DATA ENVELOPMENT ANALYSIS

A TECHNIQUE FOR MEASURING THE EFFICIENCY OF GOVERNMENT SERVICE DELIVERY

Steering Committee for the Review of Commonwealth/State Service Provision

1997

ISBN: 0 646 33533 2

Secretariat
Steering Committee for the Review of Commonwealth/State Service Provision
Industry Commission
LB2 Collins Street East Post Office
Melbourne VIC 8003

Level 28
35 Collins Street
Melbourne VIC 3000

Telephone: 03 9653 2100 Facsimile: 03 9653 2199
E-mail: IC@mail.indcom.gov.au

An appropriate citation for this report is:

Steering Committee for the Review of Commonwealth/State Service Provision 1997, *Data Envelopment Analysis: A technique for measuring the efficiency of government service delivery,* AGPS, Canberra.

PREFACE

Effective government service provision benefits from the support of rigorous measurement techniques. Data Envelopment Analysis (DEA) is an analytical tool that can assist in the identification of best practices in the use of resources among a group of organisations. Such identification can highlight possible efficiency improvements that may help agencies to achieve their potential.

The aim of this paper is to promote a better understanding of the DEA technique. It explains DEA's conceptual underpinnings, how to interpret the output from DEA models and its strengths and weaknesses. Also, through the use of case studies on hospitals, dental services, police, motor registries, and corrective services, this paper provides a practical guide to developing and refining a DEA model and interpreting of results.

This paper is directed at those responsible for providing government services and those accountable for their delivery in a cost effective manner. It should encourage people to think about how more detailed and rigorous analysis of performance can assist in improving the efficiency with which resources are used to provide essential services to the community.

DEA can be a very useful analytical technique by providing an important 'first step' tool in comparative analysis. But users also need to recognise its limitations as an input to the development of public policy. Its theoretical predictions of potential efficiency gains may not be translatable into actual gains when factors such as service quality, fundamental differences between individual services and the costs of implementing changes are fully accounted for. Non-efficiency objectives such as access and equity are also important policy considerations for government, against which efficiency benefits will inevitably be balanced.

The Steering Committee wishes to thank the service agencies that were involved in the case studies and their staff for their enthusiasm and assistance. I would also like to thank, on behalf of the Steering Committee, the DEA Working Group which was responsible for preparing the paper.

Bill Scales, AO
Chairperson
Steering Committee for the Review
of Commonwealth/State Service Provision

CONTENTS

DEA WORKING GROUP

Jurisdiction	*Member*
New South Wales	
NSW Treasury	Mr John Pierce (Convenor)
NSW Treasury	Mr Roger Carrington
NSW Treasury	Ms Nara Puthucheary
NSW Treasury (from January 1997)	Ms Kathryn Kang
Victoria	
Department of the Treasury, Victoria	Mr Barry Bond
Department of the Treasury, Victoria	Mr Dick Elvins
Department of the Treasury, Victoria	Ms Kerry Macdonald
Queensland	
Queensland Treasury	Ms Trish Santin
Queensland Treasury	Ms Jennifer Pallesen (until January 1996)
Queensland Treasury	Mr Rick Stankiewicz (until January 1996)
Tasmania	
Department of the Treasury and Finance	Mr Richard Mackey (until March 1996)
Northern Territory	
Department of the Chief Minister	Mr Phil Temple
NT Treasury	Mr Don Parker
Commonwealth Government	
Department of Finance	Mr Richard Mackey (from April 1996)
Local Government	
Local Government & Shires Associations of NSW	Mr Murray Kidnie
Local Government & Shires Associations of NSW	Ms Lorraine Slade
Secretariat	
Industry Commission	Mr Jeff Byrne
Industry Commission	Mr Rob Bruce
Industry Commission	Mr Tendai Gregan
External reviewers	
Australian Bureau of Statistics	Mr Steven Kennedy
Tasman Asia Pacific	Dr Denis Lawrence
University of Georgia	Professor Knox Lovell

ABBREVIATIONS

CRS	Constant returns to scale
DEA	Data envelopment analysis
RTA	Roads and Traffic Authority (NSW)
SCNPMGTE	Steering Committee on National Performance Monitoring of Government Trading Enterprises
SCRCSSP	Steering Committee for the Review of Commonwealth/State Service Provision
TFP	Total factor productivity
WIES	Weighted inlier equivalent separations

GLOSSARY

Allocative efficiency	Whether, for any level of production, inputs are used in the proportion which minimises the cost of production, given input prices.
Benchmarking	The process of comparing the performance of an individual organisation against a benchmark, or ideal, level of performance. Benchmarks can be set on the basis of performance over time or across a sample of similar organisations, or against some externally set standard.
Best practice	In this context, the set of management and work practices which results in the highest potential, or optimal, quantity and combination of outputs for a given quantity and combination of inputs (*productivity*) for a group of similar organisations. Best practice can be identified at a number of levels, including organisational, national and international.
Congestion	A situation in which an organisation has unwanted or surplus inputs and would incur a net cost to reduce those inputs. For example, redundancy payments associated with reducing staff levels will result in a net cost to an organisation if they are higher than the savings in wages for a given period.
Cost efficiency	Where an organisation is *technically efficient* and *allocatively efficient* and, hence, produces a given quantity, quality and mix of outputs at minimum possible cost given existing knowledge of technologies and people’s preferences.
Data Envelopment Analysis (DEA)	A *linear programming* technique which identifies *best practice* within a sample and measures *efficiency* based on differences between observed and BEST PRACTICE units. DEA is typically used to measure *technical efficiency*.
Decision Making Units	The organisations or units being examined in a DEA study. In public sector studies, these units may not be commercial or profit-making entities.
Dynamic efficiency	Success with which producers alter technology and products following changes in consumer preferences and productive opportunities.

Effectiveness	Degree to which the outputs of a service provider achieve the stated objectives of that service — for example, the extent to which hospitals are meeting the demand for non-elective surgery. In the case of government service providers, the government normally sets such objectives.
Efficiency	Degree to which the observed use of resources to produce outputs of a given quality matches the optimal use of resources to produce outputs of a given quality. This can be assessed in terms of *technical efficiency* and *allocative efficiency*.
External operating environment	Factors which affect the providers of outputs that are not in the direct control of managers — for example, weather, client wealth and in some cases input prices.
Human services	Services provided by core government agencies, such as education, health, welfare and justice.
Linear program	A set of linear mathematical equations for which a solution can be obtained subject to an upper bound (maximisation) or a lower bound (minimisation).
Non-scale technical efficiency	Proportion of *technical efficiency* which cannot be attributed to divergences from optimal scale (*scale efficiency*); sometimes known as managerial efficiency or pure technical efficiency.
Outputs	Goods and services provided to entities or persons outside the production unit.
Partial productivity indicator	Ratio of the quantity of an output (or the combined quantities of a subset of total outputs) to the quantity of an input (or the combined quantities of a subset of total inputs) where some inputs or outputs are not included. For example, output per employee does not include the other inputs required to produce the output, such as raw materials, semi-finished goods and capital.
Peers	In DEA studies, the group of best practice organisations with which a relatively inefficient organisation is compared.
Production frontier	The curve plotting the minimum amount of an input (or combination of inputs) required to produce a given quantity of output (or combination of outputs).

Production technology	Relationships incorporated in production processes which determine the manner in which inputs can be converted to outputs.
Productivity	Measure of the physical output produced from the use of a given quantity of inputs. This may include all inputs and all outputs (*total factor productivity*) or a subset of inputs and outputs (*partial productivity*). Productivity varies as a result of differences in *production technology*, differences in the *technical efficiency* of the organisation, and the *external operating environment* in which production occurs.
Returns to scale	Relationship between output and inputs. Returns can be constant, increasing or decreasing depending on whether output increases in proportion to, more than or less than inputs, respectively. In the case of multiple inputs and outputs, this means how outputs change when there is an equi-proportionate change in all inputs.
Scale efficiency	The extent to which an organisation can take advantage of returns to scale by altering its size towards optimal scale (which is defined as the region in which there are constant *returns to scale* in the relationship between outputs and inputs).
Slacks	The extra amount by which an input (output) can be reduced (increased) to attain *technical efficiency* after all inputs (outputs) have been reduced (increased) in equal proportions to reach the *production frontier*. This is a feature of the piece-wise linear production frontier derived when using DEA.

Technical efficiency	Conversion of physical inputs such as labour services and raw materials or semi-finished goods into *outputs*. Technical efficiency is determined by the difference between the observed ratio of combined quantities of an entity's output to input and the ratio achieved by *best practice*. It can be expressed as the potential to increase quantities of outputs from given quantities of inputs, or the potential to reduce the quantities of inputs used in producing given quantities of outputs.
	Technical efficiency is affected by the size of operations (*scale efficiency*) and by managerial practices (*non-scale technical efficiency*). It is defined independent of prices and costs.
Total factor productivity (TFP)	Ratio of the quantity of all outputs to the quantity of all inputs. TFP can be measured by an index of the ratio of all outputs (weighted by revenue shares) to all inputs (weighted by cost shares).
Yardstick competition	Competition over performance levels where no market exists for the goods or services concerned. This competition relies on performance indicators.

1 IMPROVING THE PERFORMANCE OF GOVERNMENT SERVICE PROVIDERS

Measuring the efficiency of government service provision can play an important role in achieving improvements in the performance of Australia's human service delivery. Data envelopment analysis is a technique that can be used to assist in identification of best practice performance in the use of resources, highlight where the greatest gains may be made from improvements in efficiency, and help agencies achieve their potential.

1.1 Improving performance in government funded service delivery

Government funded services contribute about 10 per cent to Australia's gross domestic product and affect the daily lives of all Australians. For example, governments make available education and training, health and community services; they maintain law and order; and they provide subsidised care and shelter for people in need. Improvements in the performance of these government funded 'human' services have the potential to deliver major social and economic benefits and to improve our capacity to address social needs more effectively.

Human service delivery performance is coming under increased scrutiny as part of the ongoing process of microeconomic reform. In the absence of market contestability, comparative performance reporting provides a way of introducing competitive pressures for government service providers. Developing and reporting performance indicators is crucial to identifying performance improvements, and thus guiding decision making. Comparative performance reporting is typically undertaken cooperatively to assist all participants to improve their performance; it is not focused on determining if a specific level of performance is being achieved.

The Steering Committee for the Review of Commonwealth/State Service Provision, established in 1993, is developing and reporting jurisdiction based performance indicators for human services funded by governments. The performance indicator reports of the Steering Committee, released in 1995 and 1997, have concentrated on the areas of health, public housing, education, justice and community support services.

1.2 Who uses performance measures?

Performance varies in human service provision across jurisdictions and within jurisdictions. Providing an indication of how much performance differs and which organisations are the best performers is potentially of value to the providers of funds and the clients of these services — members of the community — as well as to those managing the service provision — governments, departments and service providers. Comparable performance measures are of most value when they:

- are linked to service objectives and aspects of provision for which there is responsibility and accountability; and
- relate to aspects of service provision for which there is limited competitive market pressure. The resulting comparisons of performance measures, or yardstick competition, can provide an alternative form of pressure for improved performance.

Concerned citizens are able to use publicly available information on the performance of different service providers to make governments more accountable for the expenditure of taxpayer funds, and to exercise client choice more effectively.

Governments can use performance measures to:

- stimulate policy development by highlighting the effect on the performance of government determined aspects of the operating environment (for example, client choice, extent of competition);
- facilitate monitoring of public sector managerial performance and improve accountability within government;
- promote ‘yardstick competition’ by providing a means of comparing the performances of those responsible for similar aspects of service provision where there is little direct competition in input and/or output markets;
- analyse the relationships between agencies and between programs, to allow governments to coordinate policy across agencies (for example, the interrelationships between policing, courts and correctional services); and
- assist the resource allocation/budgeting process by providing a means of allocating funding based on agreed plans for improved performance, rather than on the assumption that performance should equal past levels.

Comparative performance measurement is also a powerful management tool for both agency managers (such as department heads) and individual service provider managers (for example, hospital or police station managers).

Managers can use performance measurement to:

- identify differences in performance; and

- focus attention on other organisations which may be performing better. More detailed exercises may help identify the practices being used by the other organisations, along with the extent to which those practices could usefully be adopted in the operations under review.

The objective of comparative performance measurement is to facilitate a program to improve performance, not to provide a simple grading of service providers.

Identifying major gaps in performance can provide the impetus for an organisation to fundamentally rethink how it does things. There has been much focus on the 'continuous improvement' approach to managing government organisations, but this approach may limit managers to looking at only small changes to procedures. If a quantum leap is needed, rather than an incremental change, then monitoring can help managers focus on where a major overhaul of the organisation's operations may be required. Most organisations have at least some aspect of their operations from which others can learn, and in the absence of direct competition, sharing information is the best way of transferring best practice.

The process of performance measurement has the value of identifying performance variations, and hence providing encouragement and direction for performance improvement. There are also two wider benefits of the process which can be equally important in supporting performance improvements.

First, measuring performance requires a clear understanding and articulation of the resources being used, and the outputs being produced, in the process of providing a service. Making the inputs and outputs transparent can allow a critical assessment of why particular resources are being used to provide particular outputs, clarifying service provision objectives and priorities.

Second, attempting to measure performance provides a heightened awareness of data shortcomings for managers and policy makers. If data deficiencies are catalogued and advertised the quality of data may be improved, and the ability to better measure performance enhanced.

There may be a hesitancy to try new approaches to measuring the performance of human services given concerns that data are not of sufficient quality. However, as these points illustrate, a useful start can usually be made on performance measurement with data that are currently available — waiting for the perfect data may lead to extensive delay, and the use of available data is often a catalyst for developing better quality data.

No single performance measure or technique can provide the whole answer; quantitative analysis involves significant assumptions and limitations. Consequently, it may be desirable to use the results of several approaches, both quantitative and qualitative, to judge how a particular agency is performing overall and what needs to be done. Thus, the Steering Committee is interested in the development and application of new techniques and approaches to performance measurement.

1.3 Concepts of effectiveness, efficiency and productivity

Assessing the performance of human service providers is a complex task. The Steering Committee has developed a framework of effectiveness and efficiency with which to assess government funded service provision.

Effectiveness is the extent to which outputs of service providers meet the objectives set for them by governments. An example is whether hospitals meet the waiting time targets set for elective surgery.

Efficiency is the success with which an organisation uses its resources to produce outputs — that is the degree to which the observed use of resources to produce outputs of a given quality matches the optimal use of resources to produce outputs of a given quality. This can be assessed in terms of technical, allocative and dynamic efficiency. (Definitions of the different types of efficiency are developed in Chapter 2 and are included for reference in the Glossary.)

Improving the performance of government service relies on both efficiency and effectiveness. A government service provider might increase its measured efficiency at the expense of the effectiveness of its service. For example, a hospital might reduce the inputs used for cleaning but service the same number of patients. This could increase the apparent efficiency of the hospital but reduce its effectiveness in providing satisfactory outcomes for patients. Therefore, it is important to develop effectiveness indicators for government service providers. For example, are patients being re-admitted to hospitals at unacceptable rates? Effectiveness is more fully discussed in the Steering Committee report for 1997 (SCRCSSP 1997).

This paper focuses on the assessment of the technical efficiency of government service providers. Technical efficiency is determined by the difference between the observed ratio of combined quantities of an organisation's output to input and the ratio achieved by best practice. Producing the maximum output or consuming the minimum inputs, as compared to what is technically feasible, is an essential step for service providers to be able to best meet their objectives.[1]

[1] See Pestieau and Tulkens (1993) for a fuller discussion of the relationship between technical efficiency and the ability of public enterprises to achieve their objectives.

The technical efficiency of an organisation depends on its level of productivity, or the ratio of its outputs to inputs.[2] The following could lead to productivity improvements:

- the adoption of technological advances; and/or
- increases in efficiency through for example the removal of unnecessarily restrictive work practices or better management; and/or
- a change in the operating environment in which production occurs.

All agencies use a range of inputs, including labour, capital, land, fuel and materials, to produce services. If an agency is not using its inputs in a technically efficient manner, it is possible to increase the quantities of outputs without increasing inputs, or to reduce the inputs being used to produce given quantities of outputs.

The Steering Committee has begun the task of assessing efficiency by compiling information on simple, partial productivity indicators — ratios of output to input which do not include all outputs and inputs. Efficiency indicators reported so far include, for example, measures of recurrent expenditure and staff and/or capital per unit of a particular output.

Partial productivity measures and recurrent costs per unit of output are used widely because they are simple to calculate, but they need to be interpreted with caution. By definition these measures are always only partial in that they do not account for the relationships and trade-offs between different inputs and outputs. This is a significant limitation in their application to government service delivery, which typically involves multiple inputs and outputs. For example, if labour inputs are replaced by capital inputs, labour productivity is likely to increase while capital productivity declines. To assess whether the agency has become more efficient overall, output must be measured against both labour and capital inputs.

Several partial productivity measures and recurrent costs per unit of output may be used collectively to obtain a broad picture of efficiency. However, the presentation of a large number of partial measures will typically be difficult to comprehend and interpret if some indicators move in opposite directions over a given period of time. This reinforces the value of more comprehensive summary measures of efficiency. Partial measures may provide important information on specific aspects of operation, but it is important to see how the agency is performing overall relative to comparable organisations producing similar outputs.

---

[2] Productivity is an absolute concept, measured by the ratio of outputs to inputs, while efficiency is a relative concept, measured by comparing the actual ratio of outputs to inputs with the optimal ratio of outputs to inputs.

Given the shortcomings of partial indicators, governments are adopting more comprehensive indicators of performance for government trading enterprises and government service providers. In the case of government trading enterprises, the techniques adopted include total factor productivity (TFP) indexing (see Box 1) and data envelopment analysis (DEA).

Box 1: Performance measurement in government trading enterprises

The Steering Committee on National Performance Monitoring of Government Trading Enterprises was established following the July 1991 Special Premiers' Conference to develop a national performance monitoring scheme for government trading enterprises across the Commonwealth, states and territories. The Committee has developed a consistent suite of financial and non-financial performance indicators of efficiency and effectiveness, which are reported by participating government trading enterprises annually (see SCNPMGTE 1996).

To measure the performance of a government trading enterprise, increasing use is being made of total factor productivity (TFP) indexing — a procedure which combines all outputs and inputs into a comprehensive measure of overall productivity. Important to this process, the Steering Committee on National Performance Monitoring of Government Trading Enterprises published a guide to using TFP and examples of its application in several case studies (SCNPMGTE 1992).

An international benchmarking program for key Australian infrastructure industries, started at the Bureau of Industry Economics in 1991 and now being undertaken by the Industry Commission, has used DEA, TFP and partial productivity measures for benchmarking.

However, the TFP technique is not generally applicable to service provision, because it requires a price for each output's and inputs and output prices often cannot be identified for many government services. Thus, DEA is being used more for government service providers. As well as being able to handle multiple services and inputs, DEA does not require information on the price of services or inputs, making it particularly applicable to government service provision.

DEA is a linear programming technique that identifies the apparent best providers of services by their ability to produce the highest level of services with a given set of inputs, or to produce given services with the least amount of inputs. Other service providers receive an efficiency score that is determined by their performance relative to that of the best performers. The technique can also determine whether the main source of inefficiency is the scale of operations or the managerial capabilities and effort of the service provider. Further, it can incorporate variables to account for environmental factors which might influence the productivity of a service provider but which are beyond its control — for example, the education or wealth of clients.

Like any empirical technique, DEA has limitations of which practitioners need to be mindful (these are discussed in more detail in the following chapters). DEA results provide the maximum benefit when they are interpreted with care. In general, they should be considered as a starting point for assessing the efficiency of the service providers within a sample. Indications of possible sources of relative inefficiency can guide further investigation to determine why there are apparent differences in performance. This information can be used to inform the managers of individual service providers, administrators and policy makers.

Finally, it is important to recognise that performance measures will inevitably evolve through time. Gaining experience in formulating and using the measures, the agency will refine the set to better meet its requirements. Agencies might start off with relatively simple measures and progress to more sophisticated measures as they gain experience and as they initiate the collection of better quality data.

1.4 Objectives of this paper

The Steering Committee established a Working Group to promote the application of DEA to government services. The DEA Working Group has produced this information paper to further understanding of the technique and to outline its application in a number of case studies focussing on how relative efficiency is measured and on how results can be used to enhance performance.

The following chapter provides a relatively non-technical explanation of the principles behind DEA. This is followed by a simple example of how to calculate DEA in Chapter 3, and an overview of five case studies of the application of DEA (highlighting practical issues encountered) in Chapter 4. Chapter 5 contains the case studies, which cover Victorian hospitals, Queensland oral health services for students, and NSW corrective services, police patrols, and motor registries. Appendix A supports the second and third chapters, and Appendix B lists software options for applying DEA.

Performance measurement of human service delivery is still in its infancy. Much remains to be done in getting the necessary data systems consistently in place, and in resolving the precise nature and value of many human service outputs and how to measure them. DEA is useful for improving the performance of government service delivery by advancing our understanding of key efficiency drivers and identifying examples of good practice.

2 WHAT IS DATA ENVELOPMENT ANALYSIS?

Data envelopment analysis provides a means of calculating apparent efficiency levels within a group of organisations. The efficiency of an organisation is calculated relative to the group's observed best practice. This chapter explains the basic concepts behind DEA and provides a simple graphical example. Some extensions to the DEA model, allowing the sources of inefficiency to be identified, are also discussed.

2.1 Data envelopment analysis and different efficiency concepts

Typically using linear programming, DEA calculates the efficiency of an organisation within a group relative to observed best practice within that group. The organisations can be whole agencies (for example, Departments of Health), separate entities within the agency (for example, hospitals) or disaggregated business units within the separate entities (for example, wards).[1]

To discuss DEA in more detail it is necessary to look at the different concepts of efficiency. The most common efficiency concept is *technical efficiency*: the conversion of physical inputs (such as the services of employees and machines) into outputs relative to best practice. In other words, given current technology, there is no wastage of inputs whatsoever in producing the given quantity of output. An organisation operating at best practice is said to be 100 per cent technically efficient. If operating below best practice levels, then the organisation's technical efficiency is expressed as a percentage of best practice. Managerial practices and the scale or size of operations affect technical efficiency, which is based on engineering relationships but not on prices and costs.

Allocative efficiency refers to whether inputs, for a given level of output and set of input prices, are chosen to minimise the cost of production, assuming that the organisation being examined is already fully technically efficient. Allocative efficiency is also expressed as a percentage score, with a score of 100 per cent indicating that the organisation is using its inputs in the proportions which

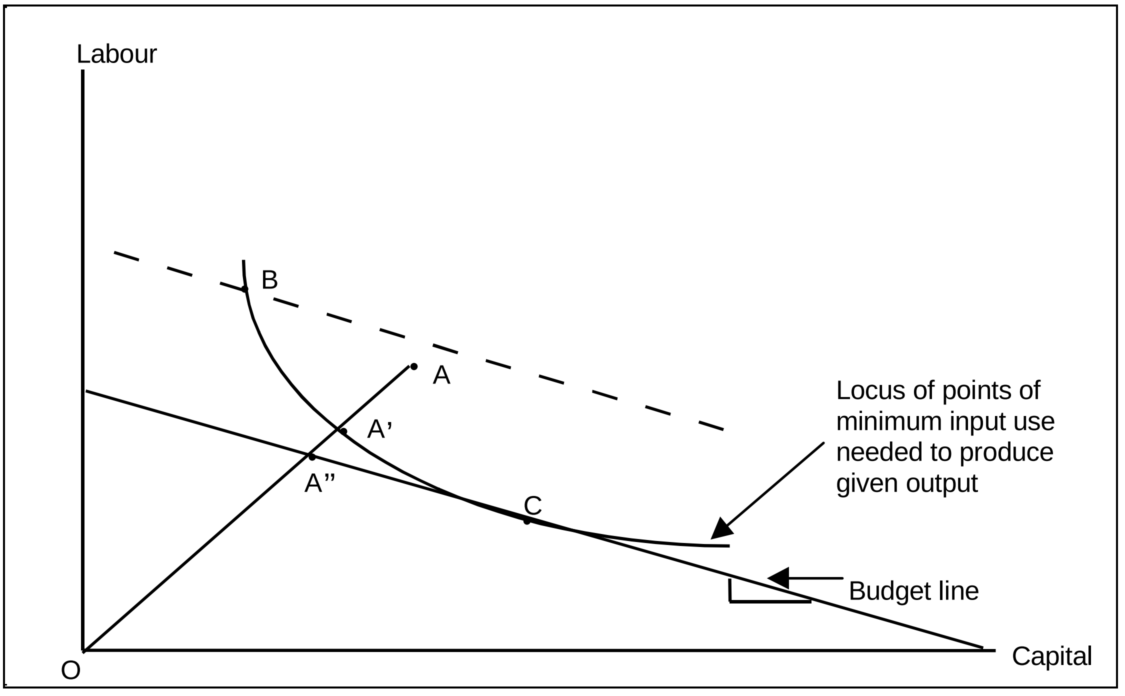
---

[1] Given that DEA is particularly well suited to government service and other non-profit organisations, as well as private sector firms, the individual units examined are often referred to as decision-making units rather than firms.

would minimise costs. An organisation that is operating at best practice in engineering terms could still be allocatively inefficient because it is not using inputs in the proportions which minimise its costs, given relative input prices.

Finally, *cost efficiency* refers to the combination of technical and allocative efficiency.[2] An organisation will only be cost efficient if it is both technically and allocatively efficient. Cost efficiency is calculated as the product of the technical and allocative efficiency scores (expressed as a percentage), so an organisation can only achieve a 100 per cent score in cost efficiency if it has achieved 100 per cent in both technical and allocative efficiency.

Figure 2.1: Illustration of different efficiency concepts



These concepts are best depicted graphically, as in Figure 2.1 which plots different combinations of two inputs, labour and capital, required to produce a given output quantity. The curve plotting the minimum amounts of the two inputs required to produce the output quantity is known as an isoquant or efficient frontier. It is a smooth curve representing theoretical best engineering practice. Producers can gradually change input combinations given current technological possibilities. If an organisation is producing at a point on the

[2] Cost efficiency is sometimes extended to include a third measure called dynamic efficiency: the degree to which producers respond to changes to technology and products following changes in consumer preferences and productive opportunities.

isoquant then it is technically efficient. The straight line denoted as the budget line plots combinations of the two inputs that have the same cost. The slope of the budget line is given by the negative of the ratio of the capital price to the labour price. Budget lines closer to the origin represent a lower total cost. Thus, the cost of producing a given output quantity is minimised at the point where the budget line is tangent to the isoquant. At this point both technical and allocative efficiencies are attained.

The point of operation marked A would be technically inefficient because more inputs are used than are needed to produce the level of output designated by the isoquant. Point B is technically efficient but not cost efficient because the same level of output could be produced at less cost at point C. Thus, if an organisation moved from point A to point C its cost efficiency would increase by (OA–OA")/OA. This would consist of an improvement in technical efficiency measured by the distance (OA–OA')/OA and an allocative efficiency improvement measured by the distance (OA'–OA")/OA'. Technical efficiency is usually measured by checking whether inputs need to be reduced in equal proportions to reach the frontier. This is known as a ‘radial contraction’ of inputs because the point of operation moves along the line from the origin to where the organisation is now.

2.2 Operationalising the concepts

The smooth curve in Figure 2.1 representing theoretical best practice typically cannot be calculated from observed data. Rather, data usually are only available on a group of organisations which gives limited information on theoretical best practice. First, it is unknown whether any of the organisations in the group, or sample, are achieving outright best practice. Second, the sample points will not cover all of the range of possible input combinations.

There are several ways to use the data from the sample to try and approximate the smooth curve in Figure 2.1. Early attempts used ordinary least squares regression techniques, that plot an average curve through the sample points. However, this was not satisfactory because an individual organisation’s efficiency was compared with an average level of performance in the sample rather than an estimate of best practice within the sample. This led to attempts to approximate best practice in the sample by estimating frontiers.

The two techniques used to estimate the frontier are DEA and stochastic frontier analysis. The focus in this report is on DEA, which is a deterministic means of constructing a ‘piece-wise linear’ approximation to the smooth curve of Figure 2.1 based on the available sample. In simple terms, the distribution of sample points is observed and a ‘kinked’ line is constructed around the outside of them, ‘enveloping’ them (hence the term data envelopment analysis).

Stochastic frontier analysis is an alternative approach using regression techniques. It tries to take account of outliers which either are very atypical or appear to be exceptional performers as a result of data measurement errors. The relevance of stochastic frontier analysis to budget sector applications is limited to those situations in which a single overall output measure or relatively complete price data are available. This is not often the case for service providers, so stochastic frontier analysis is not covered in this information paper. (See Fried, Lovell and Schmidt 1993 for a discussion of stochastic frontiers.)

DEA is often only used to calculate the technical efficiency of government services. The DEA approach to calculating technical efficiency can be shown with a simple numerical example: a sample of five hospitals that use two inputs — nurses and beds — to produce one output — treated cases. Obviously the inputs and outputs of a real hospital are considerably more complex, but this simplification may be a good starting point for actual as well as illustrative examples — for instance, the input ‘beds’ might serve as a proxy for the amount of capital inputs used by the hospital. The hospitals are likely to be of differing sizes; to facilitate comparisons, input levels must be converted to those needed by each hospital to produce one treated case. The hospital input and output data are presented in Table 2.1.

