'performance indicators'. Their main drawback if that they do only provide a partial picture of performance and given the lack of equitable weighting systems to prioritise them, only of limited usefulness. Regression analysis would seem to be a better way of measuring technical efficiency, a regression line can be estimated and those DMUs that are above or below average technical efficiency can be identified. It is better than PIs because it can incorporate multiple inputs. But, the use of regression analysis does not equate with the theoretical definition of the production function as being a maximal concept. There will be positive as well as negative residuals and these lie above the regression line. A more correct way of interpreting the production function is that it is a frontier and this is what DEA estimates. There are econometric methods of calculating production frontiers but there is a danger that a parametric form may be imposed on the data which may not be warranted. DEA would thus seem to be a more theoretically correct method of measuring efficiency because it estimates a non-parametric production frontier.

This is not to say that DEA does not have its drawbacks or that there are not practical problems with its use. These have been explored through the conduct of two case studies. The results of which will be summarised and the implications for these organisations discussed before information realised about DEA is presented.

In Chapter 4 the three factors that made DEA useful in improving performance were established, inefficient DMU's can be identified, targets set for resource improvements and peer groups identified whose management practices can be observed to achieve best-practice (there was a fourth factor, using virtual inputs and outputs to examine the dimensions in which performance is derived and this is examined in the context of AEBs in Chapter 7). Post Office Counters is actually using DEA as an evaluative technique and was found to be using it in this way. These three stages were demonstrated in my case study of Counters.

The data used in the analysis was derived from 1281 Crown Post Offices for a 13 week period from September-November 1989. The input was labour (in hours) and the outputs were a quality variable (inverted average weighting time) and the outputs were ten different categories of transactions. It was calculated that scale efficiency was not a problem and the main body of results was derived from pure technical efficiency using the variable returns to scale, input minimisation program. It was found that overall Counters was inefficient as it had a mean technical efficiency of 0.814. Even so this concealed wide disparities of performance, only 13.82% achieved a rating of 1 and the worst DMU had a rating of 0.386. This variation was further explored on a regional basis and it was found that different regions had widely varying mean efficiencies. The most inefficient region was London with a mean efficiency of 0.705. The main reasons for this being the case were high turnover of staff, and the entrenchment of outmoded working practices due to the stronger unions in London.

The case study of the Area Electricity Boards produced similarly revealing results. A time-series approach was taken in this case on a pooled data set for 12 AEBs from 1969-88. Inputs used were labour (number of employees), the size of the mains distribution network (in circuit-kilometres) and transformer capacity (MVA). The outputs used were the maximal demand for electricity (in KW) and the quantity of electricity supplied (in KWH) to the domestic, commercial and industrial sectors. Whilst efficiency varied over time between individual AEBs, much more interesting was the change in average technical efficiency, over all AEBs, over time. It was surmised that technical efficiency changed in relation to economic conditions in the economy. To test this hypothesis, technical efficiency was correlated with regional GDP and it was found that there was a strong relationship. It was deduced that the reason for this was related to spare capacity in the use of resources. Capacity adjustments lagged behind changes in demand for electricity, in a slump electricity demand decreases and spare capacity increases, thus technical efficiency decreases. Conversely, in a boom when electricity demand increases, spare capacity decreases and technical efficiency rises. It was concluded that this had implications for the body now regulating the newly privatised AEBs because price-cap reviews will be related to measured productive efficiency.

Through the work that was conducted on these case studies, a number of issues were examined. They relate to the applicability and robustness of DEA as a method of measuring public sector performance. The key areas were clustering of DMUs to create more homogenous cross-sections and better peer groups, in the absence of hypothesis testing the robustness of the frontier must be established in other ways and also, can a time-series approach be used with DEA?

Clustering is a useful idea in theory, grouping DMUs by some common factor means that variables that could not be incorporated because the information is not available can be included. Efficiency rankings should thus be more accurate if a new frontier is estimated for each cluster. This is what was done with the CPOs which were clustered on a geographical basis by region. This was a perfectly valid basis on which to divide them, given the disparities of efficiency between regions and the fact that peer groups would be more localised and accessible. However, any gain in accuracy of efficiency ratings achieved by clustering was lost due to the fact that with smaller cross-sections measured efficiency rises. This also has implications for the accuracy of targets which will be reduced. It should have been possible to allocate new regional peer groups to inefficient CPOs. But, due to the fact that there are now a smaller number of inefficient DMUs, not many CPOs measured as being inefficient in the all-inclusive analysis can be given new peer groups. However, clustering can still be considered worthwhile because the DMUs that are allocated new regional peer groups are the most inefficient.