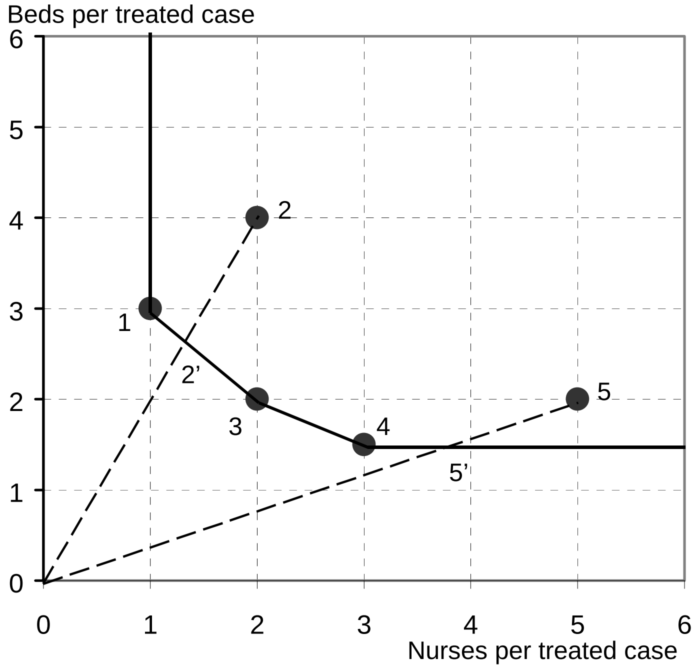
Table 2.1: Illustrative hospital data

Hospital	*Nurses*	*Beds*	*Treated cases*	*Nurses per treated case*	*Beds per treated case*
1	200	600	200	1	3
2	600	1200	300	2	4
3	200	200	100	2	2
4	600	300	200	3	1.5
5	500	200	100	5	2

The five hospitals range in size from 200 to 1200 beds, and there is a similarly large range in the numbers of nurses, beds, treated cases, and nurses per treated case and beds per treated case. Given the large discrepancies in the five hospitals’ characteristics it is not obvious how to compare them or, if one is found to be less efficient, which other hospital it should use as a role model to improve its operations. The answers to these questions become clearer when the data for nurses per treated case and beds per treated case are plotted in Figure 2.2, where data are abstracted from differences in size.

The hospitals closest to the origin and the two axes are the most efficient, so a ‘kinked’ frontier can be drawn from hospital 1 to hospital 3 to hospital 4. For the moment, the parts of the frontier above hospital 1 and to the right of hospital 4 are drawn by extending the frontier beyond these points parallel to the respective axes. The kinked frontier in Figure 2.2 envelopes all the data points and approximates the smooth isoquant in Figure 2.1 based on the information available from the data.

Figure 2.2: Illustrative hospital input–output data



Which are the most efficient or best practice hospitals in the sample? Hospitals 1, 3 and 4 are on the efficient frontier, so are assumed to be operating at best practice. However, hospitals 2 and 5 are north-east of the frontier, so are considered to be less efficient. This is because they appear to be able to reduce their input use and still maintain their output level compared with the performance of the best practice hospitals. For example, hospital 2 could reduce its use of both inputs by one third before it would reach the efficient frontier at point 2’. Similarly, its technical efficiency score is given by the ratio 02’/02 which is equal to 67 per cent in this case. This is because the ‘hypothetical’ hospital 2' has a value of 1.33 for nurses per treated case and a value of 2.67 for beds per treated case. In terms of actual input levels, hospital 2 would have to reduce its number of nurses from 600 to 400 and its number of beds from 1200 to 800. At the same time, it would have to maintain its output of 300 treated cases before it would match the performance of the hypothetical best practice hospital 2'.

But how is the hypothetical best practice hospital 2' derived? It is formed by reducing the inputs of hospital 2 in equal proportions until reaching the best practice frontier. The frontier is reached between hospitals 1 and 3 in this case, so the hypothetical hospital 2' is a combination, or weighted average, of the operations of hospitals 1 and 3. If hospital 2 is looking for other hospitals to use as role models to improve performance, then it should examine the operations of hospitals 1 and 3 because these are the efficient hospitals most similar to itself. In DEA studies these role models are known as the organisation's 'peers'.[3]

The other less efficient hospital — hospital 5 — is in a different situation. It is north-east of the efficient frontier, but contracting its inputs in equal proportions leads to the hypothetical hospital 5', which still lies to the right of hospital 4 on the segment of the frontier which was extended parallel to the nurses per treated case axis. Thus, the peer group for hospital 5 solely consists of hospital 4 because it is the only one which 'supports' that section of the frontier on which the hypothetical 5' lies. But hospital 5' is not fully efficient because the number of nurses per treated case can be reduced, while the number of beds per treated case is held constant, thus moving from 5' back to 4. That is, to maximise its efficiency given the available data, hospital 5 has to reduce one input more than the other. In this special case, a radial contraction of inputs means that the frontier is reached, but a further reduction of one of the inputs can be achieved without a reduction in output. This extra input reduction available is known in DEA studies as an input 'slack'. Thus, it is important in DEA studies to check for the presence of slacks as well as the size of the efficiency score.

It is relatively easy to implement this simple example of data envelopment analysis in a two-dimensional diagram. However, with a larger number of inputs and outputs and more organisations, it is necessary to use mathematical formulae and computer packages. Using the same principles, an example of how to implement a more complex analysis is given in Chapter 4 and the technical details behind DEA are briefly presented in Appendix A. Before moving on to look at some extensions to the basic DEA model outlined above, some of the questions DEA can help agency managers answer are briefly reviewed.

---

3 The term 'peers' in DEA has a slightly different meaning to the common use of the word peer. It refers to the group of best practice organisations with which a relatively less efficient organisation is compared.

2.3 What questions can DEA help us answer?

By providing the observed efficiencies of individual agencies, DEA may help identify possible benchmarks towards which performance can be targeted. The weighted combinations of peers, and the peers themselves may provide benchmarks for relatively less efficient organisations. The actual levels of input use or output production of efficient organisations (or a combination of efficient organisations) can serve as specific targets for less efficient organisations, while the processes of benchmark organisations can be promulgated for the information of managers of organisations aiming to improve performance.

The ability of DEA to identify possible peers or role models as well as simple efficiency scores gives it an edge over other measures such as total factor productivity indices.

Fried and Lovell (1994) listed the following as questions that DEA can help to answer for managers:

- How do I select appropriate role models to serve as possible benchmarks for a program of performance improvement?
- Which production facilities are the most efficient in my organisation?
- If all my operations were to perform according to best practice, how many more service outputs could I produce and how much could I reduce my resource inputs by, and in what areas?
- What are the characteristics of efficient operating facilities and how can they guide me in choosing locations for expansion?
- What is the optimum scale for my operations and how much would I save if all my facilities were the optimum size?
- How do I account for differences in external circumstances in evaluating the performance of individual operating facilities?

The simple model of DEA already outlined can satisfy the first four of these questions. To answer the last two, some extensions to the model are needed.

2.4 Extensions to the DEA model

By making the DEA model a little more complicated, the range of topics it can explore is increased. Particularly interesting is the decomposition of the technical efficiency score into components resulting from: the scale of operations; surplus inputs which cannot be disposed of; and a residual or ‘pure’ technical efficiency. A further extension which is often important is to allow for differences in operating environments; this involves trying to adjust for factors which might be beyond managers' control, and which thus possibly give some organisations an artificial advantage or disadvantage. Each of these issues is addressed in turn below. A technical treatment of these topics is presented in Appendix A.

2.4.1 Scale efficiency

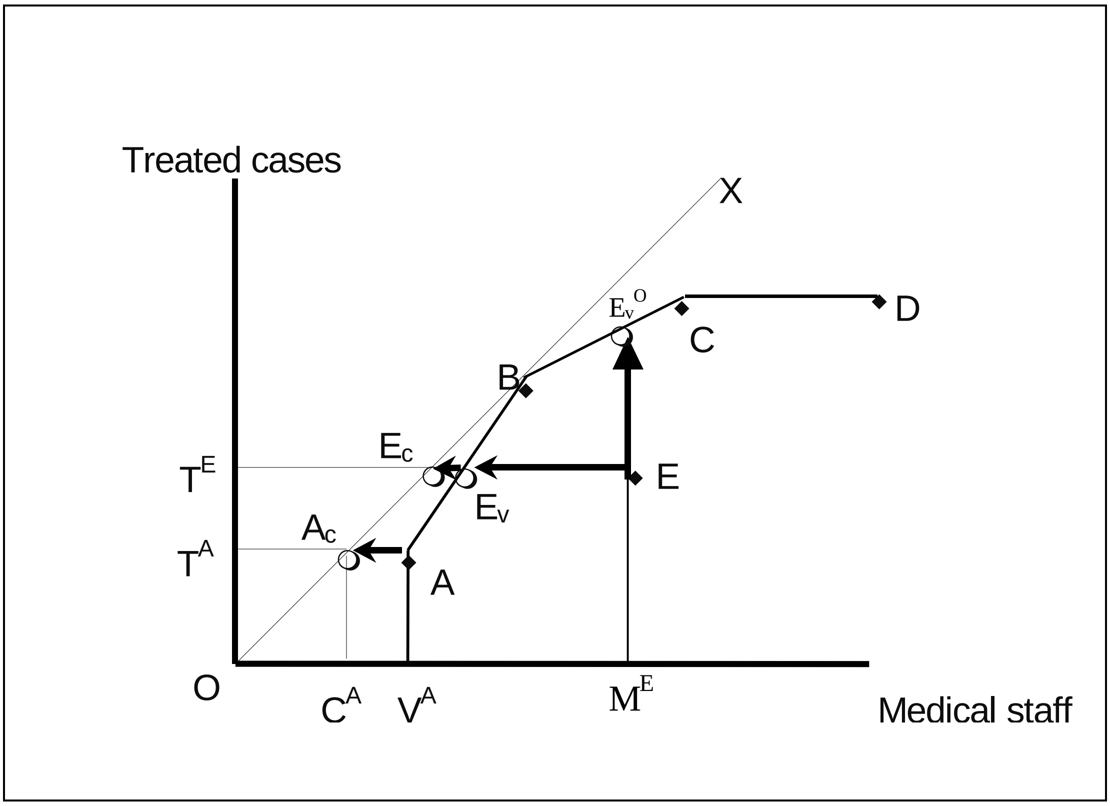
The simple example presented in Section 2.2 was based on the assumption of constant returns to scale. Given this assumption, the size of the organisation is not considered to be relevant in assessing its relative efficiency. Small organisations can produce outputs with the same ratios of input to output as can larger organisations. This is because there are no economies (or diseconomies) of scale present, so doubling all inputs will generally lead to a doubling in all outputs. However, this assumption is inappropriate for services which have economies of scale (or increasing returns to scale). In these services, doubling all inputs should lead to more than a doubling of output because producers are able to spread their overheads more effectively or take advantage of purchasing items in bulk. For other services, organisations might become too large and diseconomies of scale (or decreasing returns to scale) could set in. In this case, a doubling of all inputs will lead to less than a doubling of outputs. It would be to an agency's advantage to ensure that its operations are of optimal size — neither too small if there are increasing returns nor too large if there are decreasing returns to scale.

If it is likely that the size of service providers will influence their ability to produce services efficiently, the assumption of constant returns to scale is inappropriate. The less restrictive variable returns to scale frontier allows the best practice level of outputs to inputs to vary with the size of the organisations in the sample. This is demonstrated using the simplified one input (medical staff), one output (treated cases) example shown in Figure 2.3.

The constant returns to scale frontier is the straight line emanating from the origin (OBX), determined by the highest achievable ratio of outputs to inputs in the sample, regardless of size. The variable returns to scale frontier (V^AABCD) passes through the points where the hospitals have the highest output to input ratios, given their relative size, then runs parallel to the respective axes beyond the extreme points. The scale efficiency of an organisation can be determined by comparing the technical efficiency scores of each service producer under constant returns to scale and variable returns to scale.

The distance from the respective frontier determines technical efficiency under each assumption. The distance between the constant returns and the variable returns frontiers determines the scale efficiency component. Technical efficiency resulting from factors other than scale is determined by the distance from the variable returns frontier. Thus, when efficiency is assessed under the assumption of variable returns, the efficiency scores for each organisation indicate only technical inefficiency resulting from non-scale factors. Technical efficiency scores calculated under variable returns, therefore, will be higher than or equal to those obtained under constant returns.

Figure 2.3: The production frontier and returns to scale



This can be demonstrated using the examples in Figure 2.3.

- Hospital B is the only one that has no scale or non-scale inefficiency under either assumption. It represents the *optimal scale* within the sample.
- Hospitals A, C and D are scale inefficient but do not have any inefficiency resulting from non-scale factors under the variable returns assumption. For example, the scale efficiency score of hospital A is determined by the ratio of distances T^AA_c/T^AA, which is less than one. Hospital A has increasing returns to scale because it would approach the optimal scale in the sample if it increased its size. Hospitals C and D are producing outputs with decreasing returns to scale and are too large to be considered scale efficient, with hospital D being the furthest from optimal scale.
- The technical inefficiency of hospital E under constant returns (T^EE_c/T^EE) is made up of both scale inefficiency (T^EE_c/T^EE_v) and non-scale technical inefficiency (T^EE_v/T^EE).

2.4.2 Input and output orientation

Another issue that can be illustrated in Figure 2.3 is the question of output and input orientation. The examples so far have been input oriented — that is, by how much can inputs be reduced while maintaining the same level of output? However, the corresponding output-oriented question could be equally important — by how much can output be increased while keeping the level of inputs constant? The latter question may be more relevant for many government service providers, particularly those supplying human services, as the community often wants more of these services while budgetary pressures make it difficult to increase inputs.

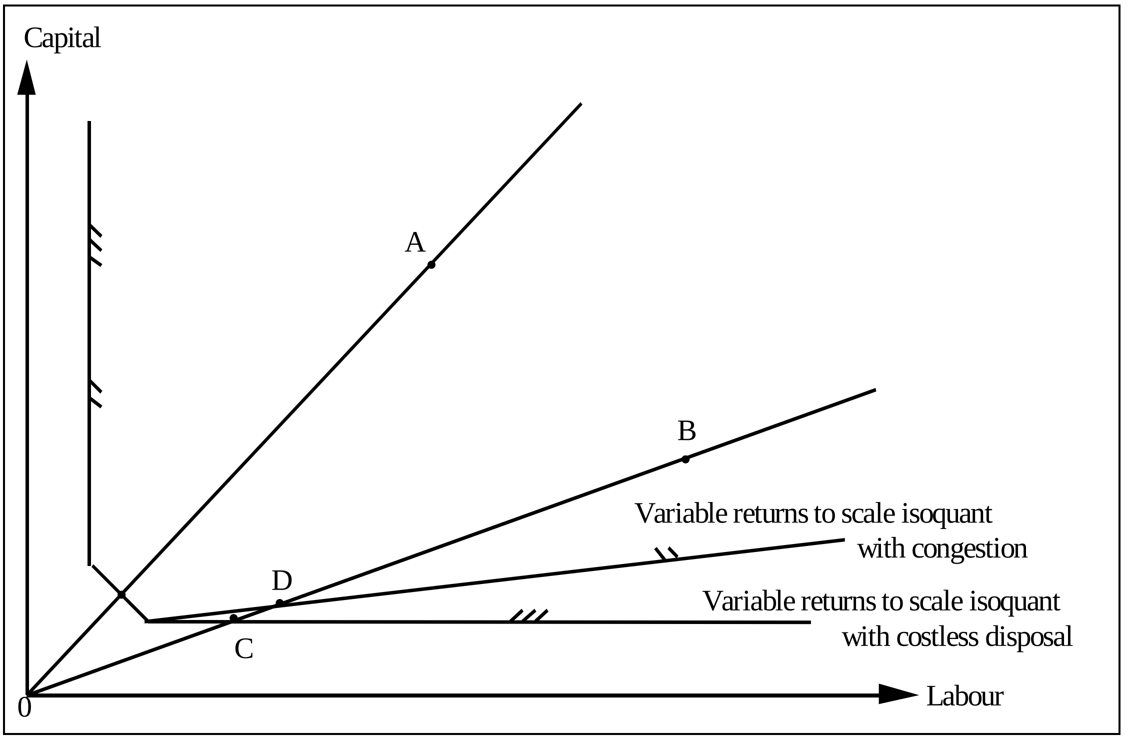
In Figure 2.3 the input-oriented technical efficiency score for hospital E under variable returns to scale was given by the ratio of distances T^EE_v/T^EE. The technical efficiency score for hospital E, using an output orientation and again assuming variable returns to scale, is given by the ratio of distances $M^EE/M^EE_v^O$. If an organisation is technically inefficient from an input-oriented perspective, then it will also be technically inefficient from an output-oriented perspective. However, the values of the two technical efficiency scores typically will differ, as will the presence and extent of slacks.

Depending on whether an input-saving or output-expanding orientation is utilised, the peers for hospital E will also differ. Its peers are hospitals A and B under input orientation but hospitals B and C under output orientation. This reflects the fact that hospital E can learn different things from the two sets of peers. Hospital C is better at producing more output from a roughly similar input level to that of hospital E, while hospital A produces less output than does hospital E but uses considerably fewer inputs.

2.4.3 Input congestion

In many situations agencies will not be able to dispose of 'unwanted' inputs. For example, union restrictions may inhibit reductions in the workforce, or government controls may prohibit reducing certain inputs. To cover situations such as these, Färe, Grosskopf and Lovell (1983) introduced the concept of congestion efficiency (See Figure 2.4). In the previous examples, the constant returns to scale isoquant is eventually parallel to both axes (as in Figure 2.2). This reflects the assumption that an input that cannot be profitably used can be disposed of at no cost. In contrast, under congestion, the inability to dispose of unwanted inputs increases costs.

Figure 2.4: Measuring input congestion



In Figure 2.4, congestion is assumed to be present in the use of labour. Thus, instead of the frontier eventually running parallel to the labour axis, congestion is reflected in a frontier which slopes upwards from the labour axis. In Figure 2.4, an organisation operating at point B would have congestion inefficiency equal to OC/OD, whereas a firm operating at point A would be congestion efficient.

After decomposing the constant returns technical efficiency score into components resulting from scale efficiency and congestion efficiency, a residual or ‘pure’ technical efficiency score remains. This residual score largely indicates the scope for efficiency improvements resulting from less efficient work practices and poor management, but may also reflect differences between operating environments.

2.4.4 Adjusting for operating environments

The environment in which a service provider operates can have an important influence on its relative performance if other providers are operating in different environments. Many of these operating environment factors are not under the control of managers, and ignoring them in assessing performance may lead to spurious results. Climate, topography, the socioeconomic status of a neighbourhood, government restrictions and the degree of unionisation, for example, can affect performance but be beyond management control.

The efficiency score of a police station in a poor area, for example, may not be comparable with the score of a police station in a well-to-do area. This may be misleading if the level of crime is lower in well-to-do neighbourhoods and if the police stations’ output is measured by the number of incidents attended and arrests. Thus, it may be important to adjust for the impact of the respective socioeconomic environments on incidents attended and arrests. Only then may it be possible to determine which police station is being more successful at transforming inputs (the number of police and cars) into outputs (the number of incidents attended and arrests).

But how to adjust for these differences in operating environments which are beyond management control? The main options in DEA studies are to:

- only compare organisations which operate in a similar operating environment. This may sound attractive but it often dramatically reduces the potential number of comparison partners and, hence, much information is lost: for instance, the main lessons may come from organisations that operate differently in a dissimilar environment;
- only compare the organisation with other organisations in a similar or less favourable operating environment. This overcomes some of the problems of the preceding method but still ignores potential lessons from more favourable operating environments;
- include the operating characteristic as part of the DEA calculation. This method is useful where the direction of influence of the characteristic is obvious. However, the characteristic has to be a continuous variable and,

by including more variables in the analysis, the efficiency scores tend to be automatically inflated. This method was used by the Bureau of Industry Economics (BIE 1994b) to adjust for the effects of climate on the observed efficiency of gas suppliers in different countries. Suppliers in colder climates have a higher demand for gas and are able to achieve better capital use than suppliers in warmer countries. By including a measure of degree-days — the number of days per year which deviate from an average temperature by more than a specified range — the Bureau was able to make more like-with-like comparisons;

- employ a two-stage procedure which uses econometric methods to estimate the relationship between the characteristic and the efficiency scores.[4] The efficiency scores can be adjusted on the basis of this relationship. The advantages of this approach are that it can accommodate several characteristics, makes no prior assumptions about the direction of influence, and allows for tests of statistical significance — something which usually is not possible in DEA studies. The technique is similar to that used in the NSW case studies of police services and motor registries presented later in this report.

Some of these adjustment methods are examined in more detail in Appendix A.

2.5 Advantages and limitations of DEA

The main advantages of DEA are that:

- it can readily incorporate multiple inputs and outputs and, to calculate technical efficiency, only requires information on output and input quantities (not prices). This makes it particularly suitable for analysing the efficiency of government service providers, especially those providing human services where it is difficult or impossible to assign prices to many of the outputs;
- possible sources of inefficiency can be determined as well as efficiency levels. It provides a means of 'decomposing' economic inefficiency into technical and allocative inefficiency. Furthermore, it also allows technical inefficiency to be decomposed into scale effects, the effects of unwanted inputs which the agency cannot dispose of, and a residual component;

---

[4] The efficiency scores have a truncated distribution between zero and one, so it is necessary to use Tobit rather than ordinary least squares regression techniques. (See the NSW Police Patrols case study in Chapter 5 for an explanation of the regression techniques.)

- by identifying the 'peers' for organisations which are not observed to be efficient, it provides a set of potential role models that an organisation can look to, in the first instance, for ways of improving its operations. This makes DEA a potentially useful tool for benchmarking and change implementation programs. This role is strengthened by DEA's ability to incorporate differences in operating environments beyond management control and, thus, to make more like-with-like comparisons.

However, like any empirical technique, DEA is based on a number of simplifying assumptions that need to be acknowledged when interpreting the results of DEA studies. DEA's main limitations include the following.

- Being a deterministic rather than statistical technique, DEA produces results that are particularly sensitive to measurement error. If one organisation's inputs are understated or its outputs overstated, then that organisation can become an outlier that significantly distorts the shape of the frontier and reduces the efficiency scores of nearby organisations. In regression–based studies, the presence of error terms in the estimation tends to discount the impact of outliers, but in DEA they are given equal weight to that of all other organisations. It is important to screen for potential outliers when assembling the data. One useful check is to scrutinise those organisations whose output-to-input ratios lie more than about two-and-a-half standard deviations from the sample mean. This approach is used in some of the case studies presented later in the report.
- DEA only measures efficiency relative to best practice within the particular sample. Thus, it is not meaningful to compare the scores between two different studies because differences in best practice between the samples are unknown. Similarly, a DEA study that only includes observations from within the state or nation cannot tell us how those observations compare with national or international best practice.
- DEA scores are sensitive to input and output specification and the size of the sample. Increasing the sample size will tend to reduce the average efficiency score, because including more organisations provides greater scope for DEA to find similar comparison partners. Conversely, including too few organisations relative to the number of outputs and inputs can artificially inflate the efficiency scores. Increasing the number of outputs and inputs included without increasing the number of organisations will tend to increase efficiency scores on average. This is because the number of dimensions in which a particular organisation can be relatively unique (and, thus, in which it will not have similar comparison partners) is increased. DEA gives the benefit of the doubt to organisations that do not have similar comparison organisations, so they are considered efficient by

default. There are different rules as to what the minimum number of organisations in the sample should be; one rule is that the number of organisations in the sample should be at least three times greater than the sum of the number of outputs and inputs included in the specification (Nunamaker 1985).

Despite these limitations, data envelopment analysis is a useful tool for examining the efficiency of government service providers. Just as these limitations must be recognised, so must the potential benefits of using DEA (in conjunction with other measures) be explored to increase our understanding of public sector performance and potential ways of improving it.

3 HOW DO YOU CALCULATE DEA?

The focus in this chapter is on key considerations involved in specifying the outputs and inputs attributed to each organisation in a DEA study and its most appropriate coverage. The simple hospitals example of the previous chapter is extended to illustrate how to apply the DEA linear programming formulae. Key results provided by the DEA output and their uses are also discussed.

The simple example presented in the preceding chapter of five hospitals producing one output (treated cases) using two inputs (nurses and beds) served to illustrate the basic concepts behind DEA. It was also sufficiently simple to solve graphically. However, this simple method is inappropriate for more realistic and, consequently, more complex situations. Once the number of inputs and outputs and the number of organisations in the sample increases, linear programming methods are needed to calculate DEA.

This chapter expands the hospital example — twenty hospitals, two outputs and two inputs — to illustrate how these linear programming techniques are used. This sample is still much simpler than most actual studies would be, but remains of manageable size for the purpose.

3.1 Specifying outputs, inputs and coverage

The first step in a DEA study is deciding on its most appropriate scope. Sometimes the study's most important contribution is providing managers with the discipline of having to specify their outputs and inputs and how they can best be measured.

It is essential to include managers with a sound understanding of the process of the organisations being examined from the early stages of model development. The measures of outputs and inputs should be as comprehensive as possible: not including some output dimensions will disadvantage those organisations which are relatively efficient at producing those outputs. Important trade-offs invariably have to be made. Including too many different outputs and inputs — particularly if there are not many organisations in the sample — will tend to inflate the efficiency scores because there is more scope for each organisation to be relatively unique. The organisations might then be considered efficient by default. As a result, the ideal selection includes the smallest number of output and input measures that adequately captures all essential aspects of the organisation's operations.

The process of developing a final model of service production is often iterative, with different combinations of inputs and outputs, and sometimes measures of inputs and outputs, trialed before a final model is reached. This ensures the most appropriate measures, and inputs and outputs, are utilised in the assessment of relative efficiency and also allows the sensitivity of the model to different specifications to be tested.

Outputs

Government agencies deliver a wide range of outputs, and it can be difficult to specify all of them and to account for differences in their quality. However, outputs of service deliverers can generally be classified into those that are proactive and those that are reactive.

- Reactive outputs are often those most readily associated with a particular service — for example, police attending a crime scene, or a hospital providing treatment for admitted patients.
- Proactive outputs are often equally as important in the delivery of the service, but less readily identified and measured — for example, time spent by police gaining the confidence of their community, or a hospital providing an education and immunisation program. Proactive outputs are also related to providing a contingent capability for the community — for example, hospitals providing casualty departments to respond to and cope with unexpected accidents and natural disasters.

Both the reactive and proactive outputs should be taken into account. The quality of the outputs provided, relative to that of other providers in the sample, should also be considered in any efficiency assessment, or managers may be able to increase apparent efficiency at the expense of output quality. This is in addition to the need to assess the effectiveness of the overall service being provided (discussed in Section 1.3).

The quality of reactive outputs and the level and quality of proactive outputs are often reflected in the outcomes achieved by the service overall — for example: the degree to which a community feels safe within a particular area will reflect the quality of police reactions to incidents and crimes, and the degree to which police have gained the community's confidence; the quality of treatment in a hospital can be reflected in the proportion of patients returned to surgery unexpectedly; and the output and quality of an education and immunisation program is likely to be reflected in a reduced incidence of the targeted disease in the community.

Outcome indicators are often associated with the effectiveness of a service, but it is possible to use them in a DEA assessment of efficiency where it is not possible to measure the proactive outputs of service providers directly. They are also useful where there is scope for differences in the quality of outputs, to ensure that quality is not ignored in the efficiency analysis.

In the hospital example being used to illustrate DEA concepts, simply using the number of treated cases (as in the preceding chapter) will not adequately capture the full scope of the hospital's role. It is only a measure of the reactive part of the hospital's contribution to the community without accounting for the proactive side in terms of education, immunisation services and provision of a contingent capability. Concentrating on the reactive output side is unlikely to be adequate. Examples of outputs used in DEA studies which aim to capture the quality of service provision outputs are included in the case studies of hospitals, police patrols and motor registry offices (See Chapter 5).

The functions of hospitals differ markedly, with some providing basic services and others providing more resource-intensive specialist care. In efficiency comparisons, ignoring the fact that some hospitals provide more intensive care for acute cases would disadvantage a small country hospital, for example, which only provides care for non-acute cases and transfers its acute cases to larger metropolitan hospitals. To account for this aspect of hospital operations, the scenario's one output (the total number of treated cases) is changed to two outputs (the number of non-acute cases and the number of acute cases). This will produce more like-with-like comparisons. To keep the example simple, the proactive dimension, which could involve a measure of community health levels, for example, is omitted.

Labour

The desirable measure of labour inputs is that which most accurately reflects the amount of labour used by each organisation. Total hours worked might be the most suitable measure in many cases. However, many organisations do not keep records of hours worked, so the number of full-time equivalent staff is often the best available measure. Both measures are preferable to the simple number of persons employed, which may be misleading if the average number of hours worked per employee varies considerably between the organisations.

However, physical measures of labour input do not capture differences in the quality of labour. This can be addressed by disaggregating the number of hours or full-time equivalents into different types of labour, such as administrative and operational. In the example, the labour input is measured by the number of full-time equivalent nursing staff.

An alternative to using a direct measure of the quantity of labour input is to deflate each organisation's total labour costs by an appropriate labour price. To be accurate, this approach requires a good estimate of the average labour price each organisation faces: for example, an organisation that must pay overtime to employees may have relatively higher labour costs than an organisation that does not.

Capital

Measures of capital input are subject to considerable variation and can be a potential weakness in efficiency measures. There is little consensus on the best means of calculating the price and quantity (and thus cost) of capital inputs in any one period. This is a particularly important issue for those government business enterprises where capital inputs generally account for a large proportion of production inputs. Capital inputs may also be relatively important for many government service providers such as hospitals and schools.

The difficulty in measuring capital inputs is that the entire cost of a durable input purchased in one accounting period cannot be charged against that period's income. Rather, the capital item will provide a flow of services over a number of years. How much of the purchase price should be charged to each period then has to be determined, along with how interest and depreciation costs should be allocated.

There are a variety of methods for calculating the annual cost of capital and the quantity of capital input. The declining balance method is often used in government business enterprise studies, and relies on having an accurate market valuation of the organisation's assets at one point in time (see Salerian and Jomini 1994). However, many government service providers often have little information available on the value of their capital assets. As a result, many government service efficiency studies rely on simple measures of the overall capital used by each organisation. If possible, the capital measures used should provide some insights into the sources of inefficiency that may be associated with the use of capital inputs. This could include purchasing too large a quantity of capital, paying too high a price for capital, purchasing the wrong type of capital, or using an incorrect mix of other inputs with the capital available.

In the hospital example, the number of beds in the hospital was initially a proxy for the hospital's total capital inputs — buildings, land, operating theatres, x-ray equipment and so on. Clearly, this is not a very accurate proxy, but such simple measures are a useful starting point in many government service studies, provided their limitations are recognised. As noted in Chapter 1, using available data to start the process is often the best catalyst for ensuring mechanisms are put in place to systematically collect the data necessary to construct better measures. For the purposes of illustrating how to calculate DEA, the number of beds per hospital is still used to approximate each hospital's overall capital inputs.

Materials and services inputs

Ideally, the analysis should account for all the inputs used by each organisation, just as it should measure all aspects of their output. As well as labour and capital inputs, all organisations use a range of materials and services varying from electricity to pharmaceuticals in hospitals, food in prisons, and electricity to run computers in agency offices. These miscellaneous items are usually aggregated in efficiency studies and deflated by a representative price index. Ideally, the price index should account for differences in the prices faced by each organisation — otherwise, those organisations facing relatively high prices will be disadvantaged because expenditure of a given number of dollars will be translated into a larger input quantity using an average price.

In the hospital example, materials and services costs are not included separately. This is equivalent to assuming that they are used in fixed proportions to the labour and capital inputs.

Coverage

The coverage of a DEA efficiency study depends on the overall aims of the study, the availability of potential comparison partners, and the availability of data. Inevitably, trade-offs have to be made and some degree of pragmatism is always required. If an organisation is sufficiently large it may choose to start with an in-house study measuring the efficiency of different business units performing similar functions — for example, different hospitals within a health department. Alternatively, comparisons could be made at a more aggregate level but this would normally involve including similar organisations in different jurisdictions and/or countries.

Ideally, the more organisations included in the sample the better the explanatory power of the DEA model — there will be fewer organisations found efficient by default. Typically, there will also be more to learn from including a more diverse range of organisations. However, the cost of possibly including too much diversity is that comparisons may no longer be sufficiently like-with-like. This may require adjustment for differences in operating environments to ensure that the study is both fair and credible.

The appropriate scope of a study is usually a matter of what type of organisation is involved. A study being undertaken by an agency itself may concentrate on individual processes in detail, whereas one undertaken by a government monitoring agency may concentrate on overall performance at the aggregate level. In all cases, three things should be kept in mind. First, it is often better to start with the available information, rather than waiting for the perfect data set (although data needs to be reliable for valid conclusions to be drawn). Second, the limitations of the study should always be remembered, and the specification should be refined if necessary. Third, DEA is only one of a number of techniques that can be used in assessing overall performance.

3.2 DEA formula and a simple example

The remainder of this chapter contains illustrations of how to apply DEA to an extended data set, presenting the constant and variable returns to scale cases and calculating scale efficiency scores for each of the twenty hospitals.

There are several different ways to present the linear programming problem for DEA. The formulae for other DEA extensions, including input congestion and allocative efficiency, are shown in Appendix A. In most cases, they involve relatively straightforward modifications to the basic formulae described here.

The simplest general presentation for the version of DEA where assumptions include constant returns to scale, and an objective of minimising inputs for a given level of output (an input-orientated version of DEA), proceeds by solving a sequence of linear programming problems:

(1) Minimise $\mathbf{E}_n$ with respect to $w_1, ...w_N, E_n$

subject to:

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

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

$$w_j \geq 0 \qquad j = 1, \ldots, N$$

where there are N organisations in the sample producing I different outputs (y_{in} denotes the observed amount of output i for organisation n) and using K different inputs (x_{kn} denotes the observed amount of input k for organisation n). The w_j are weights applied across the N organisations. When the nth linear program is solved, these weights allow the most efficient method of producing organisation n's outputs to be determined. The efficiency score for the nth organisation, E_n*, is the smallest number E_n which satisfies the three sets of constraints listed above. For a full set of efficiency scores, this problem has to be solved *N* times — once for each organisation in the sample.

This seems a daunting formula: does it really make any intuitive sense? The less than transparent nature of the DEA formula has contributed to DEA's reputation as being a bit of a 'black box' which people have trouble understanding — and the above formula is one of the simpler ways of presenting it! But it does make intuitive sense once the maths is penetrated.

The above formula is saying that the efficiency score for the nth organisation should be minimised subject to a number of constraints. The factors that can be varied to do this are the weights w_j and the score E_n itself. The weights are used to form the hypothetical organisation lying on the frontier. The constraints are that the weighted average of the other organisations must produce at least as much of each output, as does organisation *n* (the first set of constraints above), while not using any more of any input than does organisation *n* (the second set of constraints above). The third set of constraints simply limits the weights to being either zero or positive.

Relating this back to the simple diagram in Figure 2.2, the process is simply one of looking at all the possible combinations of weights on the other organisations that will produce a point on the frontier such as 2'. The efficiency score is being minimised because it represents the smallest proportion of existing inputs that organisation *n* can use and still produce its existing output if it was using the best practice observed in the sample. It is desirable to be as close to the origin as possible to ensure being on the frontier: that is, both the weights and the efficiency scores are systematically varied to contract each organisation as close to the origin as possible while the contracted point is still a weighted average of some of the other organisations. Thus, point 2 can be contracted as far as point 2': closer to the origin than 2', the point cannot be formed as a weighted average of any of the other points and is not feasible. In the example in Figure 2.2, this gave hospital 2 an efficiency score of 67 per cent. Points 1, 3 and 4 cannot be contracted any closer to the origin while remaining weighted averages of other points, so they achieve efficiency scores of 100 per cent.

Extended hospital data set

How does this apply to the expanded hospitals example (Table 3.1)? The two outputs are the numbers of minor and acute treated cases, while the two inputs remain the numbers of (full-time equivalent) nursing staff and beds.

Table 3.1: Two output, two input hospital data

Hospital number	*Minor cases*	*Acute cases*	*Nurses*	*Beds*
1	150	50	200	600
2	225	75	600	1200
3	90	10	200	200
4	160	40	600	300
5	50	50	500	200
6	75	75	320	150
7	200	50	375	450
8	350	100	400	320
9	400	90	550	480
10	250	300	900	660
11	350	350	850	720
12	350	400	720	940
13	275	375	900	850
14	220	40	250	370
15	300	10	115	250
16	320	275	600	590
17	375	230	550	710
18	230	50	200	280
19	290	90	450	410
20	360	70	415	575

The DEA formula for the first hospital in the two output, two input, twenty hospitals example (data listed above) would be:

(2) Minimise E_1 with respect to $w_1, w_2, \ldots, w_{20}$ and E_1

subject to:

$$150w_1 + 225w_2 + 90w_3 + \ldots + 230w_{18} + 290w_{19} + 360w_{20} - 150 \geq 0$$

$$50w_1 + 75w_2 + 10w_3 + \ldots + 50w_{18} + 90w_{19} + 70w_{20} - 50 \geq 0$$

$$200w_1 + 600w_2 + 200w_3 + \ldots + 200w_{18} + 450w_{19} + 415w_{20} - 200E_1 \leq 0$$

$$600w_1 + 1200w_2 + 200w_3 + \ldots + 280w_{18} + 410w_{19} + 575w_{20} - 600E_1 \leq 0$$

$$w_1 \geq 0,\ w_2 \geq 0,\ w_3 \geq 0,\ \ldots,\ w_{18} \geq 0,\ w_{19} \geq 0,\ w_{20} \geq 0$$

The first constraint requires that the weighted average of the output of minor treated cases, less hospital 1's output of 150 minor treated cases, be greater than or equal to zero. This means that the hypothetical frontier hospital for hospital 1 has to produce at least 150 minor treated cases. Similarly, the second constraint requires that the frontier hospital for hospital 1 produce at least fifty acute treated cases. The third and fourth constraints require the hypothetical hospital to not use any more than hospital 1's 200 nurses and 600 beds, respectively.

Solving this system of equations is not trivial and requires a computer program. A number of specialised and general computer packages can be used to conduct data envelopment analysis (see Appendix B).

The results obtained from solving this DEA problem are presented in Table 3.2. The efficiency scores estimate the extent to which both inputs would need to be reduced in equal proportions to reach the production frontier. In addition, for some hospitals, after both inputs have been reduced in equal proportions, one input could be reduced still further without reducing output (these are referred to as 'slacks' in the DEA literature).[1] The table also contains the peer group for each hospital, the peer weights and the peer count — the number of times this hospital appears in the peer group of other hospitals (excluding itself).

Hospital 1 obtains an efficiency score of 0.63 or 63 per cent (see Table 3.2). That means that it appears that it could be able to reduce its number of nurses and beds by 37 per cent and still produce its 150 minor treated cases and fifty acute treated cases to operate at observed best practice. In practical terms, this means that hospital 1 would have to reduce its number of nurses by 75 to a new total of 125 and its number of beds by 224 to a new total of 376. The peer group and peer weights columns indicate that the best practice for hospital 1 is given by a weighted average of 80 per cent of hospital 15 and 20 per cent of hospital 12. However, as evident from the input slack columns, as well as reducing both nurses and beds by 37 per cent, hospital 1 has an additional 176 beds. That means that to remove all the apparent waste and inefficiency relative to hospitals 15 and 12, hospital 1 would appear to have to reduce its number of beds to a new total of 200.

Overall, six hospitals achieve efficiency scores of 100 per cent. It is evident from the peer count column that all of the apparently efficient hospitals appear in peer groups for other hospitals (and thus, none are efficient by default). However, it is far more likely that hospitals 15, 8, and 16 are truly efficient because they are peers for seven or more other hospitals in the sample. Hospitals 6, 11 and 12 each appear in only two or three peer groups, so there could be scope for them to improve their efficiency further even though they receive efficiency scores of 100 per cent.

---

1 In the example above, the model is run with the assumption that the objective is to minimise inputs for a given level of output. If the model is run with the assumption that the objective is to maximise output then slacks would reflect the amount that an output can be increased, after all outputs have been increased in equal proportions to reach the production frontier (see Figure 2.2).

Table 3.2: Constant returns to scale DEA results for the twenty hospitals

Hospital number	*Efficiency score*	*Labour slacks*	*Beds slacks*	*Peer group*	*Peer weights*	*Peer count*
1	0.63	0	176	15, 12	0.4, 0.1	0
2	0.31	0	76	15, 12	0.5, 0.2	0
3	0.39	22	0	15, 8	0.2, 0.1	0
4	0.48	123	0	15, 8	0.1, 0.4	0
5	0.50	37	0	6	0.7	0
6	1.00	0	0	6	1	2
7	0.46	0	0	8, 15, 16	0.2, 0.4, 0.1	0
8	1.00	0	0	8	1	8
9	0.75	26	0	15, 8	0.3, 0.9	0
10	0.93	0	0	11, 6	0.7, 0.8	0
11	1.00	0	0	11	1	2
12	1.00	0	0	12	1	3
13	0.94	0	0	16, 11	1.0, 0.3	0
14	0.59	0	0	15, 16, 8	0.6, 0.1, 0.1	0
15	1.00	0	0	15	1	11
16	1.00	0	0	16	1	7
17	0.90	0	0	16, 15, 12	0.3, 0.5, 0.4	0
18	0.85	0	0	8, 16, 15	0.1, 0.1, 0.6	0
19	0.71	0	0	8, 15, 16	0.6, 0.2, 0.1	0
20	0.62	0	0	15, 16, 8	0.8, 0.2, 0.2	0

At the other end of the spectrum, with the lowest observed efficiency, hospital 2 appears from the data in Table 3.1 to be grossly over-resourced relative to its output. It has the highest number of beds by far and the fifth equal highest number of nurses but only produces a modest number of minor and acute treated cases. However, it is less obvious from the raw data that the hospital with the second lowest efficiency score — hospital 3 — would be a poor performer because it is considerably smaller. This highlights the advantage of DEA as a systematic way of measuring relative efficiency within the whole sample.

3.3 Introducing scale effects

One simple addition to the DEA formulae above enables the change to variable returns scale. This change relaxes the simplistic assumption that inputs normally will move in exact proportions to the scale of operations: it allows for the existence of economies and diseconomies of scale.

The additional constraint is that the weights in the DEA formula must sum to one. From Figure 2.3 the variable returns frontier is the tight fitting frontier V^AABCD compared with the less restrictive constant returns frontier OBX. Introducing this constraint has the effect of pulling the frontier in to envelop the observations more closely. The variable returns DEA problem for the first hospital in the twenty hospital data set is given by:

(3) Minimise E_1 with respect to $w_1, w_2, \dots, w_{20}$ and E_1

subject to:

$$150w_1 + 225w_2 + 90w_3 + \dots + 230w_{18} + 290w_{19} + 360w_{20} - 150 \geq 0$$

$$50w_1 + 75w_2 + 10w_3 + \dots + 50w_{18} + 90w_{19} + 70w_{20} - 50 \geq 0$$

$$200w_1 + 600w_2 + 200w_3 + \dots + 200w_{18} + 450w_{19} + 415w_{20} - 200E_1 \leq 0$$

$$600w_1 + 1200w_2 + 200w_3 + \dots + 280w_{18} + 410w_{19} + 575w_{20} - 600E_1 \leq 0$$

$$w_1 + w_2 + w_3 + \dots + w_{18} + w_{19} + w_{20} = 1$$

$$w_1 \geq 0, w_2 \geq 0, w_3 \geq 0, \dots, w_{18} \geq 0, w_{19} \geq 0, w_{20} \geq 0$$

The measure of scale efficiency (illustrated in Figure 2.3) can be derived by taking the ratio of the constant returns to the variable returns efficiency scores. If the value of this ratio is one, then the hospital is apparently operating at optimal scale. If the ratio is less than one then the hospital appears to be either too small or too large relative to its optimum size. To determine whether it may be too small or too large requires running a third variant of DEA subject to 'non-increasing' returns. This corresponds to fitting the frontier OBCD in Figure 2.3. By comparing the variable and non-increasing returns scores for those hospitals which appear to be not at optimal scale, it is possible to identify on which part of the frontier they fall. If the variable and non-increasing returns scores are the same then the hospital would be on the segment of the frontier BCD, and thus would be too large relative to its optimum size. If the variable returns score is higher than the non-increasing returns efficiency score, then the hospital is on the segment of the frontier V^AAB, and thus would be too small relative to its optimum size. To calculate the non-increasing returns version of DEA, the constraint in (3) that the weights sum to one is replaced with the constraint that their sum must be less than or equal to one (see Appendix A).

The results for the DEA run with variable returns to scale for the twenty hospitals are presented in Table 3.3. The average size of the efficiency scores is higher in the variable returns case — 87 per cent compared with 75 per cent for constant returns (see Section 4.8.4 for a discussion of the meaning of average efficiency scores). There are now nine hospitals achieving an efficiency score of 100 per cent, although of the three additional efficient hospitals compared with constant returns, one does not appear in any peer counts. This indicates that this hospital — hospital 3 — was found apparently efficient by default because there are no other hospitals of comparable size.

Table 3.3: Variable returns to scale DEA results for the twenty hospitals

Hospital number	*CRTS efficiency*	*VRTS efficiency*	*Scale efficiency*	*Too small/ too big*	*Peer group*	*Peer count*
1	0.63	0.89	0.71	too small	15, 12	0
2	0.31	0.36	0.87	too small	15, 12	0
3	0.39	1.00	0.39	too small	3	0
4	0.48	0.63	0.77	too small	6, 15	0
5	0.50	0.75	0.67	too small	6	0
6	1.00	1.00	1.00	–	6	7
7	0.46	0.56	0.82	too small	6, 12, 15	0
8	1.00	1.00	1.00	–	8	1
9	0.75	1.00	0.75	too big	9	1
10	0.93	0.93	1.00	too big	11, 6	0
11	1.00	1.00	1.00	–	11	2
12	1.00	1.00	1.00	–	12	6
13	0.94	0.98	0.96	too big	12, 11	0
14	0.59	0.72	0.83	too small	15, 12, 6	0
15	1.00	1.00	1.00	–	15	8
16	1.00	1.00	1.00	–	16	1
17	0.90	1.00	0.90	too big	17	1
18	0.85	0.99	0.86	too small	15, 12, 6	0
19	0.71	0.74	0.97	too small	8, 16, 6, 15	0
20	0.62	0.93	0.67	too big	17, 15, 9	0

The average scale efficiency score is 86 per cent. The hospitals that are not of optimal size comprise nine that appear to be too small and five that seem too big. There are some apparent anomalies in this — for instance, hospital 2, which was identified as being the worst performer as a result of its inadequate output for a relatively large amount of inputs, is still the least efficient under variable returns but the results suggest that it is too small rather than too big. Clearly, apparent anomalies such as this would have to be followed up with more detailed analysis in an actual study.

3.4 Conclusion

This discussion has covered some of the main issues to consider before undertaking a DEA efficiency study, and an example of how to calculate DEA for a group of twenty hospitals. A more technical description of DEA and various extensions is presented in Appendix A. In Appendix B, the computer programs to calculate DEA information such as that presented in this chapter are outlined.

Chapter 4 contains an overview of case studies where DEA has been used to assess the relative efficiency of a range of human services. The case studies are presented in detail in Chapter 5.

To summarise the main issues to consider and anticipate before undertaking a DEA study, the following questions based on Fried and Lovell (1994) are worth asking:

- What should the unit of observation be — the aggregate organisation or business units within the organisation?
- What are the organisation's main outputs and inputs?
- What characteristics of the operating environment are relevant?
- What should the comparison set be — within the city, within the state, national or international?
- What time period should the study take?
- Are all outputs and inputs under management control?
- What do you tell the managers of an apparently less efficient organisation?
- What would you say if you were the manager of an apparently less efficient organisation?
- What should you do with an organisation that is apparently less efficient because it is too small or too large?

4 OVERVIEW OF THE CASE STUDIES

In this chapter specific reference is made to case studies in which data envelopment analysis was applied:

- *acute care services in Victorian hospitals;*
- *Queensland oral health services for school students;*
- *NSW correctional centres;*
- *NSW police patrols; and*
- *NSW Roads and Traffic Authority (RTA) motor registries.*

4.1 Introduction

The models used to assess efficiency are outlined below, along with practical issues that were encountered in applying DEA. The following points should be kept in mind when examining the case studies:

- the case studies are work in progress, with the ways in which the models could be improved highlighted where appropriate;
- it is not possible to compare efficiency scores across case studies — each is specific to the sample of service providers included in the study;
- the issues raised in this section are not comprehensive. The case studies (presented in full in Chapter 5) contain more detail on preparing a DEA study and interpreting results; and
- while the case studies presented in this report are based on organisations for which State governments are responsible, it would be equally appropriate to use DEA to assess efficiency at other levels of government and, where data were available and comparable, across jurisdictions.

4.2 Acute care services in hospitals in Victoria

4.2.1 DEA model

The study incorporated 109 hospitals in Victoria for 1994-95. Given differences in input data availability and expected differences in operating structures, the sample was split into metropolitan/large country hospitals (including teaching, research and base hospitals) and small rural hospitals (excluding base hospitals).

An output orientation was used to reflect the objective of hospitals to provide the highest level of care with given resources. The DEA model included the following inputs and outputs.

Inputs

- Labour, disaggregated into the number of full-time equivalent medical and non-medical officers.
- Consumables, such as pharmaceuticals and electricity, measured by recurrent non-labour expenditure.

Outputs

- Number of patients treated by each hospital expressed in terms of weighted inlier equivalent separations (WIES). This measured the number of separations (when a patient leaves the hospital) weighted by the expected resources required to treat each case. These were aggregated into three categories based on the degree of complexity of each WIES.
- Unplanned re-admissions rate (an imperfect proxy for the quality of care).[1]

4.2.2 Some results and issues for consideration

The study suggested that the relatively less efficient metropolitan/large rural hospitals may be able to increase their outputs by an average 11 per cent while holding inputs constant, with size generally having little apparent influence on efficiency.

Those small rural hospitals which appeared relatively less efficient could potentially increase all their outputs by an average 33 per cent, using the same level of inputs. These hospitals, on average, could possibly increase their outputs by a further 29 per cent if they were producing services at the optimal size in the sample.

Overall, the results suggested there was probably more scope for improvement to best practice in the sample of small rural hospitals than in the sample of metropolitan/large country hospitals. Closer analysis showed that there was a greater range of performance in small rural hospitals, and that scale efficiency was an important determinant of technical efficiency. This is likely to be a result of rural hospitals facing overall lower demand (because they have fewer clients within their catchment areas than metropolitan hospitals) yet still having to maintain a level of ‘readiness’ to meet potential demand as it arises. This type of

[1] The inverse of the unplanned re-admission rate was used to reflect fewer unplanned re-admissions being a preferable output to higher unplanned re-admissions.

information could be used to consider the appropriate sizes of rural hospitals in the context of the broader objective of providing equitable access to hospital services.

In the metropolitan/large country sample, a high proportion of hospitals were found to have the same observed efficiency, reducing the overall explanatory power of the model. This could have been as a result of the omission of capital from the model which would bias the results towards hospitals with relatively high capital usage. These hospitals were more likely to be able to produce higher outputs with lower levels of the measured inputs such as staff and raw materials. The model could be improved by including some measure of the capital input of hospitals.

4.3 Queensland oral health services for school students

4.3.1 DEA model

The study covered the provision of oral health services to school services in thirteen regions in Queensland. The two smallest regions were excluded because they were deemed not to be as comparable as the other regions. The sample size was expanded to thirty-three by including data for each of the eleven remaining regions for three years. An input orientation was considered the most appropriate by Queensland Health, including the following inputs and outputs.

Inputs

- Labour, measured by the number of days worked, disaggregated into dental officers, dental therapists and dental assistants.
- Other consumables, measured by non-salary recurrent expenditure.

Outputs

- The number of general completed treatments.

4.3.2 Some results and issues for consideration

The study found that the apparent efficiencies of the oral health care units were relatively similar. Most units achieved efficiency scores of greater than 80 per cent — that is, they may be able to reduce inputs by up to 20 per cent while maintaining the same number of completed treatments if they operate at what appears to be best practice.

The performance of individual units — whether apparently good, average or poor performers — tended to be consistent over the three–year period. However, performance of one region deteriorated from being efficient in 1992-93 to having the lowest efficiency score (70 per cent) in 1994-95. This reflected a significant decline in the number of treatments provided, combined with a relatively large increase in non-labour expenditure over the three–year period.

Further examination of why some regions appeared to perform consistently better or worse than others would be useful. Factors that could be considered are whether an important output for some regions had been excluded from the study, or whether environments differed over the time period of the study. Consistently good performers could be examined to identify the types of management practices that were more likely to lead to efficiency in providing those services.

4.4 Correctional centres in NSW

4.4.1 DEA model

There are significant differences in the resources used to run maximum and minimum security centres. Therefore the study was limited to minimum security correction centres. There were only eleven similar centres in NSW, and data for each organisation for up to the past five years was included in the study, increasing the sample size to an acceptable level. This approach was valid because the NSW Department of Corrective Services advised that there was minimal change in the management of inmates over this period.

The model was input oriented, with efficiency relative to best practice measured in terms of how inputs could be reduced without a reduction in outputs. The DEA model included the following inputs and outputs.

Inputs

- Labour — the number of full-time equivalent custodial and other correctional officers.
- Capital — the number of beds.
- Other inputs such as food, clothing and other consumable goods and services, measured by recurrent expenditure on these goods and services.

Outputs

- The number of inmates, disaggregated into those eligible for conditional leave of absence and other inmates, because management of the latter was more resource intensive.
- The number of inmate receptions in each correctional centre (a measure of the turnover of inmates — a resource intensive activity unevenly distributed over the centres).
- The number of hours spent by inmates in personal development programs (to reflect the level of these services provided to inmates).

4.4.2 Some results and issues for consideration

It was found that the correctional centres in the sample, on average, may be able to reduce their inputs by about 4 per cent without reducing outputs if they could all operate at what appeared to be best practice. If the correctional centres could achieve optimal scale, then they may be able to reduce inputs by a further 4 per cent.

This study had a relatively high proportion of correction centres that are defined as efficient by default (about 20 per cent of the managerially efficient correction centres) because it had a relatively small sample compared with the number of outputs and inputs used in the analysis. To overcome this problem, further analysis could include correction centres from other states to increase the sample size. Alternatively the number of inputs and outputs in the analysis could be reduced.

The study identified one correctional centre as having a marked reduction in apparent efficiency over the time period — from appearing to be efficient it became the apparently least efficient in the sample. Further investigation found that the centre had been converted from a male to a female facility in the year in which it was found relatively less efficient, with inmate numbers declining by around 40 per cent without a commensurate reduction in inputs.

4.5 Police patrols in NSW

4.5.1 DEA model

The study covered 163 police patrols in NSW. A patrol could include one or several police stations within a specific geographic area. An input oriented model was used to reflect the objective of police patrols to provide effective policing with minimum inputs. The DEA model included the following inputs and outputs.

Inputs

- Labour — the number of staff disaggregated into police officers and civilian employees.
- Capital — the number of police cars in each patrol.

Outputs

- Number of arrests.
- Responses to major car accidents.
- Responses to incidents measured by recorded offences.
- Number of summons served.
- The number of kilometres travelled by police cars.

The first four outputs refer to the reactive aspects of policing. The last output — kilometres travelled by police cars — covers some of the proactive, or preventative, aspects of policing. (A visible police car can reassure the community and prevent crime.)

Environmental factors

Factors identified which may affect the apparent efficiency of a patrol but which were beyond the control of management were:

- the proportion of people aged 15 – 19 years within a patrol's jurisdiction;
- the proportion of public housing in a patrol's jurisdiction; and
- whether a patrol was a country or metropolitan patrol.

Given the above inputs and outputs, patrols with higher proportions of young people and public housing were expected to appear to be relatively more efficient, because they were likely to respond to more crime and have less idle time. Country patrols, with larger, less populated areas, were expected to appear relatively less efficient compared with metropolitan patrols because they required more inputs to provide a similar service.

4.5.2 Some results and issues for consideration

Police patrols, on average, might be able to reduce their inputs by 13.5 per cent while maintaining their output levels and operating size. Scale inefficiency did not appear to be a major source of input reduction. However, if it were possible to restructure patrols to achieve their optimal size there may be further input savings, on average, of 6 per cent. The measured efficiency of police patrols did not appear to be influenced by the environmental variables using this model.

It is not clear how the quality of police work influences the level of the outputs included in the model. Crime prevention is a major output of police patrols but is difficult to measure. It is conceivable that a patrol identified as efficient by DEA, because it had a high number of crime related activities relative to its inputs, was ineffective in crime prevention. Further work is required to improve the measurement of proactive policing to fully capture this aspect of police work in efficiency measurement.

4.6 Roads and Traffic Authority motor registry offices in NSW

4.6.1 DEA model

The study covered 131 motor registry offices in NSW. An input orientation was used, given that registry managers could not control the demand for services, and thus the level of outputs. Their objective was therefore to meet the given demand with the least resources. The DEA model included the following inputs and outputs.

Inputs

- Labour, measured by the total number of hours worked by all staff.
- Capital, measured by annual expenditure on computers (a key capital input for registry offices). The Roads and Traffic Authority considered that this measure reflected the number of computers in each office because most computer equipment was acquired at the same time and expenditure for that year was used.
- Other consumables, such as licences, plates and postage, measured by expenditure.

Outputs

- Number of transactions performed in each office, weighted by the average time taken for each type of transaction.
- Average waiting time for customers, which was the relevant measurable variable reflecting the quality of service received by customers in registry offices.[2]

[2] The reciprocal of waiting time was used to reflect that a shorter waiting time was a preferred output to a longer waiting time.

Environmental factors

Two factors which were considered to be outside the control of registry office managers but which could influence the relative efficiency of each registry were whether:

- it was open for Saturday trading; and
- it processed data for motor registry agents which did not have access to computer services.

The presence of either factor was expected to increase the relative efficiency of offices, because they were likely to allow relatively more transactions to take place with the same level of staff.

4.6.2 Some results and issues for consideration

The results suggest that motor registries may be able to produce the same level of measured output with 15 per cent fewer inputs relative to best practice. Size of registry offices appears to have only a minor impact on efficiency.

Both environmental factors were found to have a positive impact on measured relative efficiency. However, these effects were not found to be significant, so the efficiency scores for motor registries were not adjusted.

Future studies of RTA registry offices will use computer terminal numbers, rather than expenditure on computers, as a proxy for capital input.

4.7 General observations

4.7.1 Coverage

The organisational unit used in all of the case studies was the unit from which services are actually delivered. At this level of decision making:

- managers are generally responsible for how inputs are used to produce outputs;
- the organisations being assessed generally have access to similar types of resources and are expected to complete similar tasks; and
- there are generally enough organisations within a jurisdiction to allow comparisons to be made (where this was not the case, time series data were included to increase the sample).