Given that DEA is a frontier rather than an average method of measuring technical efficiency, the frontier can be easily distorted by outliers. In the absence of hypothesis testing some means has to be found of checking their validity. A number of ways were looked at to try and achieve this aim. The number of 'peer group citations' (the number of DMUs for which the DMU forms the frontier) for efficient DMUs were examined. Clear results were produced, but there is some debate about the interpretation of PGCs as an indicator of robustness. Another criterion used was 'well-rounded performance' defined by how many weights (rather than evenly distributed virtual outputs) existed in the output dimensions. Conclusive results were hard to establish because the vast majority of CPOs had low numbers of weighted outputs. This suggests that the transaction groups have been brought down to a lower level of aggregation. Comparing the efficient DMUs PGCs to the number of weights each has was quite revealing because it transpired that in general the higher the number of PGCs, the less well-rounded performance was. So in the Counters context DMUs with high PGCs are forming the frontier for a large number of others because of their extreme resource use. Sensitivity analysis was used to validate those DMUs that were identified as being unusual at this point and to assess their impact on the efficiency of other DMUs. For DMUs with high PGCs, those output dimensions which define their efficiency were excluded from the analysis. On the criterion that their robustness is determined by the degree by which they become efficient, a few could not be considered to be good peers and should be excluded. Excluding a DMU that could be considered to be non-robust to assess its importance provided convincing evidence of the impact of an outlier. When DMU No.43 which has a very large peer group, was excluded and the programme rerun, most of its peers registered efficiency improvements. It is an outlier because it sells in one of the transactions categories very large volumes. This shows that results have to be related to DMU's outside the DEA context and that care has to be taken in the construction of the original model. It also highlights the far-reaching impact of an outlier, it has to accepted that DEA is not an infallible performance measure but unusual DMU's can be identified.

A new way of applying DEA is in a dynamic context. That is, pooling observations through time rather than a cross-section and running the program. This should show technical efficiency has changed over time as each DMU is measured relative to others and itself in different time periods. This is given the proviso that there is no technical progress (as was judged to be the case in the AEB application) as this will manifest itself as rising technical efficiency. There do not seem to be any drawbacks in its use and it was judged to be superior to 'window analysis', another suggested method for using DEA to measure technical efficiency over time. Window analysis can also show trends in efficiency but the efficiency rating in any one year cannot easily be determined. The results for all analyses may be different in any one year because the peer group is shifting.

DEA represents an agenda for continuing research, both theoretical and empirical. It is not proposed that this programme be discussed but there are a number of ways in which the work that has been conducted in this thesis can be built upon. There are a range of areas in which I would have liked to have developed this work further, particularly in the case of Counters.

Although on a theoretical level it was shown that DEA is superior to PIs and regression analysis as a tool for measuring efficiency, it would have been useful to show that this was the case with the data used. Some empirical work on the relative merits of the techniques has been conducted but further research is necessary.

The basis analysis for Counters could also have been improved. More information could have been yielded by including other inputs and outputs. Information about types of staff being used and their cost, in respect to their grades and the quantity of overtime would have shown if workers were being used effectively. A similar operation could have been conducted for the cost of offices, in terms of rental costs or the opportunity cost of capital for those branches owned outright. DEA could also have identified those CPOs who are occupying offices that are too large, if data on the square metres of floor space were available. Though of course examination of individual cases is important because aspects like floor plan and quality of installation will vary. Another interesting factor I would have liked to include is the effect of staff turnover on efficiency.

The basis of the clustered and regional analysis were the ten postal regions which existed before 1985. As this was the form the data was made available to me, I examined Counters using these regions rather than the new district and territorial organisation. This was not completely illogical because they were organised in this way for a long time, they should still be quite homogenous in terms of management. Nevertheless I would liked to have conducted this part of the study by dividing the data into the new territories and districts.

The analysis could also have been varied by conducting a dynamic DEA on Counters. This would have revealed information about their progress since re-organisation and how they have responded to the threat of privatisation in terms of successful management and increased efficiency. To this end it would also be interesting to follow their progress as they are now incorporating DEA into their performance measurement system. This utilisation of DEA shows that it possesses wide applicability to measuring public sector performance and I am confident that because of this it will gain increasing acceptance and use.

Glossary.

ASE Average scale efficiency.

ATE Average technical efficiency.

ATM Automatic teller machines.

AWT Average waiting time.