4.7.2 Inputs

Labour is most often measured by full-time equivalent staff, and raw materials are most often measured by recurrent expenditure on goods and services. However, it was consistently difficult to identify an appropriate and accurate measure for capital. Most often, a proxy was used as the only available comparable data. Limitations in assessing the capital input to the service provision process need to be considered when assessing results. Improving data bases on the significant levels of capital utilised in the provision of human services is necessary for improvements in the assessment of overall performance in these areas.

4.7.3 Outputs

Careful consideration needs to be given to measuring and including the proactive or preventative outputs of organisations in the analysis. Examples include the crime prevention activities of police and the public health programs of hospitals. Where these outputs are not included in the model, service providers that are highly proactive will be penalised in the efficiency assessment if these activities are effective in reducing the need to provide reactive services. Indicators of effectiveness, such as those reported by the Steering Committee in the *Report on Government Service Provision 1997,* need to be considered in conjunction with an assessment of technical efficiency.

The quality of the outputs being measured should be considered. This is often very difficult, but is necessary to ensure that higher measured efficiency has not been achieved by providing services at a lower quality than previously provided.

4.7.4 General issues to be considered in interpreting and presenting results

The simple average of efficiency scores across a sample may not necessarily indicate the potential for overall efficiency improvement. A less efficient organisation which is a large user of inputs, for example, has a greater potential (if it were to become efficient) to reduce the total inputs used across the whole sample, than does a smaller user of inputs with the same efficiency score — even though both will have the same effect on the average score. If average scores are linked to the magnitude of potential reductions in inputs or expansions in outputs overall, efficiency scores need to be weighted by the inputs or outputs in question.[3]

In addition, the efficiency scores overstate the relative efficiencies of organisations in a DEA study. The efficiency scores represent the extent to which all inputs can be reduced proportionately to reach the production frontier. But some organisations may be able to reduce some inputs even further, without reducing output.[4] It is important to consider reductions in the use of these inputs when assessing both the sample and individual organisations within the sample.

Generally, the omission of any particular input favours those organisations which use above average amounts of that input. Likewise, organisations for which a high proportion of their output is not measured will appear to be less efficient. In presenting DEA results, it is important to place any efficiency assessment in the context of the overall objectives of the organisations being assessed. There may be a relatively high level of inputs compared with outputs in some service outlets because, for instance:

- a given level of inputs is required to provide a service which is used by relatively fewer people; and
- a given level of inputs is required to ensure potential demand can be met, but this level of ready capacity is used relatively less frequently than in other areas.

These situations are most likely to occur where the catchment area of clients for the service provider is not highly populated, such as in rural and remote areas. Organisations may be technically efficient for their size, but it may not be possible for them to achieve the economies of scale within their catchment areas that are available in more populated areas. Thus, it is important to assess the importance of scale efficiency on the technical efficiency of organisations in the sample. These issues become less important if organisations are more alike, and can be accounted for to some degree by using environmental variables such as location or population density.

---

3 The average efficiency score with variable returns to scale presented in Section 3.3 for the hypothetical twenty hospitals is 87 per cent. This implies a potential reduction in beds and nurses, on average, of 13 percent across all hospitals. However, the efficiency scores are not evenly distributed across hospitals of different sizes. After taking into account the distribution of efficiency scores across hospitals of different sizes based on beds, for example, the sum of the weighted efficiency scores ([beds in hospital X/total beds] * efficiency score) indicates that the total number of beds across the sample could be reduced by 15 per cent, rather than 13 per cent.

4 These inputs are described in the DEA literature as slacks (see Figure 2.2 and the Glossary).

Finally, as the case studies illustrate, the available data for service providers' inputs and activities are often not fully consistent or comprehensive. In order to improve the data bases for service providers, there is a need to document any data deficiencies so that these may be addressed for future assessments of performance.

DEA results provide the maximum benefit when they are interpreted with care. In general, they should be considered as a starting point for assessing the efficiency of the service providers within a sample. Indications of possible sources of relative inefficiency can guide further investigation to determine why there are apparent differences in performance. This information can be used to inform the managers of individual service providers, administrators and policy makers.

5 CASE STUDIES

5.1 Technical efficiency in the hospitals of Victoria[1]

5.1.1 Summary

This report details the results of a study of the technical efficiency of a sample of acute care public hospitals in Victoria. The study uses DEA to explore relative efficiency of all hospitals in the sample.

The objectives of this study were to demonstrate the potential for using DEA as a benchmarking tool for measuring the performance of acute services in Victorian public hospitals.

Annual data for 1994-95 was provided by the Victorian Department of Human Services on 109 hospitals, including teaching hospitals. The inputs and outputs used are set out in Table 5.1.1.

Table 5.1.1: Preferred model specification

Inputs	*Outputs*
Full-time equivalent non-medical staff	WIES with intensity rate < 0.2 (Y1)
Full-time equivalent medical staff	WIES with intensity rate ≥ 0.2 and < 0.4 (Y2)
Non-salary costs	WIES with intensity rate ≥ 0.4 (Y3)
	Inverse of the unplanned re-admission rate

A weighted inlier equivalent separation (WIES) measures the number of separations of a given complexity. A WIES is similar, but not equivalent, to a diagnostic related group separation (DRGS). It measures different acute care cases by their degree of expected resource intensity, ranging from minor treatments (Y1) through to complex cases (Y3). For example, Y1 equals the total number of WIES figures for episodes of care which required 0 – 0.2 WIES3 per day during each episode. (WIES3 refers to the way in which WIES

1 Researched and written by Tendai Gregan and Rob Bruce of the Industry Commission. Comments from Dr Graeme Woodbrigade, Paul D’Arcy, Professor Knox Lovell and Dr Suthathip Yaisawarng are gratefully acknowledged. However any errors or omissions are the responsibility of the authors.

are measured in 1996). Y2 and Y3 are similarly defined, with the intensity rates given in the above Table.

The unplanned re-admission rate was included to account for the objective of hospitals to maintain acceptable standards of quality of care while seeking efficiency improvements. Unplanned re-admission rates are a proxy for the quality of care in a hospital, but are not an ideal measure. Future studies should seek to incorporate more accurate measures of the quality of care in hospitals.

The model was run using an output maximisation orientation. Initially, it was run using the full sample under the assumption of constant returns to scale. Relaxing this assumption produced a variable returns to scale model which allowed the issue of scale inefficiency to be examined. Given differences in data available at hospital level for inputs, and expected differences in operating structures, the sample was split in two: metropolitan/large country hospitals (including teaching, research and base); and small rural hospitals (excluding base hospitals). Constant and variable returns to scale model runs were then conducted for each sub-sample.

Detailed results for each model are included in Annexes A5.1.1–A5.1.5. These results include information on: technical efficiency scores; the extent and nature of scale efficiency scores; as well as actual and target values for inputs and outputs.

In summary, the difference for metropolitan/large country hospitals between the most and least efficient seems small. Twenty–four out of thirty–seven hospitals made up the efficient frontier. The average relative efficiency score for hospitals not on the frontier was 1.11, with the average hospital potentially able to increase its outputs by 11 per cent, holding all inputs constant. In addition, after increasing all outputs by 11 per cent, some large hospitals may still be able to increase one or more output by up to 25.3 per cent. Scale efficiency of 1.05 for metropolitan/large country hospitals indicates, on average, that size appears to have little influence on efficiency.

For small rural hospitals, the results suggest that the dispersion between efficient and less efficient hospitals may be wide. Fourteen out of sixty–nine hospitals made up the efficient frontier. Small rural hospitals which were not on the frontier had an average efficiency score of 1.33, and appear to be able to increase all their outputs by 33 per cent, using the same level of inputs. In addition, there appeared scope for some hospitals to increase between one and three outputs by between 4.4. per cent and 26.8 per cent. Scale efficiency of 1.29 for small rural hospitals indicates, on average, that size may have had some influence on efficiency.

The models used were developed in consultation with the Victorian Department of Human Services. Advice was sought on hospital inputs, outputs and indicators of quality of service. Initially, the sample included all hospitals, but Department input on the relevance of some peers and the relative efficiency scores indicated that there were some problems in the input data across the whole sample. The input of the Department led to the splitting of the sample, which was also supported by expected differences in the operating structures of metropolitan/large country and small rural hospitals. The subsequent models of metropolitan/large country hospitals and small rural hospitals were validated by the Department as providing a plausible analysis of the relative efficiency of Victorian hospitals.

The sensitivity of the two models was tested by changing the measure of labour inputs from full-time equivalent staff to salary costs. The efficiency scores and the hospitals appearing on the frontier varied little when this was done, indicating that staff costs appeared to be reasonably consistent within each of the sub-samples. These tests support the hypothesis that the model specifications used are a reasonable representation of the production technology used by large and small Victorian hospitals.

5.1.2 Background

DEA has been used to analyse the relative efficiency of hospitals in NSW (NSW Treasury 1994), and the United States (Banker, Das and Datar 1989, Burgess and Wilson 1993, Valdmanis 1992), among others. For an extended bibliography of DEA health studies, see Seiford (1994).

This study was conducted by the Industry Commission in consultation with the Victorian Department of Human Services. The Department is responsible for the funding, monitoring and evaluation of the State's hospitals. The Department was interested in investigating whether DEA could be used as a tool for benchmarking relative hospital efficiency. This study includes information on casemix (the WIES data) because it provides rich information on different types of hospital outputs and facilitates like-with-like comparisons.

A single year's data was used to test the feasibility of DEA as a management tool for measuring hospital efficiency. Discussions held between the Industry Commission and Department officers allowed the Commission to learn about the operations of Victorian hospitals, develop an appropriate model specification, and interpret the results.[2]

5.1.3 Data

Table 5.1.2 shows the types of data used to construct the DEA models (actual data in Annexes 5.1.1–5.1.5). The data were supplied by hospitals in returns to the Department and a casemix data-base, comprising information for 1994-95 on:

- the resources used to provide inpatient acute care services;
- the percentage of all such cases which result in the unplanned re-admission of the patient; and
- the number of inpatient acute care services, grouped by case severity and length of treatment.

Detailed definitions of each item are given below.

Table 5.1.2: Victorian hospitals data, 1994-95

Inputs	*Units*
X1: Full-time equivalent non-medical staff (metropolitan & large country hospitals only)	Number
X2: Full-time equivalent medical staff (metropolitan/large country hospitals only)	Number
X3: Total full-time equivalent staff	Number
X4: Non-salary costs	$’000
X5: Non-medical salaries (metropolitan/large country hospitals only)	$’000
X6: Medical salaries (metropolitan/large country hospitals only)	$’000
X7: Total salaries	$’000
Outputs	*Units*
Y1: WIES with intensity rate < 0.2 (minor)	Number
Y2: WIES with intensity rate ⇩ 0.2 and < 0.4 (moderate)	Number
Y3: WIES with intensity rate ⇩ 0.4 (complex)	Number
Y4: Unplanned re-admission rate	Percent

The study focused on hospital inpatient acute care services, which make up the majority of total hospital services. (Over the sample, an average of 82 per cent

[2] The Commission sincerely appreciates the support given to the project by the Department, in particular Ms Fatima Lay, Mr Tony Simonetti and Mr John Iliadis of the Acute Health Care Division.

of inputs were devoted to acute inpatient care.) The Department considered that the measures for non-acute and outpatient services, such as bed days, were not of a level that explained the output of hospitals as well as those used for acute inpatient services, which account for case severity and length of stay.

For each hospital, estimates of the inputs used to provide acute inpatient services were derived by multiplying each of the total inputs by the share of acute inpatient services in total hospital costs.[3]

DEA is sensitive to outliers, which are observations that are not typical of the rest of the data.[4] Outliers can arise from either measurement or reporting error, or may reflect significant efficiencies being achieved by particular hospitals. Alternatively, outliers may identify hospitals which use different production technologies. In the first case, outliers should be removed from the data, and in the latter instances, hospitals should be checked to determine whether they have access to an omitted input or use different technology. All the inputs and outputs in the full sample of 109 hospitals were screened for potential outliers using the technique discussed in Section 2.5. The potential outliers were referred to the Department, who advised that three hospitals had measurement errors. These three were removed to form the sample of 106 hospitals used in the model runs. The remaining potential outliers were judged to be free of measurement or recording errors, and to be comparable to the rest of the set, and were retained in the sample.

Inputs

Valdmanis (1992) and Burgess and Wilson (1993) used physical inputs, such as the number of full–time equivalent staff by skill category; the number of beds as a proxy for capital; the number of admissions; and the number of visits by

---

3 An initial analysis was carried out excluding information on non-acute hospital outputs, but including the inputs used to provide these services. This led to biased results. It was found that hospitals which provided relatively more non-acute services — as indicated by the share of non-acute services in the total budget — appeared to be relatively inefficient compared with hospitals which concentrated on providing acute care services. When inputs used to provide non-acute services were excluded by estimating the quantities of inputs used for acute services only, it was found that the efficiency scores improved for hospitals that provided relatively more non-acute services. If these estimates still contain some inputs used to provide non-acute services, then it can be expected that there will be a degree of bias against hospitals which provide relatively more non-acute services. The extent of this bias will depend on the size of estimation error. However, it was judged that any error — and thus bias — would be small, given the accuracy of the budget share data used to split acute and non-acute services.

4 See Section 2.5 for a discussion of the impact of outliers on DEA results.

physicians. In contrast, Banker, Das and Datar (1989) used cost information, broken down by labour type and non-labour resources, to measure inputs. Although physical measures are preferred to cost measures because DEA measures physical productivity, this study used both types to test whether there was a significant difference in the results.

Full-time equivalent staff

Full-time equivalent staff were used to measure labour input. Given the shift work nature of hospitals and the prevalence of part-time employment, data on the number of full-time equivalent staff gives a more accurate indication of the amount of labour used to provide services than does a simple count of the number of staff employed or the cost of labour.

The cost of labour is an alternative measure of the resources used, but the staff measure used was preferable for two reasons.

First, the Department of Human Services advised that salary expenditure per full-time equivalent staff member could be expected to vary significantly between city and rural hospitals. Where this was the case, differences in measured expenditure would reflect the prevailing regional wage rates, the level of training of staff, and the physical quantities of labour used to perform any given service. The staff measure was likely to be more homogeneous across regions than was expenditure because it was not influenced by wage rates. Greater homogeneity allows for better comparisons of the actual physical product than does cost measures.[5]

Second, for hospitals aiming to minimise costs, they had to employ the least physical quantity of each input to produce a given level of output.

The labour data was split into two classes:

1. non-medical full-time equivalent staff, directly employed by hospitals (that is, nurses, nurse assistants, cleaners, management and administration staff); and

[5] For example, if a city and country hospital both use one doctor hour to treat a patient for a broken leg, then the measure of both their physical products would be 1 (equal 1 broken leg treatment / 1 doctor hour). However, if cost data rather than quantity data is used and doctor's wages are lower in the city than in the country, then the 'productivity' of the country doctor would mistakenly appear to be lower. If the hourly wage in the city is \$45 and the country wage is \$50, then the city hospital's 'productivity', (1/[\$45×1]), 0.022, is greater than that in the country (1/[\$50×1]), 0.020. In fact, both hospitals are equally efficient in their provision of services, but the relatively higher costs in the country may reflect, among other things, a less competitive market for labour and thus higher wages.

2. medical full-time equivalent staff (doctors, specialists) directly employed by hospitals.

This broadly reflected the different skills and functions of labour used in hospitals.

The choice of non-medical and medical full-time equivalent staff was based on the traditional division of labour used in hospitals: nurses and doctors. The vast majority of 'non-medical' staff are nurses, who provide general care to patients, usually under the direction of doctors. The remainder of staff in this category provide general hospital support and administration services. 'Medical' full-time equivalent staff comprise interns (trainee doctors), doctors, surgeons and specialists directly employed as officers of the hospital.

The medical category excluded persons providing medical services to the hospital on a fee-for-service basis, who are referred to as visiting medical officers. The input of visiting medical officers, and possibly some non-medical staff, was captured in the contract fees paid to them, which were included in the non-salary costs (X4) of each hospital. Ideally, these should have been captured in a full-time equivalent measure, but such information was not available.

Salary costs

Financial information on the costs of labour was also provided. Labour costs were divided into the same categories as staff: non-medical staff salaries and medical staff salaries.

Good information on these categories was available for metropolitan/large country hospitals, but was patchy for small rural hospitals because these do not typically employ medical staff directly. Given that they use visiting medical officers, rather than salaried doctors, the data on medical full-time equivalents and the corresponding medical costs were zero. Accordingly, two separate models of small rural hospitals were used: one using total full-time equivalent and the other using total salaries. A pooled sample of all hospitals, large and small, also used total full-time equivalents as the labour input measure. However, for the reasons set out in Section 5, this sample was split into metropolitan/large country hospitals and small rural hospitals.

Non-salary costs

Inputs other than labour are important for providing acute care hospital services. These were captured in non-salary costs, which accounted for the remaining inputs — other than capital — used in the production of hospital services.

Non-salary costs = maintenance + contract costs (visiting medical officers) + electricity + gas + water + consumables (bandages, drugs, etc.) + superannuation

Fixed capital is a significant input in providing care. A measure of capital was not included in the model because comprehensive and accurate information on the stock of capital assets was not available. This has some implications, which are examined below, for interpreting the model.

Outputs

Other studies (Burgess and Wilson 1993; Valdmanis 1992) have measured outputs using the number of inpatient hospital bed days, the number of surgeries, and the number of emergency treatments. Like Banker, Das and Datar (1989), this study used casemix data. However, this study differs in that the data were adjusted for length of stay. Time adjusted casemix data was preferable to bed days because first, it is more homogeneous across hospitals, and second, it captures casemix adjusted for severity of illness and the expected resources required to treat patients.

Weighted Inlier Equivalent Separations (WIES)

WIES is a measure of case intensity (diagnostic related group) adjusted by the normalised patient length of stay (inlier equivalent separations, or IES). Formally:

$WIES = IES \times DRG\ weight$

where:

- each DRG represents a class of patients with similar clinical conditions requiring similar hospital services. A more detailed explanation of DRGs is given in the National Health Data Dictionary (NHDC 1995) and Eagar and Hindle (1994);
- *DRG weights* are an index of case complexity based on clinical history — for example, a leg fracture has a lower DRG weight than a liver transplant; and
- IES represent the 'normal' length of time for which a patient will stay in hospital, for every type of DRG.[6] A case which is in this 'normal' interval

[6] The 'normal' length of stay is given by the DRG average length of stay, which is based on historical records and current medical practice. The low boundary point (LBP) is set to one third of the DRG average length of stay and the high boundary point (HBP) is set at three times the DRG average length of stay. Values of low and high boundary points are

is called an inlier and given an IES equal to one. Cases which take less time are weighted lower and those which take more time are weighted higher.[7]

The WIES used in this study include all acute care services to inpatients: acute care; palliative care; and alcohol and drug treatment. They exclude services for: nursing home care; aged care; psychiatric and certain types of rehabilitation. (See Appendix 7 of HCS 1994.)

There are over 500 DRGs and thus WIES. To apply DEA, these WIES groups were aggregated into three categories which reflect the different casemixes handled by different types of hospitals:

- WIES with an intensity rate less than 0.2 (minor)
- WIES with an intensity rate greater than or equal to 0.2 and less than 0.4 (moderate); and
- WIES with an intensity rate greater than or equal to 0.4 (complex)

Despite the advantages of using WIES figures over traditional variables, the casemix classification system is not perfect. The casemix formulations have been upgraded continually since inception to make them as comprehensive and accurate as possible. To the extent that not all acute care activities may be captured by the WIES figures, the DEA results presented in this report should be interpreted with caution.[8]

Unplanned re-admission rate

The quality of hospital outputs is a defining characteristic of the care provided. It is difficult to measure the quality of care. The typical surrogate measures include mortality rates, re-infection rates and unplanned re-admission rates.

---

rounded to whole numbers. In addition, the maximum value of a high boundary point is limited to 100 days.

7 Specifically, a case which is less than the low boundary point of the 'normal' length of stay is given an IES equal to the actual length of stay divided by the low boundary point. Similarly, a case which is greater than the high boundary point of the 'normal' length of stay is given an IES equal to one plus the fraction given by the number of days above the one high boundary point divided by two times the DRG Average Length of Stay.

8 A model with a single output variable, total WIES, was tested and found to be unsatisfactory because it yielded inappropriate benchmarking partners. For example, it gave small rural hospitals which treat simple cases mainly as peers for large teaching hospitals treating much more complicated cases. The preferred model has WIES separated into three classes of casemix. It was judged that the increased number of outputs gave a more plausible mix of peers, and did not unduly inflate either the efficiency scores of hospitals or the number of hospitals that were efficient by default.

None of these data were readily available. However, it was found possible to construct re-admission rates for the hospitals using the Casemix database. This was done by assuming that an unexpected return to the same hospital within twenty-eight days of the patient previous episode of care (which may or may not be related to the first episode of care) was a re-admission. The lower the number of these re-admissions, the higher the quality of care arguably.

Two criticisms of using the unplanned re-admission rate as a proxy for quality are that:

- the method used to calculate the rates tends to overstate the actual rates, because many re-admissions may be clinically unrelated to the first episode of care; and
- hospitals with a more complex casemix have a higher probability of unplanned re-admissions, biasing the results against these hospitals.

However, when using DEA to measure relative efficiency, hospitals are compared only with those hospitals which produce a similar mix of outputs, given input levels, ensuring that those with higher levels of complex cases and unplanned re-admissions are compared with each other only.

This variable was included in the model in recognition of the fact that hospitals, in seeking improvements in efficiency, wish to maintain or improve standards of service. The unplanned re-admission rate has been regularly used as a quality indicator since the introduction of casemix funding in 1993.

Unplanned re-admission rates have been used as an indicator of hospital effectiveness (SCRCSSP 1995), but this study used the rates in the measurement of hospital efficiency. The assumption of the study was that an increase in output using the same quantity of inputs, and at least maintaining the same quality standards, was a true increase in efficiency, whereas the same increase in output with a fall in quality might not have meant that there had been an efficiency increase necessarily. This is because quality is a defining characteristic of any output — it is easier and less resource intensive to produce low quality rather than high quality output. Therefore, ignoring the quality dimension results in a flawed view of any measured efficiency increases. Nevertheless, care is required in interpreting these results.

This case study measured efficiency in terms of hospitals' ability to increase outputs using the same quantity of inputs, that is, the model was *output oriented*.

Because the unplanned re-admission rate is a 'negative' output (that is, an increase is undesirable), the inverse was used in the analysis. Maximising the inverse of the unplanned re-admission rate is the same as minimising the unplanned re-admission rate.

5.1.4 Model specification and orientation

Five models were run, each with a different sample and input variables, but the same four outputs (Y1 – Y4).

1. All metropolitan, large country and small rural hospitals, with inputs: total full-time equivalent staff (X3) and non-salary costs (X4).
2. Metropolitan/large country hospitals only, with inputs: non-medical full-time equivalent staff (X1), medical full-time equivalent staff (X2) and non-salary costs (X4).
3. Small rural hospitals, only with inputs: total full-time equivalent staff (X3) and non-salary costs (X4).
4. Metropolitan/large country hospitals only, with inputs: non-medical salaries (X5), medical salaries (X6) and non-salary costs (X4).
5. Small rural hospitals only, with inputs: total salaries (X7) and non-salary costs (X4).

An output orientation was chosen after consultation with the Victorian Department of Human Services. Thus, the relative efficiency of hospitals was measured on their potential to increase outputs (given their existing level of inputs) relative to best practice in the sample. There were three reasons for this choice of orientation:

- the existence of waiting lists for metropolitan acute care indicates that productivity improvements would be best directed to increasing outputs, rather than decreasing inputs;
- in rural areas, medical facilities are provided to a relatively small population, with often limited demand, on the grounds of equity of access to essential services. This means that managers of small rural hospitals may have little scope to reduce their use of inputs; and
- Victorian acute care hospitals are funded on the basis of the outputs they provide, so the incentive is to maximise outputs rather than reduce inputs.

However, funding is based on expected average resource use for particular services, so hospital managers must also ensure efficient input use to remain within budget constraints. Each hospital forms a contract with the Department of Human Services for an agreed target level of WIES. Hospitals that produce more services than this level are not funded for those extra services.

On balance, the incentive of maximising the services provided was considered to be the most appropriate driver of productivity gains.

In consultation with the Department, models 2 and 3 were determined to be the most appropriate. The results of all models are discussed in Section 5.1.5, with the analysis focused on models 2 and 3.

5.1.5 Results and conclusions

Technical efficiency scores

The principal results reported in this section were derived by imposing the assumption of variable returns to scale on each of the models outlined above (See Chapter 2).

The technical efficiency scores indicate which of the hospitals are calculated by the model to be on the efficient (best practice) frontier (those with a score of one), and which are calculated to be less efficient relative to hospitals on the frontier (those with scores greater than one). The higher the score, the higher the potential increase in output (while maintaining inputs) relative to best practice.

Technical efficiency scores only refer to relative performance *within* the sample. Hospitals given an efficiency score of one are efficient relative to all other hospitals in the sample, but may not be efficient by some absolute or world standard necessarily.

Scale efficiency scores

The impact of scale on relative efficiency was also assessed. The effect of size on efficiency was analysed using a three stage process. First, the models were run assuming constant returns to scale. Second, a comparison of the results for constant returns to scale and those for variable returns to scale allowed an assessment of whether the size of a hospital had an influence on its technical efficiency. Finally, to assess the nature of any scale inefficiency, each model was run under the assumption of non-increasing returns to scale. Comparing these final results with results for variable returns to scale enabled hospitals to be described as having increasing, decreasing or constant returns to scale. For a detailed explanation of this three stage process, see Section 2.4.1 and Appendix A.

Model 1 results: all hospitals

Annex 5.1.1 sets out the results of model 1. The variable returns to scale case had twenty-seven hospitals (25 per cent) making up the efficient frontier. None of these had scope to increase one output further, so all were truly efficient relative to all hospitals in the sample.

The average efficiency score of the seventy-nine hospitals off the frontier was 1.29, indicating that these hospitals on average may be able to increase all their outputs by 29 per cent using the same amount of inputs.

After analysing the results of this model and consulting with the Department, it was decided that the results did not accurately reflect the Department's expectations of relative efficiency within Victorian acute care hospitals. Nearly all metropolitan/large country hospitals were relatively less efficient and therefore had small rural hospitals as their peers, or benchmark partners.

This was because there are two important differences in the way that metropolitan/large country and small rural hospitals operate:

1. *Use of medical staff.* Small rural hospitals use visiting medical officers instead of salaried doctors, so they appear to use relatively fewer full-time equivalent staff to produce their outputs than do metropolitan/large country hospitals. This resulted in nearly all metropolitan and large country hospitals being off the efficient frontier, along with small rural hospitals that did have salaried doctors. In several instances, small rural hospitals who employed no doctors were significant peers for major teaching hospitals and specialist research hospitals.

2. *Costs.* The Department advised that small rural hospitals face significantly different costs from metropolitan and large country hospitals, which would affect the quantities of physical inputs they employ.

Given the data difficulties and the significant differences in operating procedures and costs faced by metropolitan/large country hospitals compared with small rural hospitals, the sample was split and models 2 and 3 were run.

Model 2 results: Metropolitan/large country hospitals

Annex 5.1.2 sets out the results of model 2. The variable returns to scale case had twenty-four hospitals (69 per cent) making up the efficient frontier. With the exception of one, all were unable to increase a single output or reduce an input further, so were apparently truly efficient relative to all hospitals in the sample.[9] The one hospital that was able to increase an output and reduce inputs, and four others were apparently 'efficient by default', meaning that they were

---

9 One hospital on the frontier appeared to have scope to reduce its use of non-medical full-time equivalent staff and non-salary costs, and increase production in the output of Y2. After consultation, it was revealed that this hospital had special research functions which may not have been fully captured in the model specification. This view was supported by the fact that this hospital did not appear as a best practice peer for any of the inefficient hospitals. Thus, the hospital was on the frontier by default.

on the frontier as a result of some unique characteristics in the use of inputs and production of outputs which were not explained by the model specification.

A key feature of this model was the high proportion of hospitals on the efficient frontier.

The average relative efficiency score of the thirteen hospitals off the frontier was 1.11, indicating that on average these hospitals could potentially increase all their outputs by 11 per cent using the same amount of inputs.

Average scale efficiency of 1.05 indicated that non-frontier hospitals, on average, might be able to increase their outputs by 5 per cent beyond their best practice targets under variable returns to scale, if they were to operate at constant returns to scale. In addition, it was found that most were apparently larger than the optimal efficient size derived by the model.

The apparent efficiency of non-frontier hospitals was also influenced by the extent to which it appears possible to reduce an input, or expand an output, after all outputs have been expanded uniformly to place the hospital on the production frontier.[10] The extent that it seems possible to reduce an input or expand an output was determined by multiplying the efficiency score of each hospital by its actual level of output or input and then determining the difference between this figure and the target level for the input or output. The total scope for changing each output or input was then expressed as a percentage of the total actual outputs (or inputs), thereby giving an indication of the relative size.

An output oriented study such as this typically reports only how much each output may be increased after all outputs have been increased in the proportion given by the efficiency score. However, this study also reports apparently excess inputs because their existence in an output oriented study indicates that there is *potential* to not only increase output to best practice levels using the same quantity of inputs, but to increase it using fewer inputs. This potential may never be realised, depending on the cause of the excessive input. Apparently excessive use of an input can reflect a low demand for hospital services in a region and the inability of managers to reduce inputs because they are bound to labour agreements or need to provide equitable access to essential services.

In addition to the potential for an average 11 per cent increase in all their outputs as indicated by the efficiency score, non-frontier hospitals may be able to increase output further in two of the four output categories. The model suggests non-frontier hospitals may be able to increase their production of Y2

[10] For an explanation see Section 2.2.

by an average of 12 per cent and the rate of unplanned re-admissions by a further 25 per cent, on average, relative to best practice.