BTH Basic transaction hour. Basic transaction hours for each type of transaction are calculated as follows:

$$\frac{\text{Number of Transactions} \times \text{counter transaction time}}{3{,}600}$$

It is a measure of the time spent handling transactions at post office counters and takes no account of back-office work.

CA Counter arrivals per hour.

CC Counter cover.

CCR Carter committee report.

CMA Communications managers association.

COCSI Crown office staffing instructions. A full-scale inspection of a Crown Office which determines the number of staff required to handle counter transactions and back-office work.

COLS Corrected ordinary least squares.

CRS Constant returns to scale.

CT Customer transaction time

CTT Counter transaction time. The average time taken by a counter clerk to carry out a transaction, measured from the beginning of a customers request to the completion of the necessary action by the clerk.

CWU Counter work unit. A measure of workload used internally in the Counters business. This measure takes no acount of back-office work.

DEA Data envelopment analysis.

DMU Decision-making unit.

DRS Decreasing returns to scale.

DST Deprins, Simar and Tulkens.

DFT Dyson, Foster and Thanassoulis.

ECCO A project for the introduction of electronic cash registers.

FMI Fianancial management initiative.

IRS Increasing returns to scale.

KWH Kilowatt hour.

MVA Megavoltamp.

ML Maximimum likelihood.

MLE Maximimum likelihood estimators.

MMC Monopolies and Mergers Commission.

NIRS Non-increasing returns to scale.

NPV Net present value.

PGC Peer group citation.

PI Performance indicator.

POC Post Office Counters.

POPOS Post Office point of sale. Standardised display of leaflets and information on behalf of clients in the public section of post offices.

PTE Pure technical efficiency.

RCWSA Revised counter and writing staffing agreement. A negotiated agreement between the Post Office and the Union of Communication Workers which facilitates greater staffing flexibility.

RUC Real unit cost.

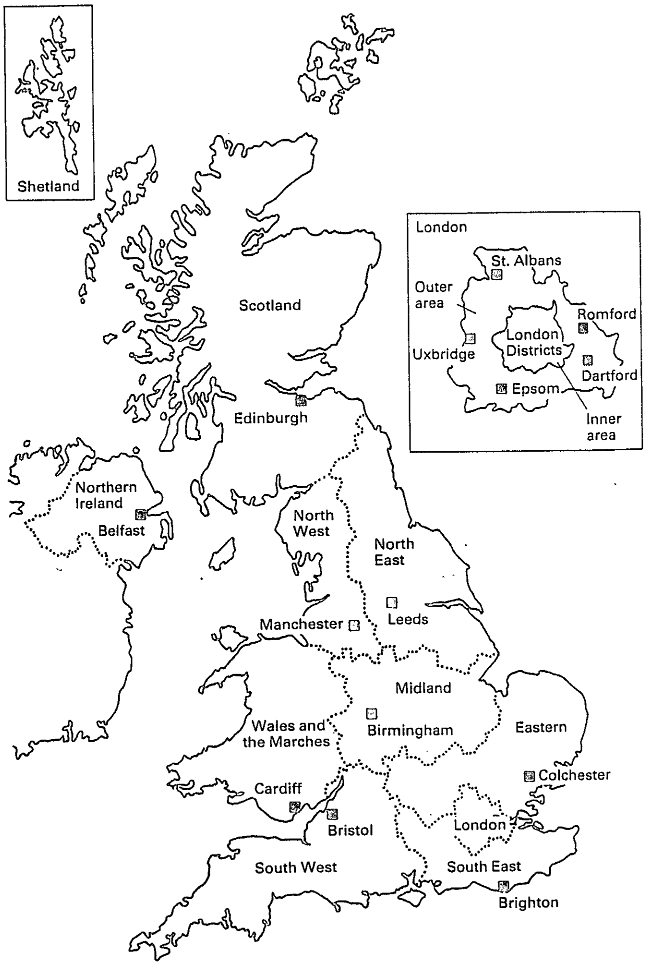
SE Scale efficiency.

TE Technical efficiency.

TFP Total factor productivity.

VRS Variable returns to scale.

Appendix 1. Map of UK postal regions.

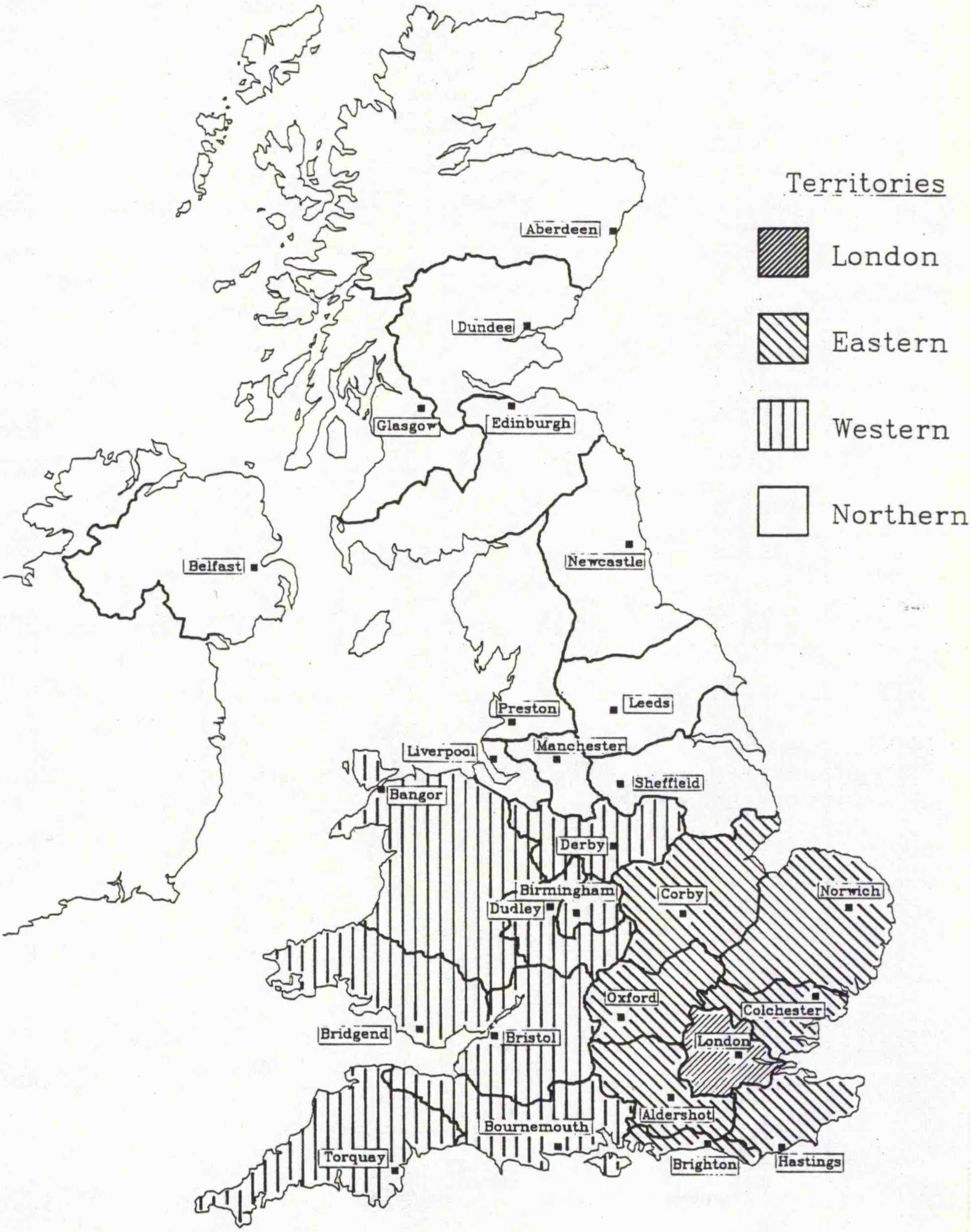


Source: MMC (1984). The Post Office Letter Post Service. Cmnd 9332. HMSO. London.

Appendix 2.

Territory	Location of territory office	District	Location of district office
London	London EC1	Central	London
		East	London
		North	London
		South-East	London
		South-West	London
		West	London
Eastern	Colchester	Aldershot	Aldershot
		Brighton	Brighton
		Colchester	Colchester
		Corby	Corby
		Hastings	Hastings
		Norwich	Norwich
		Oxford	Oxford
Western	Birmingham	Bristol	Bristol
		North and mid-Wales and the Marches.	Bangor
		Birmingham and Coventry	Birmingham
		Wessex	Bournemouth
		South Wales	Bridgend
		North Midlands	Derby
		South and West Midlands	Dudley
		Devon and Cornwall	
Northern	Leeds	Leeds	Leeds
		North of Scotland	Aberdeen
		Northern of Ireland	Belfast
		Edinburgh	Edinburgh
		Glasgow	Glasgow
		Dundee	Dundee
		Lancashire and Cumbria	Preston
		Liverpool	Liverpool
		Manchester	Manchester
		Newcastle-upon-Tyne	Newcastle
		Sheffield	Sheffield

Appendix 3. Map of territories and districts.