Non-frontier hospitals may also be able to reduce the use of some inputs. The model suggests it may be possible to reduce non-medical staff by 3 per cent, medical staff by 6 per cent and non-salary costs by 9 per cent.

Excluding the five hospitals which were apparently efficient by default, over 50 per cent of hospitals appeared to be efficient. Those hospitals on the frontier were among the larger hospitals, and were likely to be those with access to capital equipment which was unaccounted for in this model. The efficiency scores were therefore biased towards those hospitals which used more capital relative to those that produce the same output with less capital. That is, hospitals with relatively high capital intensities were more likely to make up the frontier, because the partial productivity of their other inputs will be higher relative to that of other hospitals. A model which included capital and a larger sample of large teaching and research hospitals from interstate or overseas would better lend itself to analysis of relative efficiency of these metropolitan/large country hospitals.

Model 3 results: small rural hospitals

Annex 5.1.3 sets out the results of model 3. The variable returns to scale case had fourteen hospitals (20 per cent) making up the efficient frontier. None of these could further reduce inputs, or expand outputs, so all were truly efficient relative to all hospitals in the sample.

The average efficiency score of the fifty-five hospitals off the frontier was 1.33, indicating that these hospitals, on average, may be able to increase all their outputs by 33 per cent using the same amount of inputs.

Average scale efficiency of 1.30 indicates that non-frontier hospitals, on average, may be able to increase their outputs by 30 per cent beyond their best practice targets under variable returns to scale, if they were to operate using constant returns to scale. In addition, the results suggest that most hospitals were larger than the optimal size implied by the model.

As with metropolitan/large country hospitals, the efficiency of non-frontier hospitals was influenced by the scope to further reduce individual inputs or expand outputs, beyond that reflected by the efficiency score. On average, non-frontier hospitals could expand Y2 by 23 per cent and Y3 by 27 per cent. Unplanned re-admissions could be reduced by 10 per cent and Y1 increased by 4 per cent.

The scope for expanding output of Y2 and Y3 was large, but it should be recalled that these classes of output represent more complicated cases. Many small rural hospitals would not have the facilities or qualified staff to treat Y2 and Y3 cases, and such cases would be passed on to metropolitan/large country hospitals typically. Accordingly, the apparently significant scope for increasing output that the model generates for these outputs should be interpreted cautiously. However, the remaining scope for increasing the other outputs suggests that small rural hospitals would be able to increase their number of basic treatments (Y1) and lower their unplanned re-admission rates (beyond the 33 per cent given by the average efficiency score).

The model suggests that non–frontier hospitals may be able to reduce total full-time equivalent staff by 2.8 per cent and non-salary costs by 4.7 per cent. This suggests that if less efficient hospitals operated according to best practice in the sample, then they might not only be able to expand their output using the same amount of inputs, but may be able to produce more output using 2.8 per cent less of total full-time equivalent staff and 4.7 per cent less of non-salary costs. However, this does not account for the practical limitations of reducing inputs, such as contracted labour, or for possible constraints on the demand for outputs of many small rural hospitals.

Sensitivity analysis: models 4 and 5

To test the robustness of the models to changes in the measurement of inputs, models 2 and 3 were run with salary expenditure instead of full-time equivalent staff. Detailed results are given in Annexes 5.1.4 and 5.1.5.

Changing the way in which labour was measured had a minor impact on model results. The hospitals assessed to be on the frontier were largely the same; along with the average efficiency scores, scale efficiency scores and the scope to expand some outputs and decrease inputs. This suggests that:

- wage rates appear to be reasonably consistent in each of the sub-samples; and
- the model is robust in its labour specification.

Further models were run using traditional measures of output, such as adjusted length of stay and acute and non-acute bed days. Analysis and consultations with the Department indicated, as expected, that these did not capture outputs as accurately as did the outputs used in the preferred models 2 and 3.

5.1.6 Future studies

This case study provides a sound starting point for using DEA to assess the efficiency of acute care services in Victorian hospitals. In the development of further models, areas in which the modelling could be improved include:

- the capital input of hospitals;
- a more accurate indicator of the quality of care provided by hospitals; and
- inclusion of public/private patient mix to determine the effect on the efficiency of different patient mixes.

Annex 5.1.1: Model 1 - all hospitals

	Efficiency Score	Scale Score		Actual inputs		Input targets		Actual outputs				Output targets			
HOSPITAL	*PHI (VRS)*	*SE*		*Total FTE staff*	*Non-salary costs ($'000)*	*Total FTE staff*	*Non-salary costs ($'000)*	*WIES1*	*WIES2*	*WIES3*	*Unplanned re-admission rate*[a]	*WIES1*	*WIES2*	*WIES3*	*Unplanned re-admission rate*[a]
HP2	1	1.56		2013.14	48702	2013.14	48702	12115.02	17554.22	16422.44	6.63	12115.02	17554.22	16422.44	6.63
HP3	1	1.41		1463.73	28483	1463.73	28483	3788.66	8428.51	13520.01	13.35	3788.66	8428.51	13520.01	13.35
HP4	1	1.32		1151.83	20069	1151.83	20069	9239.72	10683.23	4250.59	10.28	9239.72	10683.23	4250.59	10.28
HP5	1	1.27		424.62	8870	424.62	8870	1077.78	2511.97	4565.45	4.36	1077.78	2511.97	4565.45	4.36
HP6	1	1.56		2890.62	47790	2890.62	47790	11984.28	16655.07	21799.12	11.27	11984.28	16655.07	21799.12	11.27
HP7	1	1.35		913.22	18310	913.22	18310	5029.88	8718.33	7372.95	8.83	5029.88	8718.33	7372.95	8.83
HP39	1	1.00	*	30.45	1388	30.45	1388	192.21	530.17	724.29	1.32	192.21	530.17	724.29	1.32
HP8	1	1.45		903.5	19687	903.5	19687	6336.38	8167.83	7478.84	9.13	6336.38	8167.83	7478.84	9.13
HP10	1	1.47		1483.4	29820	1483.4	29820	10686.42	13277.6	10082.34	15.50	10686.42	13277.6	10082.34	15.50
HP11	1	1.33		516.22	10355	516.22	10355	3993.22	5014.07	4125.03	8.76	3993.22	5014.07	4125.03	8.76
HP12	1	1.59		99.93	3360	99.93	3360	1041.53	1049.94	977.04	3.09	1041.53	1049.94	977.04	3.09
HP14	1	1.52		2712.85	62241	2712.85	62241	11965.57	17454.41	26061.18	11.33	11965.57	17454.41	26061.18	11.33
HP16	1	1.15		260.62	6121	260.62	6121	2042.32	3225.78	2015.74	4.36	2042.32	3225.78	2015.74	4.36
HP18	1	1.27		175.17	5119	175.17	5119	1544.5	2153.06	1930.04	5.39	1544.5	2153.06	1930.04	5.39
HP20	1	1.35		540.22	12709	540.22	12709	3370.37	5409.52	4997.13	7.29	3370.37	5409.52	4997.13	7.29
HP21	1	1.64		1203.49	30153	1203.49	30153	9210.84	10761.21	8953.54	9.89	9210.84	10761.21	8953.54	9.89
HP41	1	1.06		61.48	1743	61.48	1743	589.63	907.89	733.33	6.47	589.63	907.89	733.33	6.47
HP43	1	1.00	*	47.11	535	47.11	535	407.77	414.09	459.04	12.64	407.77	414.09	459.04	12.64
HP31	1	1.59		933.97	25399	933.97	25399	6852.34	8689.47	7673.53	7.55	6852.34	8689.47	7673.53	7.55
HP32	1	1.61		562.69	17053	562.69	17053	3014.11	5041.9	5692.09	5.49	3014.11	5041.9	5692.09	5.49
HP35	1	1.16		308.2	8551	308.2	8551	2996.7	4160.34	2592.08	7.65	2996.7	4160.34	2592.08	7.65
HP67	1	1.00	*	39.23	268	39.23	268	381.34	321.74	264.98	8.76	381.34	321.74	264.98	8.76
HP78	1	1.00	*	19.44	654	19.44	654	406.03	341.74	267.95	7.11	406.03	341.74	267.95	7.11
HP89	1	1.00	*	3.94	76	3.94	76	120.55	44.67	39.84	11.96	120.55	44.67	39.84	11.96
HP98	1	1.00	*	43.8	301	43.8	301	439.95	377.51	331.13	12.99	439.95	377.51	331.13	12.99
HP101	1	1.00	*	32.5	97	32.5	97	134.14	65.55	79.56	15.13	134.14	65.55	79.56	15.13
HP102	1	1.27		11.6	224	11.6	224	115.11	55.31	30.24	5.14	115.11	55.31	30.24	5.14

Annex 5.1.1: Model 1 - all hospitals (continued)

HOSPITAL	*Efficiency Score* PHI (VRS)	*Scale Score* SE	*Actual inputs* Total FTE staff	*Actual inputs* Non-salary costs ($'000)	*Input targets* Total FTE staff	*Input targets* Non-salary costs ($'000)	*Actual outputs* WIES1	*Actual outputs* WIES2	*Actual outputs* WIES3	*Actual outputs* Unplanned re-admission rate[a]	*Output targets* WIES1	*Output targets* WIES2	*Output targets* WIES3	*Output targets* Unplanned re-admission rate[a]
HP29	1.01	1.82	207.81	4312	207.81	4312	1903.56	1498.89	980.27	11.14	1924.65	2233.81	1120.71	7.48
HP19	1.02	1.37	690.92	14636	690.92	14636	4157.2	6423.26	5763.47	9.53	4231.85	6538.61	5866.97	8.18
HP17	1.03	1.82	1678.45	52260	1678.45	39830.2	7489.94	10300.59	15834.22	9.20	7684.67	11367.47	16245.89	7.63
HP46	1.03	1.33	82.91	1823	82.91	1823	837.29	825.09	645.54	6.55	862.91	927.9	665.29	6.35
HP52	1.03	1.54	127.74	2920	127.74	2920	1264.75	1121.69	741.2	9.34	1297.82	1494.88	830.58	7.32
HP1	1.06	1.72	766.64	18413	766.64	14693.74	6024.45	5976.86	2298.6	7.89	6363.4	7653.81	3447.73	7.47
HP27	1.06	1.67	362.99	11793	362.99	9461.96	2406.49	3182.76	3351.36	7.84	2554.34	3602.91	3557.27	6.05
HP63	1.07	1.25	37.91	783	37.91	783	479.42	371.82	306.31	12.53	512.1	464.95	327.43	8.02
HP86	1.08	1.43	49.02	880	49.02	880	528.49	384.14	281.65	8.32	570.5	542.76	372.91	7.70
HP92	1.08	1.52	52.09	1254	52.09	1254	598.46	452.37	370.54	6.33	647.07	649.44	409.49	5.86
HP24	1.09	1.45	404.55	12359	404.55	11600.88	2490.45	4204.52	3400.87	9.92	3093.79	4569.38	3695.99	6.80
HP15	1.1	1.39	202.71	5376	202.71	5376	1672.47	1958.33	1700.72	3.92	1844.51	2530.19	1875.66	3.56
HP26	1.1	1.49	355.84	10277	355.84	9345.04	2258.65	3201.23	3211.35	6.74	2476.68	3510.25	3521.34	5.95
HP44	1.1	1.61	148.29	2654	148.29	2654	1182.27	995.63	886.83	9.03	1295.52	1448.69	971.78	8.24
HP58	1.1	1.33	53.93	1039	53.93	1039	539.58	435.03	431.47	7.35	593.62	599.27	474.68	6.68
HP60	1.12	1.35	19.1	191	7.84	191	104.16	39.18	63.17	7.23	122.28	72.37	70.61	6.47
HP100	1.12	1.01	30.14	1114	30.14	1114	225.94	433.02	499.96	3.26	297.58	486.88	562.15	2.04
HP94	1.13	1.00	12.82	288	12.82	288	135.71	132.13	154.17	10.87	167.68	150.18	174.81	6.24
HP47	1.14	1.28	26.04	815	26.04	815	409.13	296.46	212.45	9.73	463.01	421.2	312.41	7.12
HP33	1.14	1.03	43.6	970	43.6	970	356.29	491.96	444.1	5.54	444.41	561.07	506.49	4.68
HP79	1.14	1.69	168.23	4087	168.23	4087	1462.39	1257.31	910.45	8.56	1667.94	2054.26	1182.98	7.39
HP36	1.15	1.30	222.06	6645	222.06	6480.56	1533.92	2410.66	1895.01	3.69	1899.64	2774.54	2181.06	3.21
HP85	1.15	1.16	11.97	235	11.97	235	170.85	109.83	61.49	9.19	196.45	142.6	130.52	7.99
HP71	1.16	1.47	24	505	24	505	331.53	154.61	165.56	12.24	385.12	320.1	260.36	8.42
HP83	1.17	1.03	37.36	873	37.36	873	397.5	436.37	316.71	8.67	463.85	509.21	416.86	7.43
HP103	1.18	1.28	15.63	290	15.63	290	203.02	123.53	95.56	9.20	238.95	184.66	161.75	7.82
HP40	1.19	1.39	121.42	2071	121.42	2071	875.23	873.46	739.83	8.38	1043.34	1172.31	881.93	7.03
HP30	1.19	1.30	194.97	4166	194.97	4166	817.25	1579.32	1593.13	9.40	1059.48	1873.64	1890.03	2.42

Annex 5.1.1: Model 1 - all hospitals (continued)

	Efficiency Score	Scale Score	Actual inputs		Input targets		Actual outputs				Output targets			
HOSPITAL	*PHI (VRS)*	*SE*	*Total FTE staff*	*Non-salary costs ($'000)*	*Total FTE staff*	*Non-salary costs ($'000)*	*WIES1*	*WIES2*	*WIES3*	*Unplanned re-admission rate*[a]	*WIES1*	*WIES2*	*WIES3*	*Unplanned re-admission rate*[a]
HP42	1.2	1.45	117.4	2583	117.4	2583	959.87	913.57	769.19	9.83	1149.96	1338.52	921.52	8.21
HP95	1.2	1.35	85.75	2052	85.75	2052	733.66	714.1	682.57	7.59	879.65	992.03	818.39	6.33
HP34	1.21	1.30	248.38	6920	248.38	6920	1663.77	2414.44	2045.02	7.86	2034.21	2931.74	2483.17	5.85
HP62	1.21	1.47	27.01	445	27.01	445	268.73	170.98	127.37	6.87	326.01	299.46	270.32	5.67
HP74	1.21	1.45	22.6	305	13.69	305	120.69	66.38	101.18	5.65	146.51	122.68	122.83	4.66
HP76	1.21	1.16	7.68	115	5.84	115	71.98	45.12	39.54	11.11	123.41	54.39	47.67	9.22
HP99	1.22	1.64	12.43	226	12.43	226	137.96	39.69	25.44	8.04	168.75	130.83	132.55	6.57
HP9	1.23	2.27	453.1	16626	453.1	11343.66	3041.92	2166.17	2751.16	3.76	3742.01	4733.75	3384.33	3.06
HP22	1.25	1.47	286.45	6608	286.45	6608	1996.59	2176.24	1792.74	6.14	2495.76	3230.32	2240.95	4.91
HP25	1.25	1.47	530.76	12806	530.76	12806	3340.42	4055.14	3567.89	8.42	4174.04	5576.59	4458.27	6.74
HP23	1.27	1.00	12.95	167	12.95	167	84.31	90.74	97.7	12.77	186.09	126.83	123.33	10.12
HP75	1.27	1.52	247.08	5632	247.08	5632	1648.62	1646.87	1682.25	9.55	2101.2	2665.26	2144.06	7.10
HP48	1.29	1.33	119.83	2223	119.83	2223	872.49	873.23	569.76	10.13	1124.03	1212.9	734.02	7.86
HP59	1.29	1.03	38.11	933	38.11	933	314.25	402.98	358.04	11.86	434.76	520.47	462.43	4.84
HP64	1.29	1.03	37.33	675	37.33	675	356.46	347.33	274.72	13.68	461.58	449.76	365.76	8.83
HP90	1.29	1.49	18.38	286	18.38	286	216.43	98.85	91.91	18.73	278.51	209.16	175.45	10.33
HP55	1.3	1.30	115.27	2984	115.27	2984	903.6	1027.98	824.33	9.66	1170.77	1432.79	1068.06	7.06
HP73	1.31	1.12	13.24	173	13.24	173	160.33	74.12	82.39	15.34	209.7	137.54	118.22	11.43
HP93	1.31	1.15	14.83	294	14.83	294	179.26	123.56	67.25	10.20	235.3	181.18	159.08	7.78
HP84	1.32	1.10	35.06	733	35.06	733	326.56	294.25	300.1	11.81	432.54	421.63	397.5	7.87
HP45	1.33	1.25	47.4	964	47.4	964	435.45	371.31	302.99	10.86	579.52	556.65	403.23	8.16
HP69	1.33	1.00	15.51	291	15.51	291	178.18	141.65	125.51	15.29	245.27	188.06	166.63	8.50
HP53	1.36	1.39	25.75	438	25.75	438	270.87	155.9	103.27	11.43	368.59	304.37	246.81	8.40
HP80	1.36	1.05	31.94	763	31.94	763	277.96	316.31	248.89	5.90	393.5	430.07	403.04	4.34
HP72	1.37	1.23	37.61	1300	37.61	1244.69	391.38	392.16	281.84	6.71	537.44	538.51	446.09	4.89
HP87	1.41	1.28	8.3	146	6.25	146	56.82	30.87	42.28	11.19	121.79	62.13	59.75	7.92
HP91	1.41	1.20	39.68	839	39.68	839	318.06	306.83	251.07	5.83	448.99	478.48	414.82	4.13
HP96	1.41	1.67	17.95	378	17.95	378	188.17	72.58	88.97	9.21	265.97	225.43	202.71	6.51

Annex 5.1.1: Model 1 - all hospitals (continued)

	Efficiency Score	*Scale Score*	*Actual inputs*		*Input targets*		*Actual outputs*				*Output targets*			
HOSPITAL	*PHI (VRS)*	*SE*	*Total FTE staff*	*Non-salary costs ($'000)*	*Total FTE staff*	*Non-salary costs ($'000)*	*WIES1*	*WIES2*	*WIES3*	*Unplanned re-admission rate*[a]	*WIES1*	*WIES2*	*WIES3*	*Unplanned re-admission rate*[a]
HP51	1.44	1.69	26.84	568	26.84	568	276.28	113.36	127.27	9.37	397.5	351.54	296.14	6.51
HP105	1.44	1.41	32.94	482	32.94	482	261.79	174.83	128.81	8.14	375.83	349.58	310.52	5.67
HP97	1.45	1.03	26.79	529	26.79	529	189.32	229.4	174.22	8.53	360.45	333.55	303.47	5.86
HP28	1.46	1.30	269.47	6124	269.47	6124	1146.75	1902.48	1731.84	8.09	1679.25	2785.91	2536.03	2.02
HP56	1.46	1.22	124.39	2750	124.39	2750	696.43	918.69	837.8	7.10	1013.49	1336.93	1219.22	4.31
HP50	1.48	1.92	31.7	490	31.7	490	275.51	89.17	79.36	10.99	406.93	357.5	305.18	7.44
HP66	1.48	1.67	30.15	1968	30.15	946.9	338.5	159.09	117.97	13.04	502.12	483.37	354.15	7.13
HP49	1.5	1.28	77.56	1296	77.56	1296	508.28	477.76	204.86	10.37	762.65	783.95	486.47	6.92
HP61	1.53	1.00	20.51	439	20.51	439	162.48	169.48	166.73	10.64	272.14	257.36	253.19	5.39
HP65	1.53	1.18	17.89	322	17.89	322	174.61	118.82	86	12.05	267	210.58	179.53	7.88
HP82	1.53	1.03	6.31	89	5.55	89	80.13	36.19	16.78	17.15	129.98	55.34	47.79	11.21
HP88	1.53	1.72	32.92	560	32.92	560	267.18	139.98	119.72	9.36	409.66	382.24	340.8	6.11
HP37	1.6	1.52	849.95	15990	849.95	15990	2578.66	3261.6	4483.67	8.22	4113.43	5866.12	7152.27	5.15
HP57	1.61	2.04	35	743	35	743	297.75	129.24	114.41	11.27	479.2	439.46	329.44	7.00
HP68	1.66	1.82	27.81	487	27.81	487	242.37	101.73	87.18	21.14	402.66	338.41	278.96	8.83
HP106	1.69	1.00	28.69	537	28.69	537	186.51	203.79	191.94	10.33	357.29	344.03	324.03	5.32
HP54	1.81	2.44	58.45	952	58.45	952	311.58	121.62	129.48	8.68	564.24	600.51	463.84	4.79
HP81	1.83	1.00	8.45	102	8.45	102	77.77	33.38	40.06	22.22	156.54	82.35	73.06	12.06
HP77	1.96	2.00	18.15	350	18.15	350	118.46	53.51	46.08	11.11	232.7	210.25	206.16	5.66
HP70	2.08	1.52	9.46	158	9.46	158	79.15	21.22	15.29	19.31	164.96	102.73	94.36	9.26
NON-FRONTIER HOSPITALS														
TOTAL			11540	285793	11515	259100	75018	85619	81538		88985	112898	98756	
AVERAGE	1.29	1.33								9.67				6.74
No. of less efficient hospitals			79											
% less efficient hospitals			75											

Notes: * indicates hospital is the optimal size defined by Constant Returns to Scale

a the inverse of this was used as an output in the model

Slacks are derived by the product of the efficiency score and the actual input or output, subtracted from the target for the input or output.

Annex 5.1.2: Model 2 - Metropolitan/large country hospitals

	Efficiency score	Scale score	Actual inputs			Input targets			Actual outputs				Output targets			
HOSPITAL	*PHI (VRS)*	*SE*	*Non-medical FTE staff*	*Medical FTE staff*	*Non-salary costs ($'000)*	*Non-medical FTE staff*	*Medical FTE staff*	*Non-salary costs ($'000)*	*WEIS1*	*WEIS2*	*WEIS3*	*Unplanned re-admission rate*[a]	*WEIS1*	*WEIS2*	*WEIS3*	*Unplanned re-admission rate*[a]
HP1	1	1.18	715.62	51.02	18413	715.62	51.02	18413	6024.45	5976.86	2298.6	7.89	6024.45	5976.86	2298.6	7.89
HP2	1	1.18	1739.3	273.84	48702	1739.3	273.84	48702	12115.02	17554.22	16422.44	6.63	12115.02	17554.22	16422.44	6.63
HP3	1	1.05	1283.44	180.29	28483	1283.44	180.29	28483	3788.66	8428.51	13520.01	13.35	3788.66	8428.51	13520.01	13.35
HP4	1	1.00 *	1063.22	88.61	20069	1063.22	88.61	20069	9239.72	10683.23	4250.59	10.28	9239.72	10683.23	4250.59	10.28
HP5	1	1.00 *	374.17	50.45	8870	374.17	50.45	8870	1077.78	2511.97	4565.45	4.36	1077.78	2511.97	4565.45	4.36
HP6	1	1.00 *	2600.09	290.54	47790	2600.09	290.54	47790	11984.28	16655.07	21799.12	11.27	11984.28	16655.07	21799.12	11.27
HP7	1	1.00 *	812.3	100.92	18310	812.3	100.92	18310	5029.88	8718.33	7372.95	8.83	5029.88	8718.33	7372.95	8.83
HP8	1	1.05	816.95	86.55	19687	816.95	86.55	19687	6336.38	8167.83	7478.84	9.13	6336.38	8167.83	7478.84	9.13
HP9	1	1.47	425.8	27.3	16626	346.38	27.3	9594.13	3041.92	2166.17	2751.16	3.76	3049.74	3477.16	2758.23	3.75
HP10	1	1.09	1334.27	149.13	29820	1334.27	149.13	29820	10686.42	13277.6	10082.34	15.50	10686.42	13277.6	10082.34	15.50
HP11	1	1.00 *	458.25	57.97	10355	458.25	57.97	10355	3993.22	5014.07	4125.03	8.76	3993.22	5014.07	4125.03	8.76
HP12	1	1.00 *	99.93	0	3360	99.93	0	3360	1041.53	1049.94	977.04	3.09	1041.53	1049.94	977.04	3.09
HP14	1	1.08	2390.54	322.31	62241	2390.54	322.31	62241	11965.57	17454.41	26061.18	11.33	11965.57	17454.41	26061.18	11.33
HP16	1	1.00 *	255.15	5.47	6121	255.15	5.47	6121	2042.32	3225.78	2015.74	4.36	2042.32	3225.78	2015.74	4.36
HP18	1	1.00 *	165.65	9.52	5119	165.65	9.52	5119	1544.5	2153.06	1930.04	5.39	1544.5	2153.06	1930.04	5.39
HP20	1	1.01	500.1	40.12	12709	500.1	40.12	12709	3370.37	5409.52	4997.13	7.29	3370.37	5409.52	4997.13	7.29
HP21	1	1.14	1084.98	118.51	30153	1084.98	118.51	30153	9210.84	10761.21	8953.54	9.89	9210.84	10761.21	8953.54	9.89
HP24	1	1.15	391.57	12.98	12359	391.57	12.98	12359	2490.45	4204.52	3400.87	9.92	2490.45	4204.52	3400.87	9.92
HP29	1	1.00 *	204.93	2.87	4312	204.93	2.87	4312	1903.56	1498.89	980.27	11.14	1903.56	1498.89	980.27	11.14
HP30	1	1.00 *	191.31	3.67	4166	191.31	3.67	4166	817.25	1579.32	1593.13	9.40	817.25	1579.32	1593.13	9.40
HP31	1	1.19	851.89	82.08	25399	851.89	82.08	25399	6852.34	8689.47	7673.53	7.55	6852.34	8689.47	7673.53	7.55
HP32	1	1.05	540.56	22.13	17053	540.56	22.13	17053	3014.11	5041.9	5692.09	5.49	3014.11	5041.9	5692.09	5.49
HP35	1	1.00 *	291.91	16.29	8551	291.91	16.29	8551	2996.7	4160.34	2592.08	7.65	2996.7	4160.34	2592.08	7.65
HP36	1	1.08	216.96	5.1	6645	216.96	5.1	6645	1533.92	2410.66	1895.01	3.69	1533.92	2410.66	1895.01	3.69
HP17	1.01	1.09	1456.91	221.54	52260	1456.91	191.45	38216.56	7489.94	10300.59	15834.22	9.20	7543.63	10997.04	15947.71	7.64
HP19	1.02	1.01	623.58	67.34	14636	623.58	67.34	14636	4157.2	6423.26	5763.47	9.53	4223.49	6525.68	5855.37	8.10

Annex 5.1.2: Model 2 - Metropolitan/large country hospitals (continued)

	Efficiency score	*Scale score*	*Actual inputs*			*Input targets*			*Actual outputs*				*Output targets*			
HOSPITAL	*PHI (VRS)*	*SE*	*Non-medical FTE staff*	*Medical FTE staff*	*Non-salary costs ($'000)*	*Non-medical FTE staff*	*Medical FTE staff*	*Non-salary costs ($'000)*	*WEIS1*	*WEIS2*	*WEIS3*	*Unplanned re-admission rate*[a]	*WEIS1*	*WEIS2*	*WEIS3*	*Unplanned re-admission rate*[a]
HP13	1.04	1.05	778.54	91.21	17141	778.54	89.99	17141	5598.75	6678.03	6422.84	10.21	5835.91	7455.77	6694.91	9.49
HP15	1.05	1.03	191.99	10.72	5376	191.99	10.72	5376	1672.47	1958.33	1700.72	3.92	1762.53	2149.19	1792.3	3.72
HP23	1.06	1.03	242.95	4.13	5632	211.96	4.13	5514.16	1648.62	1646.87	1682.25	9.55	1739.93	2560.16	1775.42	3.86
HP27	1.06	1.10	345.99	17	11793	345.99	17	10640.68	2406.49	3182.76	3351.36	7.84	2553.65	3859.51	3556.3	5.99
HP26	1.07	1.09	341.17	14.67	10277	341.17	14.67	10277	2258.65	3201.23	3211.35	6.74	2423.34	3751.04	3445.51	5.48
HP33	1.08	1.01	164.09	4.14	4087	164.09	2.22	4087	1462.39	1257.31	910.45	8.56	1584.23	1420.67	1046.77	5.24
HP25	1.09	1.18	504.42	26.33	12806	461.44	26.33	12806	3340.42	4055.14	3567.89	8.42	3656.91	5159.86	3905.93	6.88
HP34	1.13	1.04	240.43	7.95	6920	240.43	7.95	6920	1663.77	2414.44	2045.02	7.86	1872.9	2763.39	2302.07	4.32
HP22	1.16	1.02	277.77	8.68	6608	277.77	8.68	6608	1996.59	2176.24	1792.74	6.14	2311.89	3412.51	2075.84	4.76
HP28	1.26	1.05	262.16	7.3	6124	250.19	7.3	6124	1146.75	1902.48	1731.84	8.09	1450.03	2490.56	2189.86	6.40
HP37	1.4	1.08	797.08	52.87	15990	655.48	52.87	15990	2578.66	3261.6	4483.67	8.22	3609.49	5699.57	6276.03	5.88
NON-FRONTIER HOSPITALS																
TOTAL			6227.08	533.88	169650	5999.54	500.65	154336.4	37420.7	48458.28	52497.82		40567.93	58244.95	56864.02	
AVERAGE	1.11	1.05										8.02				5.98
No. of less efficient hospitals			13													
% less efficient hospitals			35													