Source: MMC (1988). Post Office Counters Services. Cm 398. HMSO. London.

Bibliography.

Albon, R.P. (1987), 'Privatise the post: steps towards a competitive postal service,' Centre for Policy Studies, Policy Study no.82, London.

Anthony R.N. (1965), Planning and Control Systems: A Framework for Analysis, Division of Research, Graduate School of Business Administration, Harvard University, Boston.

Anthony R.N. and Herzlinger R.E. (1980), Management control in Non-Profit Organisations, Irwin, Homewood, Illinois.

Anthony R.N. and Dearden, J. (1980), Management control systems, Irwin, Homewood, Illinois.

Afriat, S. (1972), 'Efficiency estimation of production functions,' International Economic Review, 13, 2, 568-598.

Aigner, D., Lovell, C. and Schmidt, P. (1977), 'Formulation and estimation of stochastic frontier production functions,' Journal of Econometrics, 5, 21-38.

Aigner, D.J. and Chu, S.F. (1968), 'On estimating the industry production function,' American Economic Review, 58, 4, 826-839.

Audit Commission (1986), Performance review in local government. Action guide, HMSO, London.

Banker, R.D. (1984), 'Estimating most productive scale size using data envelopment,' European Journal of Operational Research, 17, 1, 35-44.

Banker, R.D. Charnes A. and Cooper W.W. (1984), 'Some models for estimating technical and scale efficiencies in data envelopment analysis,' Management Science, 30, 9, 1078-1092.

Banker R.D., Charnes A. and Cooper W.W. (1984), 'Some models for estimating technical and scale efficiencies in data envelopment analysis,' 30, 9, 1078-1092.

Banker R.D., Conrad R.F. and Strauss R.P. (1986), 'A comparative application of DEA and translog methods: an illustrative study of hospital production,' Management Science, 32, 1, 30-44.

Banker, R.D. and Maindiratta, A. (1986), 'Piecewise loglinear estimation of efficient production surfaces,' Management Science, 32, 1, 126-135.

Banker, R. and Morey R. (1986), The use of categorical variables in Data Envelopment Analysis. Management Science, 32, 12, 513-521.

Baumol, W.J., Panzar, J.C. and Willig, R.D. (1982), Contestible markets and theory of industry structure, Harcourt, Brace, Jovanovich, San Diego, California.

Beer, S. (1966), Decision and Control, London.

Beesley, M. and Littlechild, S. (1983), 'Privatisation: Principles, Problems and Priorities,' Lloyds Bank Review, 149, 1-20.

Bitran, G.R. and Valor-Sebatier, J. (1985), 'Some mathematical programming based measures of efficiency in health care institutions,' Sloan School of Management Working Paper. Massachusetts Institute of Technology. Massachusetts.

Bös, D. (1986), Public enteprise economics: Theory and application, North-Holland, Amsterdam.

Bowlin, W.F. (1987), 'Evaluating the efficiency of US Air Force real-property maintenance activities,' Journal of the Operations Research Society', 38, 2, 127-35.

Bowlin, W.F., Charnes, A., Cooper, W.W. and Sherman, H.D. (1985), 'Data envelopment analysis and regression approaches to efficiency estimation and evaluation,' Annals of Operations Research, 2, 113-138.

Boyd, G. and Fare, R. (1984), 'Measuring the efficiency of decision making units: A comment,' European Journal of Operational Research, 15, 331-332.

British Telecommunications Act, (1981), Section 6, HMSO, London.

Charnes, A. and Cooper, W.W. (1962), 'Programming with linear fractional functionals,' Naval Research Logistics Quarterly, 9, 181-185.

Charnes, A. and Cooper, W.W. (1973), 'An explicit general solution in linear fractional programming,' Naval Research Logistics Quarterly, 20, 3.

Charnes, A. and Cooper, W.W. (1984), 'The non-Archimedean CCR ratio for efficiency analysis: A rejoinder to Boyd and Fare', European journal of operational research, 15, 333-334.

Charnes, A., Cooper, W.W., Lewin, A.Y., Morey, R.C. and Rousseau, J. (1985), 'Sensitivity and stability analysis in DEA', Annals of operations research, 2, 139-156.

Charnes, A., Cooper, W.W. and Rhodes, E. (1978), 'Measuring the efficiency of decision making units,' European Journal of Operational Research, 2, 6, 429-444.