Annex 5.1.3: Model 3 – Small rural hospitals

	Efficiency score	*Scale score*	*Actual inputs*		*Input targets*		*Actual outputs*				*Output targets*			
HOSPITAL	*PHI (VRS)*	*SE*	*Total FTE staff*	*Non-salary costs ($'000)*	*Total FTE staff*	*Non-salary costs ($'000)*	*WEIS1*	*WEIS2*	*WEIS3*	*Unplanned re-admission rate* [a]	*WEIS1*	*WEIS2*	*WEIS3*	*Unplanned re-admission rate*[a]
HP39	1	1.06	61.48	1743	61.48	1743	589.63	907.89	733.33	6.47	589.63	907.89	733.33	6.47
HP41	1	1.00 *	30.45	1388	30.45	1388	192.21	530.17	724.29	1.32	192.21	530.17	724.29	1.32
HP43	1	1.79	148.29	2654	148.29	2654	1182.27	995.63	886.83	9.03	1182.27	995.63	886.83	9.03
HP44	1	1.00 *	47.11	535	47.11	535	407.77	414.09	459.04	12.64	407.77	414.09	459.04	12.64
HP46	1	1.00 *	32.5	97	32.5	97	134.14	65.55	79.56	15.13	134.14	65.55	79.56	15.13
HP52	1	1.00 *	19.44	654	19.44	654	406.03	341.74	267.95	7.11	406.03	341.74	267.95	7.11
HP55	1	1.37	82.91	1823	82.91	1823	837.29	825.09	645.54	6.55	837.29	825.09	645.54	6.55
HP67	1	1.25	93.6	1643	93.6	1643	621.72	887.73	801.96	8.11	621.72	887.73	801.96	8.11
HP78	1	1.00 *	3.94	76	3.94	76	120.55	44.67	39.84	11.96	120.55	44.67	39.84	11.96
HP89	1	1.59	127.74	2920	127.74	2920	1264.75	1121.69	741.2	9.34	1264.75	1121.69	741.2	9.34
HP98	1	1.00 *	43.8	301	43.8	301	439.95	377.51	331.13	12.99	439.95	377.51	331.13	12.99
HP101	1	1.00 *	39.23	268	39.23	268	381.34	321.74	264.98	8.76	381.34	321.74	264.98	8.76
HP102	1	1.67	115.27	2984	115.27	2984	903.6	1027.98	824.33	9.66	903.6	1027.98	824.33	9.66
HP104	1	1.27	11.6	224	11.6	224	115.11	55.31	30.24	5.14	115.11	55.31	30.24	5.14
HP56	1.02	1.75	124.39	2750	124.39	2271.85	696.43	918.69	837.8	7.10	951.54	933.24	851.07	6.37
HP40	1.04	1.59	121.42	2071	109.72	2071	875.23	873.46	739.83	8.38	906.65	904.82	766.39	8.08
HP92	1.04	1.56	52.09	1254	52.09	1254	598.46	452.37	370.54	6.33	623.33	565.25	431.95	6.08
HP38	1.05	1.02	26.42	1139	26.42	1139	156.39	398.55	362.44	1.64	175.72	428.59	575.82	1.56
HP63	1.05	1.27	37.91	783	37.91	783	479.42	371.82	306.31	12.53	504.96	433.08	335.75	8.46
HP86	1.05	1.47	49.02	880	49.02	880	528.49	384.14	281.65	8.32	554.58	492.1	390.6	7.93
HP42	1.06	1.64	117.4	2583	117.4	2419.85	959.87	913.57	769.19	9.83	1013.02	991.85	811.78	8.08
HP58	1.08	1.35	53.93	1039	53.93	1039	539.58	435.03	431.47	7.35	581.35	562.69	464.87	6.82
HP95	1.09	1.49	85.75	2052	85.75	2052	733.66	714.1	682.57	7.59	800.38	940.2	744.64	6.96
HP48	1.11	1.54	119.83	2223	107.65	2223	872.49	873.23	569.76	10.13	965.5	966.32	718.38	9.15
HP60	1.12	1.35	19.1	191	7.84	191	104.16	39.18	63.17	7.23	122.28	72.37	70.61	6.47

Annex 5.1.3: Model 3 - Small rural hospitals (continued)

	Efficiency score	*Scale score*	*Actual inputs*		*Input targets*		*Actual outputs*				*Output targets*			
HOSPITAL	*PHI (VRS)*	*SE*	*Total FTE staff*	*Non-salary costs ($'000)*	*Total FTE staff*	*Non-salary costs ($'000)*	*WEIS1*	*WEIS2*	*WEIS3*	*Unplanned re-admission rate* [a]	*WEIS1*	*WEIS2*	*WEIS3*	*Unplanned re-admission rate* [a]
HP94	1.12	1.28	26.04	815	26.04	792.09	409.13	296.46	212.45	9.73	458.36	389.27	296.79	7.21
HP100	1.12	1.01	30.14	1114	30.14	1114	225.94	433.02	499.96	3.26	297.58	486.88	562.15	2.04
HP79	1.13	1.01	12.82	288	12.82	288	135.71	132.13	154.17	10.87	167.68	150.18	174.81	6.24
HP47	1.14	1.03	43.6	970	43.6	970	356.29	491.96	444.1	5.54	444.41	561.07	506.49	4.68
HP85	1.15	1.16	11.97	235	11.97	235	170.85	109.83	61.49	9.19	196.45	142.6	130.52	7.99
HP71	1.16	1.47	24	505	24	505	331.53	154.61	165.56	12.24	385.12	320.1	260.36	8.42
HP83	1.17	1.03	37.36	873	37.36	873	397.5	436.37	316.71	8.67	463.63	508.97	416.88	7.43
HP103	1.18	1.28	15.63	290	15.63	290	203.02	123.53	95.56	9.20	238.95	184.66	161.75	7.82
HP62	1.21	1.47	27.01	445	27.01	445	268.73	170.98	127.37	6.87	326.01	299.46	270.32	5.67
HP74	1.21	1.45	22.6	305	13.69	305	120.69	66.38	101.18	5.65	146.51	122.68	122.83	4.66
HP76	1.21	1.16	7.68	115	5.84	115	71.98	45.12	39.54	11.11	123.41	54.39	47.67	9.22
HP99	1.22	1.64	12.41	226	12.41	226	137.96	39.69	25.44	8.04	168.7	130.86	132.69	6.57
HP75	1.26	1.01	12.95	167	12.95	167	84.31	90.74	97.7	12.77	186.09	126.83	123.33	10.12
HP64	1.28	1.04	37.33	675	37.33	675	356.46	347.33	274.72	13.68	456.7	445.01	367.05	8.70
HP59	1.29	1.03	38.11	933	38.11	933	314.25	402.98	358.04	11.86	434.76	520.47	462.43	4.84
HP90	1.29	1.49	18.38	286	18.38	286	216.43	98.85	91.91	18.73	278.51	209.16	175.45	10.33
HP72	1.3	1.30	37.61	1300	37.61	1100.37	391.38	392.16	281.84	6.71	508.03	509.05	412.36	5.17
HP45	1.31	1.28	47.4	964	47.4	964	435.45	371.31	302.99	10.86	569.2	508.05	396.05	8.21
HP73	1.31	1.12	13.24	173	13.24	173	160.33	74.12	82.39	15.34	209.7	137.54	118.22	11.43
HP84	1.31	1.11	35.06	733	35.06	733	326.56	294.25	300.1	11.81	429.27	436.39	394.49	8.44
HP93	1.31	1.15	14.83	294	14.83	294	179.26	123.56	67.25	10.20	235.3	181.18	159.08	7.78
HP69	1.33	1.00	15.51	291	15.51	291	178.18	141.65	125.51	15.29	245.27	188.06	166.63	8.50
HP53	1.36	1.05	31.94	763	31.94	763	277.96	316.31	248.89	5.90	393.5	430.07	403.04	4.34
HP80	1.36	1.39	25.75	438	25.75	438	270.87	155.9	103.27	11.43	368.59	304.37	246.81	8.40
HP49	1.38	1.39	77.56	1296	70.45	1296	508.28	477.76	204.86	10.37	702.55	660.36	522.73	7.51
HP91	1.39	1.22	39.68	839	39.68	839	318.06	306.83	251.07	5.83	442.63	454.49	422.05	4.19

Annex 5.1.3: Model 3 - Small rural hospitals (continued)

	Efficiency score	*Scale score*	*Actual inputs*		*Input targets*		*Actual outputs*				*Output targets*			
HOSPITAL	*PHI (VRS)*	*SE*	*Total FTE staff*	*Non-salary costs ($'000)*	*Total FTE staff*	*Non-salary costs ($'000)*	*WEIS1*	*WEIS2*	*WEIS3*	*Unplanned re-admission rate*[a]	*WEIS1*	*WEIS2*	*WEIS3*	*Unplanned re-admission rate*[a]
HP87	1.41	1.28	8.3	146	6.25	146	56.82	30.87	42.28	11.19	121.79	62.13	59.75	7.92
HP96	1.41	1.67	17.95	378	17.95	378	188.17	72.58	88.97	9.21	265.97	225.43	202.71	6.51
HP51	1.44	1.41	32.94	482	32.94	482	261.79	174.83	128.81	8.14	375.83	349.58	310.52	5.67
HP105	1.44	1.69	26.84	568	26.84	568	276.28	113.36	127.27	9.37	397.5	351.54	296.14	6.51
HP66	1.45	1.03	26.79	529	26.79	529	189.32	229.4	174.22	8.53	360.45	333.55	303.47	5.86
HP97	1.45	1.72	30.15	1968	30.15	878.09	338.5	159.09	117.97	13.04	490.95	418.87	314.75	7.28
HP50	1.48	1.92	31.7	490	31.7	490	275.51	89.17	79.36	10.99	406.93	357.5	305.18	7.44
HP65	1.52	1.01	20.51	439	20.51	439	162.48	169.48	166.73	10.64	272.14	257.36	253.19	5.39
HP61	1.53	1.18	17.89	322	17.89	322	174.61	118.82	86	12.05	267	210.58	179.53	7.88
HP82	1.53	1.72	32.92	560	32.92	560	267.18	139.98	119.72	9.36	409.12	380.35	341.41	6.12
HP88	1.53	1.03	6.31	89	5.55	89	80.13	36.19	16.78	17.15	129.98	55.34	47.79	11.21
HP57	1.59	2.08	35	743	35	743	297.75	129.24	114.41	11.27	472.25	416.45	337.19	7.11
HP68	1.66	1.82	27.81	487	27.81	487	242.37	101.73	87.18	21.14	402.66	338.41	278.96	8.83
HP106	1.69	1.00	28.69	537	28.69	537	186.51	203.79	191.94	10.33	357.29	344.03	324.03	5.32
HP54	1.74	2.56	58.45	952	58.45	952	311.58	121.62	129.48	8.68	543.15	531.85	485.14	4.98
HP81	1.82	1.01	8.45	102	8.45	102	77.77	33.38	40.06	22.22	156.54	82.35	73.06	12.06
HP77	1.96	2.00	18.15	350	18.15	350	118.46	53.51	46.08	11.11	232.7	210.25	206.16	5.66
HP70	2.08	1.52	9.46	158	9.46	158	79.15	21.22	15.29	19.31	164.96	102.73	94.36	9.26
NON-FRONTIER HOSPITALS														
TOTAL			1962.18	41598	1906.37	39644.25	16975.67	14466.23	12221.35		21903.43	20780.96	18028.18	
AVERAGE	1.33	1.30								10.39				7.14
No. of less efficient hospitals			55											
% less efficient hospitals			80											

Notes : * indicates hospital is the optimal size defined by Constant Returns to Scale

a the inverse of this was used as an output in the model

Slacks are derived by the product of the efficiency score and the actual input or output, subtracted from the target for the input or output.

Annex 5.1.4: Model 4 – Metropolitan/large country hospitals

	Efficiency score	*Scale score*	*Actual inputs*			*Input targets*			*Actual outputs*				*Output targets*			
HOSPITAL	*PHI (VRS)*	*SE*	*Non-medical salaries ($'000)*	*Medical salaries ($'000)*	*Non-salary costs ($'000)*	*Non-medical salaries ($'000)*	*Medical salaries ($'000)*	*Non- salary costs ($'000)*	*WEIS1*	*WEIS2*	*WEIS3*	*Unplanned re-admission rate*[a]	*WEIS1*	*WEIS2*	*WEIS3*	*Unplanned re-admission rate*[a]
HP2	1	1.19	70781	22636	48702	70781	22636	48702	12115.02	17554.22	16422.44	6.63	12115.02	17554.22	16422.44	6.63
HP3	1	1.05	44275	23291	28483	44275	23291	28483	3788.66	8428.51	13520.01	13.35	3788.66	8428.51	13520.01	13.35
HP4	1	1.00 *	42997	9132	20069	42997	9132	20069	9239.72	10683.23	4250.59	10.28	9239.72	10683.23	4250.59	10.28
HP5	1	1.00 *	12689	5528	8870	12689	5528	8870	1077.78	2511.97	4565.45	4.36	1077.78	2511.97	4565.45	4.36
HP6	1	1.00 *	68551	17730	47790	68551	17730	47790	11984.28	16655.07	21799.12	11.27	11984.28	16655.07	21799.12	11.27
HP7	1	1.00 *	30419	8515	18310	30419	8515	18310	5029.88	8718.33	7372.95	8.83	5029.88	8718.33	7372.95	8.83
HP10	1	1.08	52685	13802	29820	52685	13802	29820	10686.42	13277.6	10082.34	15.50	10686.42	13277.6	10082.34	15.50
HP11	1	1.00 *	19931	6390	10355	19931	6390	10355	3993.22	5014.07	4125.03	8.76	3993.22	5014.07	4125.03	8.76
HP12	1	1.00 *	3526	0	3360	3526	0	3360	1041.53	1049.94	977.04	3.09	1041.53	1049.94	977.04	3.09
HP14	1	1.10	86757	27362	62241	86757	27362	62241	11965.57	17454.41	26061.18	11.33	11965.57	17454.41	26061.18	11.33
HP16	1	1.00 *	10210	396	6121	10210	396	6121	2042.32	3225.78	2015.74	4.36	2042.32	3225.78	2015.74	4.36
HP18	1	1.00 *	7998	644	5119	7998	644	5119	1544.5	2153.06	1930.04	5.39	1544.5	2153.06	1930.04	5.39
HP19	1	1.00 *	19943	6008	14636	19943	6008	14636	4157.2	6423.26	5763.47	9.53	4157.2	6423.26	5763.47	9.53
HP21	1	1.00 *	54996	208	30153	54996	208	30153	9210.84	10761.21	8953.54	9.89	9210.84	10761.21	8953.54	9.89
HP26	1	1.04	11650	986	10277	11650	986	10277	2258.65	3201.23	3211.35	6.74	2258.65	3201.23	3211.35	6.74
HP29	1	1.00 *	6815	203	4312	6815	203	4312	1903.56	1498.89	980.27	11.14	1903.56	1498.89	980.27	11.14
HP30	1	1.00 *	5903	284	4166	5903	284	4166	817.25	1579.32	1593.13	9.40	817.25	1579.32	1593.13	9.40
HP31	1	1.16	36800	3282	25399	36800	3282	25399	6852.34	8689.47	7673.53	7.55	6852.34	8689.47	7673.53	7.55
HP32	1	1.01	20018	1445	17053	20018	1445	17053	3014.11	5041.9	5692.09	5.49	3014.11	5041.9	5692.09	5.49
HP35	1	1.00 *	11299	1943	8550	11299	1943	8550	2996.7	4160.34	2592.08	7.65	2996.7	4160.34	2592.08	7.65
HP36	1	1.03	7678	458	6645	7678	458	6645	1533.92	2410.66	1895.01	3.69	1533.92	2410.66	1895.01	3.69
HP8	1.01	1.03	33599	7825	19687	33599	7825	19687	6336.38	8167.83	7478.84	9.13	6395.87	8244.52	7549.06	8.24
HP15	1.01	1.01	6815	649	5376	6815	649	5376	1672.47	1958.33	1700.72	3.92	1696.78	1986.8	1725.44	3.66
HP24	1.01	1.11	14886	1048	12359	14886	1048	10625.44	2490.45	4204.52	3400.87	9.92	2814.74	4238.69	3428.5	5.32
HP17	1.03	1.04	50991	18822	52260	50991	13894.35	35555.65	7489.94	10300.59	15834.22	9.20	8555.86	12209.24	16381.78	7.52
HP13	1.04	1.03	29539	8476	17141	29539	8386.06	17141	5598.75	6678.03	6422.84	10.21	5839.76	7491.07	6699.33	9.58
HP1	1.05	1.22	29352	5091	18413	29352	5091	15311.47	6024.45	5976.86	2298.6	7.89	6316.22	7522.39	3730.99	7.52

Annex 5.1.4: Model 4 - Metropolitan/large country hospitals (continued)

	Efficiency score	Scale score	Actual inputs			Input targets			Actual outputs				Output targets			
HOSPITAL	*PHI (VRS)*	*SE*	*Non-medical salaries ($'000)*	*Medical salaries ($'000)*	*Non-salary costs ($'000)*	*Non-medical salaries ($'000)*	*Medical salaries ($'000)*	*Non-salary costs ($'000)*	*WEIS1*	*WEIS2*	*WEIS3*	*Unplanned re-admission rate[a]*	*WEIS1*	*WEIS2*	*WEIS3*	*Unplanned re-admission rate[a]*
HP9	1.05	1.54	16579	2716	16626	16579	2716	11222.81	3041.92	2166.17	2751.16	3.76	3196.86	3980.37	3499.2	3.58
HP20	1.05	1.00	20608	5920	12709	20608	5920	12709	3370.37	5409.52	4997.13	7.29	3528.23	5662.89	5231.19	6.49
HP23	1.05	1.00	8501	181	5632	8470.39	181	5632	1648.62	1646.87	1682.25	9.55	1736.95	2314.06	1772.38	3.91
HP34	1.05	1.04	8539	545	6920	8539	545	6920	1663.77	2414.44	2045.02	7.86	1753.1	2544.08	2154.82	3.90
HP22	1.07	1.05	8925	644	6608	8925	644	6507.42	1996.59	2176.24	1792.74	6.14	2144.12	2590.11	1925.21	4.16
HP33	1.09	1.00	5623	273	4087	5623	151.4	4011.11	1462.39	1257.31	910.45	8.56	1589.91	1366.95	1003.23	5.55
HP25	1.13	1.08	17287	2327	12806	17287	2327	12806	3340.42	4055.14	3567.89	8.42	3771.94	5185.44	4028.79	7.36
HP27	1.21	1.06	15969	1360	11793	15969	1360	11793	2406.49	3182.76	3351.36	7.84	2900.29	3835.84	4039.04	4.12
HP28	1.28	1.02	8956	571	6124	8956	571	6124	1146.75	1902.48	1731.84	8.09	1523.75	2428.48	2210.66	6.34
HP37	1.58	1.00	29589	5545	15990	23115.69	5545	15990	2578.66	3261.6	4483.67	8.22	4063.03	5674.92	7064.63	5.22
NON-FRONTIER HOSPITALS																
TOTAL			250458	52471	187109	243954.1	47331.81	161723.5	41769.12	50428.01	51869.17		46920.02	62805.84	59741.25	
AVERAGE	1.11	1.06										7.93				5.79
No. of less efficient hospitals			21													
% less efficient hospitals			56.76													

Notes : * indicates hospital is the optimal size defined by Constant Returns to Scale

a the inverse of this was used as an output in the model

Slacks are derived by the product of the efficiency score and the actual input or output, subtracted from the target for the input or output.

Annex 5.1.5: Model 5 – Small rural hospitals

	Efficiency score	Scale score	Actual inputs		Input targets		Actual outputs				Output targets			
HOSPITAL	*PHI (VRS)*	*SE*	*Total salaries ($'000)*	*Non-salary costs ($'000)*	*Total salaries ($'000)*	*Non-salary costs ($'000)*	*WEIS1*	*WEIS2*	*WEIS3*	*Unplanned re-admission rate*[a]	*WEIS1*	*WEIS2*	*WEIS3*	*Unplanned re-admission rate*[a]
HP39	1	1.00 *	1169.0	1388	1169	1388	192.21	530.17	724.29	1.32	192.21	530.17	724.29	1.32
HP41	1	2.33	2878.0	1743	2878	1743	589.63	907.89	733.33	6.47	589.63	907.89	733.33	6.47
HP43	1	1.12	817.0	535	817	535	407.77	414.09	459.04	12.64	407.77	414.09	459.04	12.64
HP44	1	3.13	5022.0	2654	5022	2654	1182.27	995.63	886.83	9.03	1182.27	995.63	886.83	9.03
HP46	1	2.63	3007.0	1823	3007	1823	837.29	825.09	645.54	6.55	837.29	825.09	645.54	6.55
HP52	1	3.13	4798.0	2920	4798	2920	1264.75	1121.69	741.20	9.34	1264.75	1121.69	741.2	9.34
HP66	1	1.00 *	165.0	1968	165	1968	338.50	159.09	117.97	13.04	338.5	159.09	117.97	13.04
HP67	1	1.00 *	422.0	268	422	268	381.34	321.74	264.98	8.76	381.34	321.74	264.98	8.76
HP89	1	1.00 *	147.0	76	147	76	120.55	44.67	39.84	11.96	120.55	44.67	39.84	11.96
HP98	1	1.00 *	693.0	301	693	301	439.95	377.51	331.13	12.99	439.95	377.51	331.13	12.99
HP101	1	1.00 *	191.0	97	191	97	134.14	65.55	79.56	15.13	134.14	65.55	79.56	15.13
HP102	1	1.20	408.0	224	408	224	115.11	55.31	30.24	5.14	115.11	55.31	30.24	5.14
HP104	1	2.08	2997.0	1643	2997	1643	621.72	887.73	801.96	8.11	621.72	887.73	801.96	8.11
HP55	1.01	3.45	4944.0	2984	4941.7	2749.35	903.60	1027.98	824.33	9.66	1211.84	1040.82	834.63	9.14
HP56	1.02	3.13	4212.0	2750	4212	2278.6	696.43	918.69	837.80	7.10	961.42	941	852.84	6.98
HP95	1.03	2.63	2910.0	2052	2910	1760.34	733.66	714.10	682.57	7.59	754.1	734	701.59	4.48
HP40	1.04	2.78	3795.0	2071	3721.51	2071	875.23	873.46	739.83	8.38	906.65	904.82	766.39	8.08
HP92	1.04	2.56	1737.0	1254	1737	1117.98	598.46	452.37	370.54	6.33	619.78	567.84	438.14	6.12
HP38	1.05	1.05	1053.0	1139	1006.21	1139	156.39	398.55	362.44	1.64	175.72	428.59	575.82	1.56
HP42	1.05	3.23	4189.0	2583	4189	2449.27	959.87	913.57	769.19	9.83	1008.46	959.81	808.12	5.00
HP86	1.1	2.22	1578.0	880	1578	880	528.49	384.14	281.65	8.32	580.47	525.82	456.57	7.58
HP48	1.11	2.78	4157.0	2223	3771.93	2223	872.49	873.23	569.76	10.13	965.5	966.32	718.38	9.15
HP60	1.12	1.28	344.0	191	285.64	191	104.16	39.18	63.17	7.23	122.28	72.37	70.61	6.47
HP58	1.13	2.13	1784.0	1039	1784	1039	539.58	435.03	431.47	7.35	611.18	566.06	488.72	6.49
HP78	1.18	1.89	1248.0	654	1152.07	654	406.03	341.74	267.95	7.11	477.52	453.64	431.54	6.04
HP47	1.19	1.82	1519.0	970	1519	970	356.29	491.96	444.10	5.54	451.21	583.58	533.26	4.67
HP94	1.19	1.85	972.0	815	972	610.8	409.13	296.46	212.45	9.73	487.54	423.3	330.2	8.16

Annex 5.1.5: Model 5 - Small rural hospitals (continued)

	Efficiency score	Scale score	Actual inputs		Input targets		Actual outputs				Output targets			
HOSPITAL	*PHI (VRS)*	*SE*	*Total salaries ($'000)*	*Non-salary costs ($'000)*	*Total salaries ($'000)*	*Non-salary costs ($'000)*	*WEIS1*	*WEIS2*	*WEIS3*	*Unplanned re-admission rate*[a]	*WEIS1*	*WEIS2*	*WEIS3*	*Unplanned re-admission rate*[a]
HP63	1.2	1.96	1363.0	783	1363	783	479.42	371.82	306.31	12.53	573.15	496.29	384.11	10.48
HP74	1.21	1.32	483.0	305	414.4	305	120.69	66.38	101.18	5.65	146.51	122.68	122.83	4.66
HP76	1.21	1.16	267.0	115	202.29	115	71.98	45.12	39.54	11.11	123.41	54.39	47.67	9.22
HP83	1.23	1.96	1385.0	873	1385	805.02	397.50	436.37	316.71	8.67	490.2	538.14	467.97	7.03
HP100	1.27	1.35	1369.0	1114	1369	1114	225.94	433.02	499.96	3.26	351.21	551.84	637.15	2.56
HP72	1.34	2.08	1435.0	1300	1435	907.12	391.38	392.16	281.84	6.71	525.13	526.18	471.4	5.00
HP99	1.34	1.49	487.0	226	394.98	226	137.96	39.69	25.44	8.04	184.75	123.97	91.95	6.00
HP75	1.35	1.05	371.0	167	262.95	167	84.31	90.74	97.70	12.77	215.43	152.35	132.21	9.43
HP59	1.37	1.75	1378.0	933	1378	924.72	314.25	402.98	358.04	11.86	431.58	553.44	491.72	4.90
HP49	1.38	2.38	2459.0	1296	2183.43	1296	508.28	477.76	204.86	10.37	702.55	660.36	522.73	7.51
HP62	1.38	1.69	947.0	445	611.25	445	268.73	170.98	127.37	6.87	369.8	368.87	352.61	5.00
HP85	1.39	1.28	471.0	235	398.33	235	170.85	109.83	61.49	9.19	237.14	176.44	138.13	6.62
HP45	1.4	2.22	1626.0	964	1626	964	435.45	371.31	302.99	10.86	607.54	549.21	436.89	7.78
HP71	1.4	1.61	873.0	505	873	505	331.53	154.61	165.56	12.24	465.52	408.48	338.21	8.72
HP87	1.41	1.28	293.0	146	229.41	146	56.82	30.87	42.28	11.19	121.79	62.13	59.75	7.92
HP64	1.43	1.67	1305.0	675	1269.05	675	356.46	347.33	274.72	13.68	510.01	496.95	432.2	9.56
HP91	1.46	1.96	1529.0	839	1305.96	839	318.06	306.83	251.07	5.83	463.61	485.98	491.93	4.00
HP103	1.46	1.41	550.0	290	441.25	290	203.02	123.53	95.56	9.20	295.96	248.77	209.33	6.31
HP80	1.51	1.61	1166.0	763	1114.86	763	277.96	316.31	248.89	5.90	419.55	477.43	478.26	3.91
HP84	1.52	1.61	1228.0	733	1228	733	326.56	294.25	300.10	11.81	496.67	476.1	456.42	7.76
HP82	1.53	1.03	217.0	89	167.57	89	80.13	36.19	16.78	17.15	129.98	55.34	47.79	11.21
HP105	1.53	1.75	1012.0	482	778.32	482	261.79	174.83	128.81	8.14	401.05	403.46	396.05	5.32
HP73	1.55	1.06	411.0	173	286.95	173	160.33	74.12	82.39	15.34	248.04	180.45	149.8	9.91
HP79	1.56	1.09	500.0	288	402.64	288	135.71	132.13	154.17	10.87	312.71	271.75	240.6	6.96
HP51	1.59	1.79	890.0	568	890	568	276.28	113.36	127.27	9.37	439.41	418.14	374.68	5.89
HP53	1.59	1.54	858.0	438	808.17	438	270.87	155.90	103.27	11.43	430.11	401.56	375.74	7.20
HP50	1.6	1.75	1106.0	490	890.07	490	275.51	89.17	79.36	10.99	441.18	414.02	389.25	6.86

Annex 5.1.5: Model 5 - Small rural hospitals (continued)

	Efficiency score	Scale score	Actual inputs		Input targets		Actual outputs				Output targets			
HOSPITAL	PHI (VRS)	SE	Total salaries ($'000)	Non-salary costs ($'000)	Total salaries ($'000)	Non-salary costs ($'000)	WEIS1	WEIS2	WEIS3	Unplanned re-admission rate[a]	WEIS1	WEIS2	WEIS3	Unplanned re-admission rate[a]
HP88	1.64	1.96	1069.0	560	954.01	560	267.18	139.98	119.72	9.36	438.13	426.76	411.59	5.71
HP93	1.64	1.33	606.0	294	443.96	294	179.26	123.56	67.25	10.20	294.57	248.84	210.41	6.21
HP57	1.71	2.08	1174.0	743	1174	743	297.75	129.24	114.41	11.27	509.38	465.8	382.34	6.59
HP54	1.75	2.70	1814.0	952	1628.84	952	311.58	121.62	129.48	8.68	543.83	525.69	508.32	4.97
HP96	1.75	1.67	663.0	378	497.21	378	188.17	72.58	88.97	9.21	329.86	311.16	284.04	5.25
HP97	1.8	1.39	1000.0	529	812.32	529	189.32	229.40	174.22	8.53	393.52	413.18	413.43	4.73
HP81	1.84	1.01	247.0	102	187.46	102	77.77	33.38	40.06	22.22	152.41	80.01	73.82	12.06
HP90	1.88	1.10	709.0	286	541.95	286	216.43	98.85	91.91	18.73	406.13	346.61	295.17	9.98
HP61	1.92	1.37	636.0	322	460.07	322	174.61	118.82	86.00	12.05	335.58	297.21	258.11	6.27
HP106	1.94	1.20	717.0	537	717	537	186.51	203.79	191.94	10.33	372.15	396.01	372.98	5.00
HP68	2	1.59	1001.0	487	1000.86	487	242.37	101.73	87.18	21.14	485.32	423.38	378.13	10.56
HP69	2.04	1.08	534.0	291	438.16	291	178.18	141.65	125.51	15.29	362.76	312.15	262.78	7.51
HP65	2.1	1.16	829.0	439	605.38	439	162.48	169.48	166.73	10.64	370.33	367.38	349.76	5.07
HP77	2.24	1.67	591.0	350	482.45	350	118.46	53.51	46.08	11.11	264.82	240.11	217.3	4.97
HP70	2.28	1.37	298.0	158	276.59	158	79.15	21.22	15.29	19.31	180.79	112.74	90.63	8.45
NON-FRONTIER HOSPITALS														
TOTAL			67365	40284	62767.5	38577.85	17043.20	14648.88	12371.33		23695.40	22768.90	20418.37	
AVERAGE	1.47	1.59								10.28				6.76
No of less efficient hospitals			56											
% less efficient hospitals			81.16											

Notes : * indicates hospital is the optimal size defined by Constant Returns to Scale

a the inverse of this was used as an output in the model

Slacks are derived by the product of the efficiency score and the actual input or output, subtracted from the target for the input or output.