Charnes, A., Cooper, W.W. and Rhodes, E. (1979), 'Short communication. Measuring the efficiency of decision making units,' European Journal of Operational Research, 3, 4, 339.

Charnes, A., Cooper, W.W. and Rhodes E. (1981), 'Data envelopment analysis as an approach for evaluating program and managerial efficiency with an illustrative application to the program follow through experiment in US public school education,' Management Science, 27, 6, 668-697.

Charnes, A., Clark, C.T., Cooper, W.W., and Golany B. (1985), 'A developmental study of data envelopment analysis in measuring the efficiency of maintentenance units in the US air force', Annals of Operations Research, 2, 95-112.

Cubbin, J. and Ganley. J (1987), 'Total factor productivity measurement in British education: a non-parametric approach. Working paper, London business school.

Curwen P.J. (1986), Public enteprise. A modern approach. Wheatsheaf, Brighton.

Cyert, R.M. and March, J.G. (1963), A behavioural theory of the firm, Prentice-Hall Inc., Eaglewood Cliffs, New Jersey.

Deprins D., Simar L. and Tulkens H. (1984), Measuring labor-inefficiency in post offices. In Marchand M., Peastieau, P. and Tulkens H. (Eds.), (1984), 'The performance of public enteprises: concepts and measurement.' North-Holland. Amsterdam.

Doble, M. and Weyman-Jones T. (1991), 'Measuring productive efficiency in the area electricity boards of England and Wales. A dynamic approach,' Working Paper, Public Sector Economics Research Centre, University of Leicester.

Drake L. and Weyman-Jones T.G. (1991), Technical and scale efficiency in UK building societies. Loughborough University Working Paper.

Dyson, R.G., Foster M.J. and Thanassoulis E. (1987), 'Relative efficiency assessments using data envelopment analysis: an application to data on rates departments,' Journal of the Operational Research Society, 38, 5, 397-411.

Epstein M.K. and Henderson J.C. (1987), 'Data envelopment analysis for managerial control and diagnosis,' Sloan School of Management Working Paper. Massachusetts Institute of Technology. Massachusetts.

Fare R., Grosskopf, S. and Lovell, C.A.K. (1983), 'The structure of technical effciiency,' Scandinavian Journal of Economics, 85, 2, 181-190.

Fare, R., Grosskopf, S., and Lovell, C.A.K. (1985) The Measurement of Efficiency of Production, Kluwer- Nijhoff Publishing, Boston.

Farrel, M.J. (1957), 'The measurement of productive efficiency.' Journal of the Royal Statistical Society, A, 120, 253-266.

Feldstein, M.S. (1967), Economic analysis for health service efficiency: Econometric studies of the British national health service, North-Holland, Amsterdam.

Field K. (1990), Production efficiency of british building societies. Applied Economics, 22, 415-246.

Financial Times, (1986), 19 October.

Flamholz, E.G. (1979), 'Behavioural aspects of accounting/control systems,' in Kerr, S. (ed.) Organisational Behaviour, Grid, Columbus, Ohio.

Førsund, F.R., Lovell, C.A.K. and Schmidt, P. (1980), 'A survey of frontier production functions and their relationship to efficiency measurement,' Journal of Econometrics, 13, 5-27.

Ganley J.A. (1989), Relative efficiency measurement in the public sector with data envelopment analysis. Unpublished Phd. thesis, London school of economics, London.

Golany B. and Roll Y. (1989), An application procedure for DEA. Omega: International Journal of Management Science. 17, 3, 237-250.

Greene, W.H. (1980), 'Maximum likelihood estimation of econometric frontier functions,' Journal of Econometrics, 13, 27-56.

Greene, W.H. (1982), 'Maximum likelihood estimation of stochastic frontier production models,' Journal of Econometrics, 18, 285-289.

Grosskopf, S. (1986), 'The role of the reference technology in measuring productive efficiency,' Economic Journal, 96, 499-513.

Grosskopf, S. and Valdamis V. (1987), Measuring hospital performance. A non-parametric approach. Journal of Health Economics. 6, 89-107.

Hyman, H.H. (1942), 'The psychology of states,' Archives of Psychology, 269.

Jesson, D. and Mayston, D. (1989), 'Measuring performance in authorities and schools,' in, Harrison, A. and Gretton, J. (Eds.)(1989), Education and Training U.K. 1989, Policy Journals, Newbury.

Jesson D., Mayston D. and Smith, P. (1987), 'Performance assessment in the education sector: educational and economic perspectives,' Oxford Review of Education, 13, 3, 249-266.

Johnston, J. (1960), Statistical cost analysis, McGraw-Hill, New York.

Jondrow, J., Lovell, C.A.K., Materov, I. and Schmidt, P. (1982), 'On the estimation of technical inefficiency in the stochastic frontier production function model,' Journal of Econometrics, 19, 233-238.

Kanter, R.M. (1987), The change masters, Simon and Schuster, London.

Kennedy C, Thirlwall, A.P. (1972), 'Surveys in applied economics: technological progress,' Economic Journal, 82, 11-72.

Kirthisingha, H.A., Land, A. and Shutler, M.F. (1990), 'Branch operation,' Working Paper, Monopolies and Mergers Commission, London.

Kopp, R. (1981), 'The measurement of productive efficiency: a reconsideration,' Quarterly Journal of Economics, 96, 477-503.

Kopp, R. and Diewert, E. (1982), 'The decomposition of frontier cost function deviations into measures of technical and allocative efficiency,' Journal of Econometrics, 19, 319-331.

Lawler, E.E. and Rhode, J.G. (1976), Information and Control in Organisations, Goodyear, Pacific Palisades, California.

Levitt, M.S. and Joyce, M.A.S. (1987), 'The growth and efficiency of public spending,' Cambridge University Press, Cambridge.

Lewin, A.Y. and Minton, J.W. (1986), 'Determining organisational effectiveness,' Management Science, 5, 32.

Lewin, A. and Morey, R. (1981), 'Measuring the relative efficiency and output potential of public sector organisations: An application of data envelopment,' International Journal of Policy Analysis and Information Systems, 5, 267-285.

Lewin, A. Morey R. and Cook. T. (1982), 'Evaluating the administrative efficiency of courts,' Omega: The International Journal of Management Science, 10, 410-411.

Lovell, C.A.K., Walter L.C. and Wood L.L. (1989), 'Exploring the distribution of DEA efficiency scores,' Working paper.

March, J.G. and Simon, H.A. (1958), Organisations, John

Wiley and Sons, New York.

Marchand, M., Pestieau P. and Tulkens H. (1984,a), 'The performance of public enteprises: normative, positive and empirical issues,' 243-267, in Marchand, M., Pestieau P. and Tulkens H. (1984,b), (eds.), The performance of public enteprises: Concepts and measurement, Elsevier, North-Holland, Amsterdam.

Mayston, P. and Smith, P. (1987), 'Measuring efficiency in the public sector,' Omega: International Journal of Management Science, 15, 3, 181-189.

Meeusen, W. and Broeck, van den. (1977), 'Efficiency estimation from Cobb-Douglas production functions with composed error,' International Economic Review, 18, 2, 435-444.

Merchant, K. (1985), Control in Organisation, Pitman, Marshfield, Massachusetts.

Merton, R.K. (1957), 'The role-set: Problems in sociological theory,' British Journal of Sociology, 8, 106-120.

Molyneux, R. and Thompson, D. (1986), The efficiency of the nationalised industries since 1978, IFS Working Paper No.100, The Institute for Fiscal Studies.

Monopolies and Mergers Commission (1980), Inner London letter post, HC 515, HMSO, London.

Monopolies and Mergers Commission (1984), The Post Office letter post service, Cmnd. 9332, HMSO, London.

Monopolies and Mergers Commission (1986), The Post Offices procurement activities, Cmnd 9912, HMSO, London.

Monopolies and Mergers Commission (1988), Post Office Counters Services, Cm 398, HMSO, London.

Morgan, G. (1986), Images and Organisation, Sage, London.

National Audit Office (1986), Report of the comptroller and auditor general, The Rayner scrutinies, 1979 to 1983, HMSO, London.

National Economic Development Council, (1976), A study of UK nationalised industries, NEDO.

Nunamaker, T. (1985), 'Using data envelopment analysis to measure the efficiency of non-profit organisations: A critical evaluation,' Managerial and Decision Economics, 6, 1, 50-58.

Ouchi, W.G. (1979), 'A conceptual framework for the design of organisational control mechanisms,' Management Science.

Parkan, C. (1987), 'Measuring the efficiency of service operations: an application to bank branches,' Engineering Costs and Production Economics, 12, 237-242.

Parris, H., Pestieau, P. and Saynor, P. (1986), Public Enteprise in Western Europe, Croom Helm, London.

Pestieau, P. (1989) 'Measuring the Performance of Public enteprises. A must in times of privatization,' Annals of Public and Cooperative Economics, 60, 3, 293-305.

Post Office Reports and Accounts.