5.2 Assessment of the performance of oral health services for Queensland school students

5.2.1 Introduction [11]

The Queensland Treasury is undertaking pilot studies in the Queensland public sector to apply DEA. DEA is particularly useful to public sector managers because it does not require inputs or outputs to be priced.[12]

The first study, in conjunction with Queensland Health, applied DEA to determine the relative performance of units providing oral health services to Queensland students from 1992-93 to 1994-95. Oral (or dental) health services are administered through thirteen geographical regions and undertaken in fixed and mobile dental clinics which visit each school at least once a year. The aim is to examine and treat each child to achieve acceptable oral health.

Data requirements

DEA measures the efficiency of service providers relative to those included in the sample only, so more observations will lead to better results usually. When there are few observations, the service providers being compared are more likely to be unique in the combinations of inputs used and outputs produced, and the model will determine a larger number of the providers as efficient. Increasing the number of inputs or outputs in the analysis exacerbates this problem because there is more potential again for providers to be unique within the sample.

There is no strict minimum number of observations required to undertake a DEA but a general rule for the minimum sample size is the sum of the number of inputs and the number of outputs multiplied by three. For example, a study such as oral health services, which has five inputs and one output, would require a minimum of eighteen observations: that is five inputs plus one output, multiplied by three. Relative efficiency scores tend to decrease as the sample size increases, improving the explanatory power of the model.

---

[11] By Patrizia Santin-Dore and Jennifer Pallesen of the Economics Division of the Queensland Treasury. The assistance of Steve Shackcloth and Ian Proud from Queensland Health in providing data and feedback on the model’s results is particularly appreciated.

[12] For a brief overview of the background, theory and application of DEA see Santin-Dore, P. and Pallesen, J. 1995, ‘Data envelopment analysis: an overview’, *Queensland Economic Review*, September Quarter, pp. 34–37.

DEA can be applied:

- *in cross-section*, comparing a number of organisations at one point in time;
- *as a time-series*, measuring the performance of a particular organisation over time; or
- *as panel data*, combining cross-section and time–series data, (that is, comparing a number of organisations over time).

Using panel data is a good way to increase the sample size. However, if there has been a significant change in technology over the sample period, it is difficult to assess whether increases in productivity reflected in rising average efficiency scores each year are a result of improvements in technical efficiency or technological change. Expenditure data also needs to be deflated by an appropriate price index.

5.2.2 Model specification for oral health services

For the oral health services study, annual data for the three-year period 1992-93 to 1994-95 was provided by Queensland Health on the thirteen Queensland regions, giving a sample of 39 units.[13]

Readily available data on oral health services are listed in Table 5.2.1. There were no significant changes in technology over the study period.

Table 5.2.1: Oral health services data available

Variable	*Units*
Inputs	
Dental officer days	Number
Total expenditure on dental officers	$
Dental therapist days	Number
Total expenditure on dental therapists	$
Dental assistant days	Number
Total expenditure on dental assistants	$
Total expenditure	$
Labour related costs	$
Non-labour related costs (= total expenditure less labour related costs)	$
Student enrolments	Number
Output	
Treatments completed	Number

[13] Panel data were used in this study because this offered the largest possible sample size of thirty-nine observations.

One of the most important steps in undertaking a DEA study is choosing the inputs and outputs to be used in the model. The inputs and outputs must relate to the objectives of the organisation, be consistent across organisations, and be quantifiable.

The objectives or desired outcomes of oral health services were to provide dental treatment to as many students as required services and to undertake preventative care.[14] The only data available on oral health services for measuring output were the number of treatments completed.[15] This sufficiently measures the first objective. The second objective of preventative care was not assessed because it requires a measure of service quality which was not available. The treatments completed in each region vary in complexity, time and resources used. However, because the different types of treatments completed in each region were not recorded, all treatments were regarded as equal by the model.

The inputs used to provide oral health services are labour and capital.[16] The labour inputs were divided into the number of dental officer days, dental therapist days and dental assistant days. Labour inputs were measured in physical quantities rather than dollars because wage rates vary between regions, and between dental officers, therapists and assistants. If salary expenditures had been used, then differences in expenditure would reflect not only the physical quantity of labour used to perform any given service, but also the prevailing wage rates.

For example, if dentists in city and country regions spent one hour to treat a patient for a filling, then the measure of each of their physical products in each case would be equal to one. However, if cost data were used and the country dentist's wages were lower, then the productivity of the city dentist would appear to be lower mistakenly.

Capital input was measured by non-labour related costs which were calculated by subtracting labour related costs from total expenditure. Queensland Health determined that this was the only way of measuring capital with the available data. (Also, when using dollars over a number of years, the data needs to be deflated using an appropriate price index.)

---

14 Over the study period, oral health services were provided to: preschool to year 7 in 1992-93; preschool to year 8 in 1993-94; and preschool to year 9 in 1994-95.

15 Treatments are any dental procedures performed on patients.

16 Capital is used to refer to all non-labour related costs.

The number of student enrolments in each region was also included to account for differences in potential demand.

The model was run in two formats:

- input minimisation, holding output constant and determining the minimum level of inputs necessary to achieve that level of output; and
- output maximisation, holding inputs constant and determining the maximum output that can be produced for that given level of inputs.

For each of these formats, the model was run with the assumption of constant returns to scale initially (that is, output increases in equal proportion to an increase in inputs, for example, a 10 per cent increase in inputs results in a 10 per cent increase in output). By holding returns to scale constant, it is assumed that all regions are operating at a scale appropriate to their situation.

This assumption was then relaxed to allow for variable returns to scale (that is an increase in inputs can result in a greater or lesser increase in output). Under this assumption, the model's efficiency scores are adjusted to remove differences resulting from operating at a less efficient scale, with any remaining inefficiency attributable to other factors.

Variations of the model specification

The model was run a number of times using different combinations of inputs and regions with the same output. This allowed the assumptions of constant returns to scale and variable returns to scale and their impact on the results to be tested and compared.

In consultation with Queensland Health, it was determined that the most appropriate variation was the one which excluded the two smallest regions (because they were deemed too small to be directly comparable with the other regions) and total enrolments (because this was not an input over which service providers had direct control).

Therefore, the preferred variation of the model was specified as follows: inputs were the number of dental officer days, dental therapist days and dental assistant days, and non-labour related costs; and output was the number of treatments completed. Given the exclusion of the two smallest regions, the sample size was reduced to thirty-three units. The details below focus on this specification.

The model was then run in input and output maximisation mode with constant returns to scale and variable returns to scale applied in each case, giving four sets of results.

5.2.3 Results and conclusions

Technical efficiency scores

The technical efficiency scores indicate which of the regions are deemed to be efficient (those given a score of 100) and which are deemed to be less efficient relative to those that are efficient (those with scores of less than 100). The lower the score, the less efficient the region rates relative to the most efficient.

It is important to note that the scores are relative — that is, those given a score of 100 are efficient relative to the rest of the regions in the model, but might not be operating efficiently by some absolute standard or standards elsewhere necessarily.

In summary, all runs showed that the vast majority of the regions in the sample appear to be performing reasonably well. The gap between the efficient and less efficient regions was relatively small, with most regions achieving technical efficiency scores of higher than 80 per cent.

Overall performance of regions

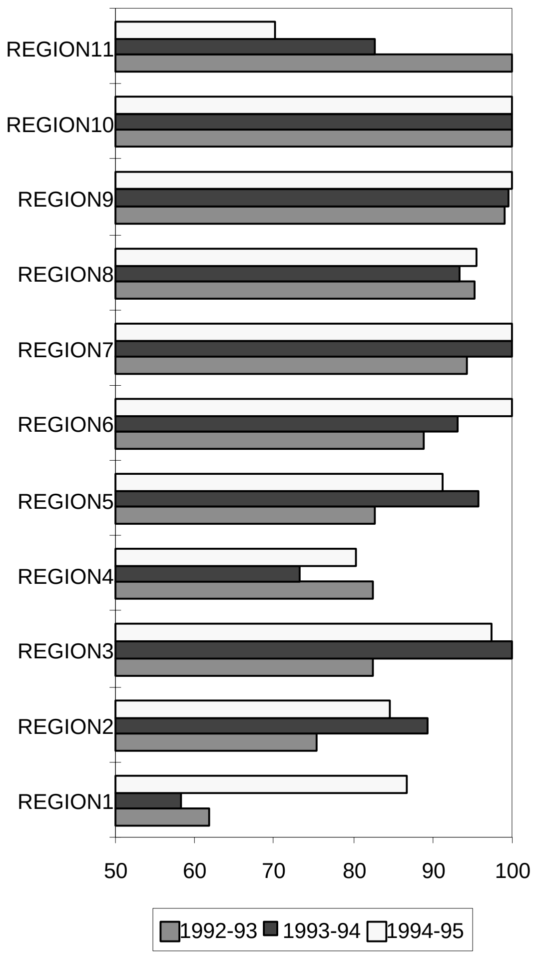
Results from the input minimisation model suggested that:

- assuming constant returns to scale, nine out of thirty-three units operated relatively efficiently, that is: region 10–93[17], region 11–93, region 3–94, region 7–94, region 10–94, region 6–95, region 7–95, region 9–95 and region 10-95;
- assuming variable returns to scale, inefficiency in seven units could be attributed to operating at an inappropriate scale of operation, that is, they were either too big or too small to operate efficiently. These units were region 4–93, region 7–93, region 9–93, region 5–94, region 9–94, region 4–95 and region 5–95; and
- the remaining seventeen units were technically less efficient.

Technical efficiency scores for the input minimising cases under constant returns to scale and variable returns to scale are depicted graphically in Figures 5.2.1 and 5.2.2, respectively.

[17] Region 10–93 means region 10 in 1992–93. Similarly, region 10–94 means region 10 in 1993–94 and region 10–95 means region 10 in 1994–95.

Figure 5.2.1: Technical efficiency scores for input minimisation under constant returns to scale



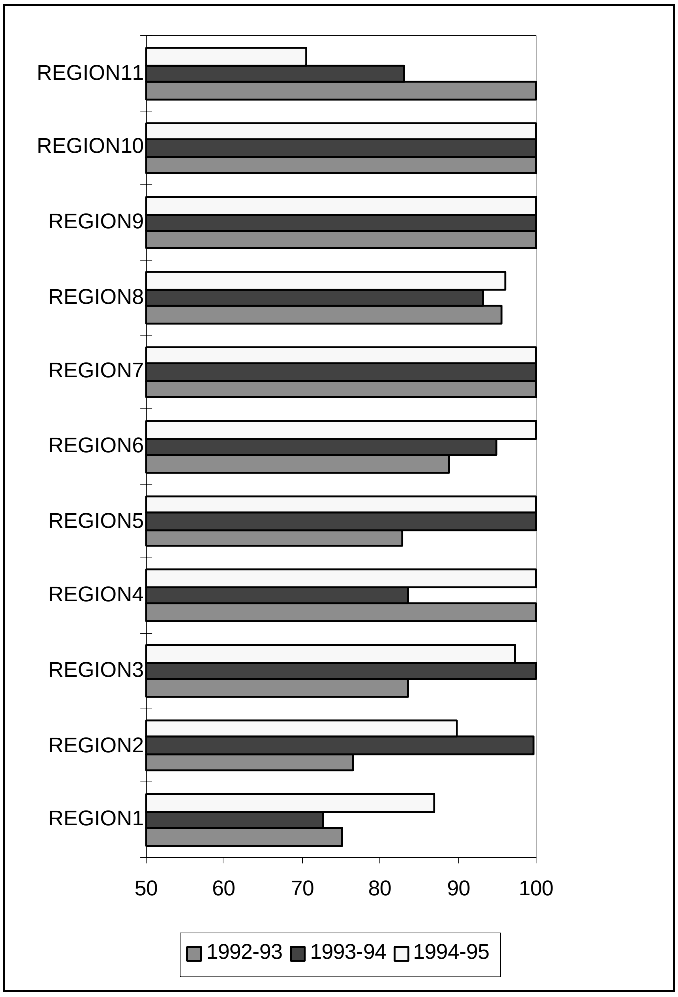
Under variable returns to scale (Figure 5.2.2), regions 1, 2 and 8 seemed to consistently perform less efficiently over the three years. Regions 3, 4, 5 and 6 improved their relative performance over the period, while regions 7, 9 and 10 seemed to be efficient in all three years.

However, the performance of region 11 deteriorated from a 100 per cent ranking in 1992-93 to the lowest ranking of 70 per cent in 1994-95. This reflected:

- a significant decline in the number of treatments provided, from over 19 000 in 1992-93 to just over 14 000 in 1994-95 (partly a result of a fall in the number of enrolments for the region); and

- a large increase in non-labour related costs, from around $76 000 in 1992-93 to over $416 000 in 1994-95. There is evidence that some items of non-labour related costs not included in 1992-93 were included in following years, which may explain the large increase.

Figure 5.2.2: Technical efficiency scores for input minimisation under variable returns to scale



Changing the orientation of the model to one of output maximisation had little effect on the technical efficiency scores and almost no effect on the rankings of the regions.

Peers and targets

DEA can also suggest peers and target input and output levels for each region. Peer regions are those which have been ranked as efficient by the model, and which a less efficient region may seek to use as a guide for improving its performance. In calculating targets, the actual level of input used is compared with the target input level calculated by the model, along with the percentage improvement needed to achieve the target. For example, the model suggested region 10–93 and region 7–95 as peers for region 1, which performed less efficiently over the three years. Region 1 can look at the input and output levels of these peers to gain insights into how it can improve its performance.

5.2.4 Conclusions

In summary, while the analysis does not test Queensland providers against an external standard, the vast majority of the units in the sample appear to be providing oral health services at better than 80 per cent efficiency. Alterations to the combination of inputs used did not significantly affect the results of the model, although the initial inclusion of the two smallest regions did result in these being determined as more efficient than the other regions, distorting the overall results.

For the preferred model specification, the gap between the efficient and less efficient units was relatively small. Regions generally showed improvement over the three–year period of the study.

Improving the performance of government service providers such as oral health services should not be based on efficiency alone. A government service provider might increase its efficiency by sacrificing the effectiveness of its service, so it is important to develop effectiveness indicators as well.

Forward — the NSW case studies

NSW Treasury and major budget sector agencies are beginning to use DEA to help establish benchmarks to improve the efficiency of government service provision.

NSW Treasury has completed a number of DEA studies on the technical efficiency of NSW police patrols (local police districts), corrective services, and motor registries. It has also completed a pilot study on the technical efficiency of local courts in NSW. Studies have commenced to provide insights into the technical efficiency of NSW hospitals and metropolitan fire brigades.

The DEA studies presented in this information paper were prepared by various members of a Treasury team that is developing performance indicators for the budget sector.[18]

The views expressed in these studies are the authors' and do not reflect those of NSW Treasury, the participating government agencies or the NSW Government necessarily.

5.3 Technical efficiency of corrective services in NSW

5.3.1 Introduction

The number of inmates in Australia has grown steadily over recent years, reflected in higher rates of imprisonment. Governments are ensuring that inmates serve longer sentences by abolishing prison remission (SCRCSSP 1995). In 1988, the NSW Government increased the sentences for crimes and abolished prison remission. Consequently, the daily average number of inmates in prisons (correction centres) rose from 4124 in 1987-88 to 6279 in 1994-95. The information presented in Table 5.3.1 indicates that the rate of imprisonment in NSW increased from 101.9 inmates per 100 000 adults in 1988-89 to 135.9 inmates per 100 000 in 1994-95. Only the Northern Territory and Western Australia have a higher rate of imprisonment. NSW Government expenditure on

[18] The team has included Roger Carrington, Nara Puthucheary, Deirdre Rose and Suthathip Yaisawarng. John Pierce, Secretary of NSW Treasury, supervised the project while he was the Executive Director of State Economic Revenue Strategy and Policy. NSW Treasury would like to thank the NSW Police Service, the Department of Corrective Services, and the Roads and Traffic Authority for their assistance in preparing the studies, and Tim Coelli (University of New England) for his useful suggestions on earlier drafts of the studies.

prisons and corrective services increased from $239 million to $367 million over 1988-89 to 1994-95 (ABS 1995a) — a real increase of 30 per cent.[19]

Table 5.3.1: Estimated total prisoners per 100 000 adults, 1988-89 to 1994-95

	NSW [1]	*Vic*	*Qld*	*SA*	*WA*	*Tas*	*NT*	*ACT* [2]
1988-89	101.9	68.1	116.0	77.9	135.5	76.9	363.0	10.6
1989-90	115.0	69.8	106.6	81.5	138.9	70.1	351.3	11.1
1990-91	129.3	69.1	101.5	87.2	152.3	70.8	394.5	11.1
1991-92	134.2	66.9	94.9	97.2	155.3	76.1	397.8	9.4
1992-93	135.9	66.8	89.0	101.5	150.0	74.5	373.4	7.5
1993-94	137.9	73.9	94.6	108.7	165.1	71.9	384.6	8.6
1994-95	135.9	71.8	109.2	118.6	164.8	74.2	393.9	8.6

1 NSW figures exclude periodic detainees.
2 ACT figures are only remandees. ACT sentenced prisoners are included in NSW figures.
Source: Steering Committee for the Review of Commonwealth/State Service Provision (1995).

The NSW Government has developed several policies to reduce the cost and improve the effectiveness of corrective services. Inmates are held in the lowest possible level of security where appropriate, so over 50 per cent of inmates are held in minimum security correction centres. Security posts in maximum security correction centres are being replaced with electronic security equipment, and alternatives to incarceration, such as community service orders, are now available to the courts to deal with fine defaulters. A home detention scheme commenced in 1996-97 for suitable minimum security inmates; it will impose liberty restrictions similar to incarceration. Several personal development programs such as vocational education and training, drug and alcohol rehabilitation, and work release programs help inmates prepare for their return to the community.

This paper reports the progress in measuring the technical efficiency of corrective services in NSW using DEA (Lovell 1993). Farrell (1957) suggested that there are two aspects to overall economic efficiency — technical efficiency and allocative efficiency. Technical efficiency describes the physical relationship between inputs and outputs; it reflects the ability of an organisation to produce maximum outputs given inputs, or to produce certain outputs with least inputs. Allocative efficiency, on the other hand, measures the optimal mix of inputs (or outputs) for an organisation given observed prices.

---

19 The implicit price deflator for Government final consumption expenditure in New South Wales in 1989-90 (ABS 1995b) is used to derive this figure.

The study does not consider the effectiveness of correction centres to reduce recidivism. Recidivism not only depends on the attitudes and skills acquired by inmates while in the correctional system (which may be of limited influence for inmates serving short sentences), but is also influenced by factors outside the control of the corrective system such as family and community support (for example, access to public housing and other social services, employment, and vocational education and training).

Section 5.3.2 explains the production of corrective services and Section 5.3.3 presents the initial results of the technical efficiency of corrective services. Section 5.3.4 contains conclusions about the technical efficiency of corrective services, and outlines further initiatives to improve the analysis.

5.3.2 The production of corrective services

The Department of Corrective Services carries out the orders of the courts by incarcerating inmates until they are lawfully discharged. It aims to manage inmates in a humane, safe and secure environment, and provide personal development programs for inmates that focus on the causes of their crime and help them return to the community as law-abiding and productive citizens. These broad objectives of corrective services provide a focus to specify the outputs of corrective services.

In NSW, correction centres are classified as maximum, medium or minimum security. As inmates enter the corrective system, they are classified to the lowest appropriate security rating which is consistent with the Department’s responsibility to protect the community. This study focuses on minimum security correction centres, which account for the majority of corrective facilities in NSW, and hold most of the state’s inmates.

Other categories of correction centres are not included in the study because they use different technology and resources to manage inmates. For example, maximum security correction centres have complex electronic surveillance systems and some of the older facilities have watch towers staffed by armed guards. Moreover, a ready-armed response is available to deal with serious incidents, such as escapes by inmates. Medium security correction centres have a ready-armed response only during the day. In both maximum security and medium security correction centres, inmates are held in secure cells and enclosed in a secure perimeter to prevent escape. By contrast, inmates in minimum security correction centres face limited physical barriers and may require little supervision by custodial officers.

One measure of confined inmates in a minimum security correction centre is the average daily inmate population. However, such a centre can hold several categories of inmates who require different inputs for their incarceration. The various categories are:

- inmates who are held in secure cells overnight but face limited physical barriers during the day;
- inmates who are not restrained by physical barriers but are supervised by staff when entering the community. Their cells are typically not secured, although accommodation blocks are locked overnight; and
- inmates who are allowed to enter the community without supervision and whose accommodation is less secured compared with the accommodation for the other categories of inmates.

The last category is cheaper to manage than the first category of inmates for several reasons.[20] They have conditional leave of absence to participate in work release, education or weekend release programs, allowing them to enter the community without custodial supervision. Inmates participating in work release programs pay board which further reduces the cost of their incarceration. Moreover, these inmates are held in less secured accommodation overnight, which costs less to construct than does secured accommodation. To account for this, we include the average daily number of inmates that are eligible for conditional leave of absence without custodial supervision and the average daily number of other inmates in the analysis.

A third measure of confinement — the number of inmates received by a correction centre — is considered in the analysis because it affects the operating costs of the correction centre. New receptions and discharges require additional resources to provide reception and induction programs, administrative processing and pre-release programs. Moreover, new inmates, even if transferred from other centres, require additional supervision and support while adapting to new circumstances.

The Department of Corrective Services provides a range of personal development programs, such as vocational education and training and drug and alcohol rehabilitation, to help prepare inmates for their return to the community. The time that inmates spend in personal development programs measures the quantity of personal development training that inmates receive; it does not reflect how well these programs prepare the inmates for their return to the

---

[20] There are some differences in the cost of managing the second category of inmates and the other categories, but not as large as the difference in the cost of managing secured inmates and inmates eligible for conditional leave of absence.

community. Information is available only on the time that inmates spend in education programs or industrial activities, such as saw milling and farming, which provide them with additional vocational skills and some income. To reduce the number of variables in the analysis, the time that inmates spend in these activities is combined to measure the personal development programs delivered by the Department of Corrective Services. The motivation for reducing the number of variables in the analysis is explained in Section 5.3.3.

A minimum security correction centre uses a variety of inputs to manage inmates, including staff, capital, food, clothing, and other consumer goods and services, such as power and water. The operation of a correction centre requires several types of labour such as custodial officers, industrial officers (who teach vocational skills), medical and other professional officers, and maintenance and clerical staff. Custodial officers and industrial officers are the largest group of employees. They oversee security in the correction centre and deliver the main development program, inmate employment. Therefore, for the purposes of this study, custodial and industrial officers are classified as custodial officers and are measured by full-time equivalent staff numbers. Medical staff are excluded from the study because the health budget for correction centres is provided by the NSW Department of Health. There is no information on the other labour used in minimum security correction centres.

Capital is measured by the design capacity for each correction centre, which is the number of inmates that a correction centre can hold. This is usually determined by the number of beds in a correction centre. Other recurrent expenditure (that is, less wages and salaries) measures inputs such as food, clothing and other consumer goods and services.

In summary, the outputs included in this study on corrective services were the average daily number of inmates eligible for conditional leave of absence, the average daily number of other inmates, the number of inmate receptions, and the number of hours that inmates spend in personal development programs. The inputs included the number of beds in a correction centre, full-time equivalent custodial staff numbers, and other recurrent expenditure.

The study has similar features to a DEA study on correctional institutions in New York State (Yaisawarng and Vogel 1992). That study developed several models of corrective services that have comparable inputs to those used in this study. However, it assumed that some inputs, like recurrent expenses, were adjustable while holding others, such as capital, constant. Further, the study focused on confinement as the output of correctional institutions. Yaisawarng and Vogel used either a single measure (the average daily inmate population) or multiple measures to reflect the different inputs required to manage violent inmates, less violent inmates, and people awaiting sentence or arraignment.

Ganley and Cubbin (1992) considered outputs other than confinement in their study of correctional institutions in the United Kingdom, but did not consider the personal development of inmates. In addition to outputs of confinement (that distinguish between sentenced and unsentenced inmates), they included a quality indicator for confinement (the degree of overcrowding), and attempted to measure the quality of the supervision provided by correctional officers by the number of serious offences committed by inmates. Ganley and Cubbin used labour costs and non-labour costs to measure inputs: they considered these costs could be adjusted.

5.3.3 Technical efficiency of corrective services

There are seventeen medium security and minimum security correction centres in NSW that incarcerate minimum security inmates. Six correction centres were excluded from the study. They included the correction centres which hold both medium security and minimum security inmates (because they use substantially different resources to manage the medium security inmates) and the minimum security correction facilities attached to medium security or maximum security centres (because it is not possible to isolate their inputs).

The remaining eleven minimum security correction centres were a relatively small sample to generate sensible DEA results if data for only one year were used in the analysis. Therefore, panel data for 1990-91 to 1994-95 were used to increase the sample. This was possible because the Department of Corrective Services argued that there was minimal change in the technology (which included inmate management) used in minimum security correction centres over this period. Most correction centres were evaluated five times, that is, once for each year. However, two correction centres were converted from medium security in 1993-94, so they were evaluated only for each year that they were a minimum security centre.

There is incomplete information for some correction centres on the time that inmates participate in education programs in a particular year. To overcome this problem, the proportion of total inmate hours to the number of months for which information is available was assumed to prevail over the year. An implicit price deflator for NSW Government consumption (ABS 1996) was used to convert the nominal other recurrent expenditure into real other recurrent expenditure.

DEA has several limitations. It is sensitive to outliers, and the omission of important services or inputs produces biased results. Further, DEA includes random occurrences in its measures of efficiency — for example, errors in the data. To reduce the risk of data errors influencing the results of the study, the data were screened for potential outliers using descriptive statistics for the services (outputs) and the inputs and also by identifying service–input ratios that were further than two-and-a-half standard deviations from their sample means. Potential outliers that were identified this way were referred to the Department of Corrective Services for comment. The Department subsequently confirmed that these observations were not a result of errors in the data. Table 5.3.2 presents descriptive statistics for each service and input variable in the sample.

Table 5.3.2: Descriptive statistics for minimum security correction centres[a]

	Mean	*Standard deviation*	*Minimum*	*Maximum*
Outputs				
Inmates eligible for conditional leave of absence (no.)	47.74	6.57	7	187
Other inmates (no.)	134.60	12.09	28	408
Receptions (no.)	704.32	90.09	111	2 386
Personal development programs (hours)	120 565	10 009.66	31 657.10	338 014
Inputs				
Beds (no.)	181.55	17.20	64	453
Custodial staff (full time equivalent no.)	65.67	5.30	30	150.50
Real other expenditure ($'000)	1 003.77	66.51	290.60	2 261.37

[a] Forty-seven observations.

The Department of Corrective services has little control over the number of inmates that it manages. It must incarcerate and help rehabilitate inmates with least inputs. Therefore, an input-oriented DEA model is used to determine the technical efficiency of minimum security correction centres. A similar approach is used by Yaisawarng and Vogel (1992) and Ganley and Cubbin (1992). Details on the method to calculate the technical efficiency of correction centres is presented in Appendix A and Lovell (1993).

Table 5.3.3: Summary statistics: efficiency of minimum security correction centres

	Technical efficiency	*Pure technical efficiency*	*Scale efficiency*
No. of observations	47	47	47
Mean	0.930	0.964	0.963
Standard deviation	0.018	0.012	0.009
Minimum value	0.593	0.666	0.757
Number of efficient correction centres	20	29	20

The results presented in Table 5.3.3 suggest that pure technical efficiency (managerial efficiency) and scale efficiency were equal sources of lower efficiency for correction centres.

Correction centres may be able to produce, on average, the same outputs with approximately 4 per cent fewer inputs. However, this result needs to be interpreted with care. Managerial efficiency is influenced by the sample and the number of outputs and inputs included in the study. If DEA cannot compare a correction centre with other correction centres, then it is deemed efficient by default, which tends to increase the average managerial efficiency score. This study had a relatively high proportion of correction centres that were apparently efficient by default (about 20 per cent of the managerial efficient correction centres) because it had a relatively small sample compared with the number of outputs and inputs used in the analysis. To overcome this problem, future analysis could include correction centres from other states to increase the sample, or alternatively the number of outputs or inputs used in the analysis could be reduced.