Post Office Review Committee (C.F. Carter, Chairman) (1977), Report, Cmnd 6850, HMSO, London.

Prakash, P. and Rappaport, A. (1977), 'Information inductance and its significance for accounting,' Accounting, Organisations and Society, 1, 2.

Pryke, R. (1980), 'Public enteprise in practise: the British experience of nationalisation during the past decade', in Baumol, W. (ed.), Public and private enteprise in a mixed economy, St. Martins Press, New York.

Rangan N., Aly H.Y., Pasurka C. and Grabowski R. (1988) The technical efficiency of US banks. Economic Letters, 28, 169-175.

Rees, R. (1984), Public enteprise economics, Weidenfeld and Nicholson, London.

Rees, R. (1985), 'The theory of principal and agent,' Part One and Part Two, Bulletin of Economic Research, 37 (no. 1 and 2).

Richmond, J. (1974), 'Estimating the efficiency of production,' International Economic Review, 15, 2, 515-521.

Saaty, T.L. (1980), The analytic hierarchy process: planning, priority setting, resource allocation., McGraw-Hill, New York.

Schmidt, P. (1986), 'Frontier production functions,' Econometric Reviews, 4, 289-238.

Schmidt, P. and Lovell, C. (1979), 'Estimating technical and allocative inefficiency relative to stochastic production and cost frontiers,' Journal of Econometrics, 9, 343-366.

Seiford, L.M. (1990), 'A bibliography of data envelopment analysis (1978-1990),' Working Paper, Department of Industrial Engineering and Operations Research, University of Massachusetts, Amhurst.

Sengupta J. and Sfeir R. (1986), Production frontier estimates of scale in public schools in California. Economics of Education Review. 5, 3, 297-307.

Senior, I. (1983), Liberating the Letter. A proposal to privatize the Post Office, Institute of Economic Affairs, Research Monograph No.38.

Sexton T.R., Silkman R.H. and Hogan A.J. (1986), Data envelopment analysis: critique and extensions. In Silkman R.H. (Ed.) (1986), 'Measuring efficiency: an assessment of data envelopment analysis'. Jossey-Bass. San Fransisco.

Sherman, H.D. (1981), Measurement of hospital technical efficiency: a comparative evaluation of data envelopment analysis and other approaches for locating inefficiency in health care organizations, DBA thesis, Graduate School of Business Administration, Harvard University, Boston.

Sherman H.D. and Gold F. (1985), 'Bank branch operation efficiency: Evaluation with data envelopment analysis,' Journal of Finance, 9, 297-316.

Siegel, S.S. (1956), Nonparametric Statistics for the Behavioural Sciences, McGraw-Hill, New York.

Silkman, R.H. (ed.)(1986) Measuring Efficiency: an assessment of data envelopment analysis, Jossey Bass Inc., San Fransisco.

Smullen J. (1989), The application of data envelopment analysis to branch planning and management appraisal. Presented at the European Working Group of Operational Research in Banking (Eurobanking), Heemskerk, The Netherlands, May (1989).

Taylor F.W. (1911), Shop management, Harpers and Bros, New York.

Thomas, D.L. (1985), Auditing the efficiency of regulated companies through the use of data envelopment analysis: An application to electric co-operatives, Phd thesis, University of Texas Graduate School of Business, Austin.

Treasury and Civil Service Committee, (1982), Efficiency and effectiveness in the civil service, 3rd report, session 1981-82. Cmnd 8616. HMSO. London.

Tulkens, H. (1986), 'The Performance Approach in Public Enteprise Economics: An introduction and an example,' The Annals of Public and Co-operative Economy, 57, 4, 429-53.

UK Electricity Council, (1989), Handbook of Electricity Supply Statistics, The Electricity Council, London.

United Nations Department of Technical Co-operation for Development (1989), Role and extent of competition in improving the performance of public enteprises, United Nations, New York.

Vancil, R.F. (1973), 'What kind of management control do you need?' Harvard Business Review.

Vassioglou M. Giokas D. (1990), A study of the relative efficiency of bank branches: an application of data envelopment analysis. Journal of the Operational Research Society, 41, 7, 591-597.

Vicker, J. and Yarrow, G.(1988), Privatization. An Economic Analysis, MIT Press, London.

Waterson, M. (1988), Regulation of the firm and natural monopoly, Blackwell, Oxford.

Westwick, C.A. (1987), How to use management ratios, Gower, Aldershot.

Weyman-Jones, T.G. (1991), 'Productive efficiency in a regulated industry: the Area Electricity Boards of England and Wales', Energy Economics, 13, 2.