About 13 per cent of correction centres appeared to require larger reductions in inputs, compared with the average reduction in inputs, to become managerially efficient. The least efficient correction centre would appear to have to reduce its inputs by around 33 per cent. This correction centre was converted from a male facility to a female facility in 1994-95. Inmate numbers declined by about 40 per cent without a commensurate reduction in inputs. Before its conversion to a female facility, the centre appeared to be managerially efficient in 1990-91 and 1991-92, was above average managerial efficiency in 1992-93, and was marginally below average managerial efficiency in 1993-94.

The average apparent scale efficiency of correction centres was 96 per cent, which suggests they might be able to reduce inputs by a further 4 per cent to achieve optimal scale. The information presented in Table 5.3.3 suggests that about 43 per cent of the correction centres were scale efficient. About 60 per cent of the scale inefficient correction centres exhibited increasing returns to scale. The Department of Corrective Services might wish to examine the scope to combine these centres into larger centres that possess optimal scale. The remaining correction centres exhibited decreasing returns to scale and would appear to require a reduction in size to achieve optimal scale.

The managerial efficiency of a minimum security correction centre is overstated if, after an equi-proportionate decrease in all inputs, some excess inputs (or slacks) are still present. About 36 per cent of the correction centres had at least one such input. The information presented in Table 5.3.4 suggests that correction centres with particular excess inputs, on average, may be able to reduce the inputs further by 10.63 beds, 2.33 full-time equivalent staff and $145 620 in real other recurrent expenditure, and produce the same outputs. However, the excessive use of inputs was relatively minor compared with total inputs.

Table 5.3.4: Input slacks in correction centres

	Number of correction centres with slacks	*Mean*	*Total slacks as a percentage of total inputs*
Beds (no.)	7	10.63	1.00
Custodial staff (full-time equivalent)	5	2.33	less than 1.00
Real other expenditure ($'000)	13	145.62	4.01

5.3.4 Conclusion

The analysis suggests that minimum security correction centres, on average, might be able to produce the same outputs with 4 per cent less inputs. Moreover, if all correction centres were of optimal scale, they might be able to reduce inputs by a further 4 per cent. However, operational imperatives relating to centre location requirements and meeting the needs of specific offender groups need to be taken into consideration.

Care is required in interpreting the results because a relatively high number of correction centres were apparently efficient by default, which contributed to the high mean managerial efficiency of the correction centres. Further, there were deficiencies in some data, especially the information for inmate personal development.

The technical efficiency scores for individual correction centres, the associated information on peers, and the effectiveness indicators for corrective services developed by the Steering Committee for the Review of Commonwealth/State Service Provision (1995) provided the Department of Corrective Services with an objective framework to judge the extent to which, and speed with which, it might be able to improve the technical efficiency of minimum security correction centres. The Department discovered inefficiencies in some correction centres that it was unaware of before the DEA exercise. It will review these centres to discover the cause of the inefficiencies.

NSW Treasury will seek to include minimum security correction centres in other states in future analysis of the technical efficiency of corrective services in NSW. Increasing the sample might reduce the number of correction centres that are apparently efficient by default. Moreover, a larger sample might improve the benchmarks and targets for the efficient provision of corrective services. An alternative to setting better benchmarks and targets for corrective services is to seek further guidance from the Department of Corrective Services to reduce the number of outputs and inputs used in the analysis. However, this approach risks excluding important variables in the provision of corrective services, which would produce biased results.

Further work is required to improve some of the information used in this exercise. The Department of Corrective Services is aware of the deficiencies of its statistics on the time inmates spend in personal development programs and is examining ways to improve this information.

5.4 Performance measurement of police services in NSW

5.4.1 Introduction [21]

The purpose of this paper is to examine the technical efficiency of the Police Service in 1994-95, using a two-stage procedure. In the first stage, DEA is used to compute technical efficiency for all police patrols.[22] In the second stage, Tobit regression is used to analyse external factors or operating environments which might explain the variation in apparent technical efficiencies across

[21] This is an edited version of Roger Carrington, Nara Puthucheary, Deirdre Rose and Suthathip Yaisawarng 1997, ‘Performance Measurement in Government Service Provision – the Case of Police Services in NSW’, *Journal of Productivity Analysis* (forthcoming).

[22] A police patrol is an organisation unit which is responsible for providing services to a designated area within the community. The designated area is also referred to as the “patrol”.

police patrols. The results of this study can be used to assist the NSW Police Service in delivering better and efficient services to the community.

The motivation of this study originates from the introduction of comprehensive financial reforms by the NSW Government to help ensure government service provision, such as health, law and order and education, is efficient and effective. The main reform is contractual (or performance) budgeting, an approach whereby the Government enters into agreements with government service providers to purchase services that assist in achieving government policy objectives, rather than funding services according to historical expenditure patterns. Performance measurement is necessary to complement contractual budgeting to provide an incentive for government service providers to become more effective and efficient. According to Pierce (1995), performance indicators provide information that makes government service providers more accountable to Parliament. They also promote yardstick competition in the provision of government services that face little competition, acting as a powerful internal management tool to examine reasons for poor performance.

Section 5.4.2 discusses the operation of the NSW police patrols, Section 5.4.3 presents the empirical results, and Section 5.4.4 summarises the findings and demonstrates the use of DEA to improve the performance of a major government service provider.

5.4.2 Assessing the performance of the NSW police service

Law and order is a high priority of the NSW Government. It has sought to allay community concern over public safety by employing more police officers. The Government announced in the 1995-96 Budget that the number of police officers available for duty would increase by 650 over the following four years to 13 407.

Community perceptions on public safety are influenced by social, economic and institutional factors that are beyond the control of the NSW police service. Nevertheless, public safety is an outcome that the police service seeks to influence by providing several services to the community. Under the auspices of the Steering Committee, and in conjunction with other Australian state and federal police services, it has developed several objectives for its services:

- to protect, help and reassure the community;
- to prevent crime; and
- to enforce the law.

Effectiveness indicators for the objective of protecting, helping and reassuring the public include the number of hospitalisation and fatal road crashes, the percentage of vehicles stolen which were recovered in the same year, and the number of complaints about police behaviour. Crime rates reflect the police service's effectiveness in preventing crime. However, crime rates need to be interpreted with care because they are influenced by factors beyond the control of police. Success in bringing offenders to justice and the number of deaths in custody are examples of effectiveness indicators of law enforcement developed by the Steering Committee. The Steering Committee published information on the number of deaths in police custody only because there is a lack of comparable national information for the other indicators.

The Steering Committee developed a limited number of efficiency indicators for police services — for example, the average total police vehicle cost per kilometre. The unit cost of police vehicles in NSW declined from 35 cents per kilometre in 1992-93 to 31.5 cents per kilometre in 1994-95. However, the Steering Committee argued that there are several difficulties in developing a comprehensive suite of efficiency indicators for police services based on unit cost. Outputs for some police activities, especially crime prevention, are difficult to define, and police can deliver several services simultaneously, so it is difficult to isolate the inputs for each service, (for example, the public presence of police arresting an offender also reassures the community and prevents crime).

Effectiveness and efficiency indicators are important tools that can assist in improving the performance of police services, but this paper focuses on the efficiency of police services only. Efficiency scores were calculated for each police patrol using DEA. This technique provides a single measure of efficiency for each patrol. This efficiency score reflects the success of a patrol in producing a variety of services using a number of inputs. To a large extent, this overcomes the problem of allocating inputs to specific services (which is required for unit cost analysis). DEA identified the apparently best patrols by their ability to produce the greatest number of services with a given set of inputs, or to produce certain services with least inputs. Patrols received an efficiency score that was determined by their performance relative to the best patrols. The information on the set of best patrols (peers) that are compatible with a specific less efficient patrol is useful for the less efficient patrol to identify ways in which it might improve its efficiency.

It is important, although difficult, to specify completely and correctly the activities of units analysed. The omission of important variables produces biased results. A knowledge of police duties provides an insight into the services that a police patrol provides the community, and focuses attention on the development of an ideal model for measuring police efficiency. In the following two sub-sections, the inputs and outputs that should appear in a DEA model of the NSW police service are listed, and the available data that are the variables selected for use in the empirical analysis are discussed.

Production of police services

The NSW police service organises its command to deliver its services into four geographical regions of NSW that include twenty-five districts and 165 patrols, special agencies (such as the Drug Enforcement Agency) and regional support units (such as the Forensic Services Group). Most police services are delivered by police patrols.

A police patrol typically consists of a patrol commander, detectives, general duty police (which include beat police) and civilian employees. The patrol commander is responsible for the operations of a patrol. The commander provides leadership, develops the strategic plan for the patrol, and investigates complaints about police behaviour. General duty police conduct random breath testing of motor vehicle drivers, attend major car accidents, arrest people, issue traffic infringement notices, maintain lost property exhibits, handle missing persons inquiries, secure prisoners in custody, and respond to calls for assistance. Beat police gather intelligence on criminal activities, arrest people, control crowds at sporting events and demonstrations, obtain warrants and serve summons, maintain a public presence to prevent crime and reassure the public of their safety, and conduct follow-up interviews of victims of crime. Detectives investigate more serious criminal matters, arrest people, and attend to coronial matters. Civilians provide clerical support that includes answering telephone calls and sending radio messages. Highway patrols use the infrastructure of patrols — police stations, for example — in their endeavours to reduce the number and severity of road accidents. However, highway patrols are often a separate police unit to police patrols, so they are excluded from the analysis.

The NSW police service classifies the services of a police patrol into two broad categories: reactive policing, which covers law enforcement activities; and proactive policing, which covers activities that protect, help and reassure the public and prevent crime. Random breath testing and patrolling crime spots are examples of proactive policing.

If there were no data limitations, the DEA model of police patrols would cover all the outputs that a patrol produces to deliver services to the public. Given that there are several service activities in each category, a case mix index which reflects the quality of services would be created. Aggregate output indexes for reactive policing and proactive policing, which use appropriate case mix indexes as weights, would represent ideal output measures in the DEA model.

The specification of inputs is much more clear cut. Police patrols use several main inputs to provide services to the community: police officers,[23] civilian employees and capital equipment, such as police cars and computers.

The sample

This paper uses a sample of 163 police patrols in 1994-95.[24] The NSW police service has comprehensive data on outputs for reactive policing, but little information on proactive policing (which accounts for about 40 per cent of patrol work). It also has no information on the outputs of civilian employees. Further, the NSW police service does not have reliable information on the time that police spend on their activities. After discussions with the NSW police service, the major measured outputs of police patrols were included in the DEA study.

The reactive policing outputs of a patrol are responses to incidents, arrests, serving of summons and attendances at major car accidents. These were measured by the number of cases[25] and were included as output variables in the DEA study. However, there are several caveats associated with data for arrests credited to a patrol and a patrol's response to incidents.

The information on the arrests performed by a patrol includes arrests made by other NSW police agencies such as special operations groups and the Drug Enforcement Agency, and the Australian Federal Police. Moreover, arrests by

---

[23] The NSW police service cannot isolate the outputs and inputs for detectives and general duty police (which include beat police). It suggested that they can be combined into one category: namely, police officers.

[24] One Sydney patrol is excluded from the analysis because it is the central jail for offenders waiting to appear before a court. The NSW police service also decided to include a police water patrol into a nearby (regular) patrol for the purposes of this study.

[25] Previous studies (for example, Darrough and Heineke 1979; Gyimah-Brempong 1987, Gyapong and Gyimah-Brempong 1988, Levitt and Joyce 1987) used arrests or clearances as proxies for police outputs. This is only one aspect of the reactive policing of the NSW police service. Further, the police service rejected clearances as a meaningful measure of output of police patrols. A crime is cleared when police have sufficient evidence to lay a charge against those who committed the crime; they are not required to make an arrest or serve a summons to clear the crime. Moreover, it is possible for police patrols to increase their clearances merely by recording additional charges against the person or people that committed the crime. If clearances are used as a measure of output, then there is a risk that police patrols will focus on crimes that are easy to clear and ignore serious crimes that are harder to solve.

these police agencies to solve criminal activities in several patrols are credited to each patrol. The NSW police service cannot separate these arrests from the arrests affected by a patrol alone. Consequently, the outputs of some patrols are artificially inflated by these arrests. Wrong arrests also overstate the outputs of a patrol.

The NSW police service's information on incidents is limited to recorded offences. This information is likely to understate a patrol's response to incidents because it excludes those incidents for which police either decide that no further action is required or issue a warning to the offenders.

Output variables for proactive policing were difficult to obtain. The total kilometres travelled by police cars captures some aspects of proactive policing. It reflects a police presence in the community to reassure the public, and a visible police car prevents crime. However, it ignores the proactive policing that beat police do on foot in metropolitan patrols. Information is not available for alternate measures of proactive policing such as the number of random breath tests conducted by a patrol or the number of criminal intelligence reports filed by beat police. Darrough and Heineke (1979), Gyimah-Brempong (1987), and Gyapong and Gyimah-Brempong (1988) assumed that non-crime activities (such as traffic control and emergency first aid care) are related to the size of a community, and used the population of the community to measure these services. The NSW police service argued that the official population in a patrol does not accurately reflect the proactive policing provided by a patrol because the population of a patrol can swell considerably as people enter its jurisdiction for work or entertainment. Moreover, even if accurate figures on a population were available, it still must be unrealistically assumed that each police patrol provides a similar proportion of proactive policing relative to the other services it provides the community.

Similar to most existing studies, two types of labour input were used: number of police officers and the number of civilian employees as of 30 June 1995.[26] The number of employees assigned to a patrol included people on extended leave (for example, sick leave, long-service leave or seconded leave to other police units). Therefore, care is required when interpreting the results. A patrol may appear relatively less efficient because it had a higher proportion of its personnel on extended leave. Further, a patrol that consistently overworked its staff might appear more efficient compared with a similar patrol for which staff

[26] Actual number of hours worked is more desirable but this information was not available.

worked normal hours. Capital input was measured by the number of police cars.[27]

In summary, the DEA model of NSW police patrols included five output variables: the number of arrests, responses to offences recorded, serving of summons, attendances at major car accidents, and the kilometres travelled by police cars. The three inputs used were: the number of police officers, the number of civilian employees and the number of police cars. The sample included 163 police patrols for 1994-95. Table 5.4.1 presents descriptive statistics for each output and input in the sample.[28]

Although the efficiency scores obtained from solving DEA represent the ability of management to convert inputs into outputs at the current scale of operation, it is possible that some other external environmental factors beyond the control of the management may affect their measured efficiency. The study looked to determine which external factors had some influence upon variations in pure technical efficiency across police patrols and in which direction. A second-stage analysis was used to explain the variation in DEA technical efficiency scores from the first stage.[29] This used the Tobit procedure to estimate the relationship between pure technical efficiency scores and operating environmental factors unrelated to the inputs used in the DEA model.[30]

Specifically, the following model was estimated:

---

27 The number of computers or other equipment installed in patrols or the floor space of buildings occupied by a patrol could be included in the measure of capital if the information was available.

28 DEA is susceptible to outliers, so output–input ratios were computed for each patrol, and the value that exceeded two-and-a-half standard deviations from the sample mean was considered a potential outlier. Potential outliers were referred to the NSW police service who checked the data and confirmed there were no obvious measurement or reporting errors. Burgess and Wilson (1993) discussed the nature of outliers and their impact on DEA efficiency scores. Wilson (1995) suggested a way to detect influential observations, which is a computer–based technique and is appropriate when it is not possible to access the first-hand data or too costly to check all data points. This was not the case in this study.

29 The method differed from that used by Levitt and Joyce (1987) who directly included environmental variables in their DEA study of UK police authorities. Their method required an assumption on how each environmental variable affected efficiency. This assumption precluded the test of its impact. McCarty and Yaisawarng (1993) discussed advantages and disadvantages of these two approaches.

30 Given that the pure technical efficiency scores are censored from above at one, the ordinary least squares regression produces biased and inconsistent results (Judge et al., 1988).

$$TE^{VRS} = \alpha + Z\beta + u$$

where TE^{VRS} is a (J x 1) vector of pure technical efficiency scores for all J patrols, the scalar α and the ($R * 1$) vector β are unknown parameters to be estimated. Z is a ($J * R$) matrix of environmental factors and u is a ($J * 1$) vector of residuals.

The NSW police service suggested several environmental variables, or noncontrollable inputs, that might affect the measured efficiency of a patrol but that are beyond the control of management.[31] First, police observe that most offenders are young people aged 15 to 29 years. A patrol with a higher proportion of young people in its jurisdiction is likely to respond to more incidents compared with a similar patrol with a lower proportion of young people in its jurisdiction. Second, a patrol with a high proportion of public housing in its jurisdiction is likely to respond to more incidents than a similar patrol in an area with a lower proportion of public housing. Finally, country patrols usually cover a larger area than metropolitan patrols. They require additional cars and staff (above the level of resources justified for the services they provide to the community) to permit the NSW police service to provide an effective police presence in country areas which is comparable to the service it provides in metropolitan areas.

The proportion of young people in a patrol area and the proportion of government housing in a patrol area were derived from 1991 census data. A dummy variable was used to specify the location of a patrol, where a value of zero indicated a patrol was located in a metropolitan area and a value of one indicated a patrol was located in the country.

Patrols with a higher proportion of young people or a higher proportion of public housing in their area, or both, were expected to appear more efficient than similar patrols facing lower proportions of these socioeconomic conditions because they were relatively busy responding to more crime (that is, they had less idle time). Some of their additional work might not be reflected in measured outputs because some incidents they investigated warranted a warning to offenders only. Nevertheless, police patrols with a higher proportion of these environmental variables were expected to have higher measured outputs.

[31] Gyimah-Brempong (1989) used a sample of police departments in Florida, United States in 1982 and 1983 to study the impacts of environmental variables on the production of public safety. The author found that a higher proportion of non-whites in a police jurisdiction raised the cost of supplying police services to achieve a certain level of public safety, which is measured as the inverse of the crime rate of the community. A greater proportion of comparatively highly educated people in a locality reduced the cost of supplying police services to achieve a certain level of public safety in a locality.

Country patrols were expected to appear relatively less efficient compared with metropolitan patrols because they required more inputs to provide an effective service.

5.4.3 Performance of the NSW police patrols

An input orientated DEA model was chosen to calculate the overall technical efficiency scores for police patrols because the objective of the NSW police service is to provide effective policing with least inputs to the community. However, caution is required in interpreting the results. An apparently less efficient patrol was not necessarily an ineffective patrol. Given the DEA specification, a patrol with relatively low measured outputs might have reduced crime through better policing. However, relatively low measured outputs could have resulted from poor policing (that is, police did not pay proper attention to crime or did not report their response to crime). Alternatively, poor policing might have caused higher measured outputs — for instance, a police patrol could have increased its measured law enforcement outputs by ignoring or paying little attention to police procedures to solve crime, thus, encouraging crime which further increased the measured outputs of law enforcement. With these caveats in mind, summary statistics of the various measures of input-orientated technical efficiency are presented in Table 5.4.2.[32]

First, considering the pure technical efficiency, police patrols, on average, may be able to produce the same level of measured outputs with 13.5 per cent fewer inputs, holding the current input ratios constant. Using a Z-test, the null hypothesis that the sample mean was one at the 5 per cent level of significance was rejected. However, about one-third of the patrols appeared to need to reduce their inputs by less than the sample average if they were to become efficient. Moreover, approximately 35 per cent of patrols were apparently radially efficient.

Disaggregating the results by location reveals that country patrols, on average, appeared to be more efficient than metropolitan counterparts. This was contrary to expectations, and may be partly because the kilometres that beat police did on foot in metropolitan patrols were excluded from the measure for proactive policing. Using a Kruskal-Wallis test, it was found that the null hypothesis that there was no significant difference in the pure technical efficiency for the

32 The software DEAP developed by Coelli (1995) is used to calculate the various measures of technical efficiency.

metropolitan and country patrols at the 5 per cent level of significance could not be rejected.[33]

The mean pure technical efficiency score overstated the efficiency of police patrols, if there appears excessive use of some inputs, beyond that reflected by the efficiency scores. Table 5.4.3 presents a summary of the excess in each input after radial technical inefficiency is removed. This reveals the scope for further non-radial reductions in inputs once a police patrol operates on the production frontier. Holding the level of police services constant, on average, it would appear that police patrols could reduce their use of police cars by 1.96 cars; sixty-two patrols may be able to reduce the number of civilian employees by 2.24 persons; and eight patrols might be able to reduce number of police officers by 11.32 officers. Their excessive use of inputs accounted for 1 to 13 per cent of total inputs. The apparent excessive use of civilian employees by almost 40 per cent of patrols in the sample may be because civilian output was excluded from the specification of the DEA model. Some of the excess inputs may have been converted into other outputs provided by police patrols which were not measured in this study.

It is shown in Table 5.4.2 that average scale efficiency of 0.94 suggested further potential input savings of 6 per cent if a police patrol could operate at the constant returns to scale technology. Investigating the distribution of scale in Table 5.4.4 reveals that 18 per cent of patrols appeared to already operate at the appropriate scale. Approximately half of the patrols in the sample appear to be experiencing decreasing returns to scale, and could be reduced in size. On the other hand, about one-third of the patrols seemed to be experiencing increasing returns to scale, and these may be able to be consolidated with other small units to achieve the optimal size. The comparison of the metropolitan and country police patrols gives a slightly different picture: the apparently scale inefficient metropolitan police patrols were roughly split between increasing and decreasing returns, while there were twice as many apparently scale inefficient country police patrols operating on the decreasing returns region relative to those on the increasing returns range. However, an across-the-board downsizing of larger patrols may not be justified, and it may be more appropriate to consider the patrols on a case-by-case basis before any restructuring policy is implemented.

To determine whether environmental factors might affect the measured efficiency of police patrols, the pure technical efficiency scores were regressed

[33] The Kruskal-Wallis statistic (adjusted for ties) had a value of 2.83 for overall technical efficiency, 3.24 for pure technical efficiency and 0.73 for scale efficiency. The associated p-values for these statistics were 0.093, 0.072 and 0.394.

against the proportion of young people in a patrol, the proportion of government housing in a patrol, and the location of a patrol. The Tobit results in Table 5.4.5 explains about 6 per cent of the variation in pure technical efficiency scores, and none of the coefficients of the explanatory variables were significant at the 5 per cent level of significance. The insignificance of location confirmed the prior finding based on the Kruskal-Wallis test. Consequently, it was concluded that the measured efficiency of police patrols was not influenced by these environmental variables.

5.4.4 Conclusions

The NSW Government is implementing contractual budgeting for government service providers to encourage the delivery of efficient and effective services. Performance measurement of government service providers is necessary to allow the community to reap the full rewards of contractual budgeting. It encourages providers to improve their efficiency and effectiveness because this information makes them more accountable to Parliament, and it promotes yardstick competition in the provision of government services that face little competition.

NSW Treasury is using DEA to measure the efficiency of major government service providers. Furthermore, NSW Treasury is encouraging these agencies to acquire knowledge of the DEA so as to be able to maintain the DEA models developed by Treasury and to use the results of the studies in their corporate planning and internal resource allocation.

This study suggested that NSW police patrols, on average, might be able to produce the same measured outputs with 13.5 per cent less inputs at the current scale and using their inputs efficiently. However, the average reduction masked the fact that the apparent reduction in inputs for about one-third of the patrols to become technically efficient would be less than the average reduction, and that 35 per cent of the patrols already appeared to be technically efficient. No significant difference was found in the technical efficiency of country patrols and metropolitan patrols. The technical efficiency scores for some patrols may have overstated their technical efficiency because they had excess inputs beyond that reflected in their efficiency scores, which accounted for 1 per cent to 13 per cent of total inputs. Care is required in interpreting these results because it is unknown how the quality of police work influences the measured outputs of a police patrol. Nevertheless, the results provided indicative information on the technical efficiency of NSW police patrols. Scale inefficiency was not a major source of input reduction. However, if it is possible to restructure the police patrols, the potential input savings, on average, may be 6 per cent. The restructure of the police patrols should be implemented only after individual police patrols' responsible boundaries are carefully investigated. Some police patrols may be able to benefit from downsizing; and others from expansion.

The DEA results and the effectiveness indicators developed by the Steering Committee provide the NSW police service with an objective input to development of policies to improve the performance of police patrols. To ascertain the reasons for the apparent divergence in the technical efficiency of police patrols, NSW Treasury has provided the NSW police service with details on the measured technical efficiency of individual patrols and information on peers for each patrol which appears to be less efficient. The effectiveness indicators provide further information on the performance of police patrols. Given this information, the NSW police service could commence a series of operations audits to determine if the apparent differences in technical efficiency of police patrols is a result of poor policing, proactive policing or thorough policing by some patrols. This will allow the NSW police service to judge the extent to which, and speed with which, it may be able to improve the technical efficiency of its patrols.

Further work is required to improve the measures of some of the outputs and inputs used in this study. The measure of proactive policing (the kilometres travelled by a police car) did not capture all the aspects of this form of policing and probably disadvantaged metropolitan patrols because it ignores beat duty done on foot. The study's measures for labour did not allow for differences in the quality of labour, but this problem probably cannot be resolved quickly. Still, information on the full-time equivalent hours of police officers and civilian employees would provide a better indication of the labour input used by a police patrol compared with the number of employees in a patrol. The NSW police service recently introduced a financial system that indicates to which patrol an employee is attached for every pay of the year. This should permit the NSW police service to develop information on the full-time equivalent hours worked by police officers and civilian employees in police patrols. However, the financial system cannot provide details on the computers and other equipment installed in police patrols, so an improvement in the measure of capital (the number of police cars) is not imminent.

Table 5.4.1: Descriptive statistics for NSW police patrols [a]

	Mean	*Standard deviation*	*Minimum*	*Maximum*
Outputs				
Offences	3670.31	2345.08	360	12 395
Arrests	938.70	625.53	145	3 215
Summons	101.76	104.14	6	596
Major car accidents	450.05	284.05	31	1 663
Kilometres travelled by police cars	422 265	233 598.10	127 146	1268 555
Inputs				
Police officers	50.92	26.17	9	127
Civilian employees	6.57	5.99	0	41
Police cars	10.37	5.43	3	34

[a] 163 observations

Table 5.4.2: Summary statistics: efficiency of metropolitan and country police patrols

	Efficiency measures		
	Technical efficiency[a]	*Pure technical efficiency*[b]	*Scale efficiency*[c]
Mean			
Total	0.8129	0.8650	0.9408
Metropolitan	0.7944	0.8464	0.9399
Country	0.8408	0.8929	0.9422
Standard deviation			
Total	0.1460	0.1395	0.0787
Metropolitan	0.1551	0.1543	0.0779
Country	0.1271	0.1120	0.0804
Minimum value			
Total	0.4446	0.4477	0.5782
Metropolitan	0.4446	0.4477	0.5782
Country	0.6226	0.6422	0.6811
Number of efficient patrols			
Total	29	57	29
Metropolitan	16	31	16
Country	13	26	13
Number of observations			
Total	163	163	163
Metropolitan	98	98	98
Country	65	65	65

[a] Constant returns to scale model
[b] Variable returns to scale model
[c] Constant returns to scale technical efficiency/variable returns to scale technical efficiency

Table 5.4.3: Input slack variables

Inputs	*Number of patrols with slacks*	*Mean*	*Total slack as percentage of total input*
Police officers	8	11.32	1.09
Civilian employees	62	2.24	12.98
Police cars	10	1.96	1.15

Table 5.4.4: Summary of returns to scale of NSW police patrols

	Increasing returns to scale	*Constant returns to scale*	*Decreasing returns to scale*	*Total*
Metropolitan	39	16	43	98
Country	16	13	36	65
Total	55	29	79	163

Table 5.4.5: Results of Tobit regression

Variable	*Normalised coefficient* [a]	*Standard error*	*t - ratio*	*Regression coefficient*
Proportion of young people that live in or visit a patrol	–3.4063	2.4865	–1.3700	–0.6495
Proportion of government housing in a patrol	–2.3768	1.3310	–1.7857	–0.4532
Location of patrol	0.1354	0.1988	0.6809	0.0258
Constant	5.6237	0.7181	7.8303	1.0724
Pure technical efficiency score of a patrol [b]	5.2441	0.3917	13.387	

Log-likelihood function = –23.6772

Standard error of estimate (σ) = 0.1906

Squared correlation between observed and expected values = 0.0666

[a] Normalised coefficient = regression coefficient/ = $\beta i/\sigma$.

[b] Dependent normalised coefficient = $1/\sigma$

5.5 Technical efficiency of NSW motor registry offices

5.5.1 Introduction

The Roads and Traffic Authority (RTA) is responsible for licensing drivers, registering vehicles, promoting road safety and traffic management, and constructing, maintaining and enhancing roads in NSW. The RTA allocated about $144 million in 1995-96 to motor registries which predominantly oversee the first two responsibilities.

The RTA differs from other NSW budget sector agencies because it has access to a defined pool of funds from Commonwealth grants, user charges, and hypothecated state taxes such as the motor vehicle weight tax and fuel levies. Despite having a more defined revenue stream, the RTA is still subject to Government direction and control. Its operations must encompass current Government policies and initiatives such as improvements in resource allocation and efficiency reviews.

The RTA has an extensive array of performance indicators to monitor and improve the effectiveness and efficiency of the delivery of motor registry services. Customer satisfaction is used as an indicator of effectiveness. Efficiency indicators include total cost per weighted transaction,[34] weighted transactions per labour hour, and the time spent by customers in registries. These are useful efficiency measures but they can vary for reasons other than inefficiency, such as the scale of the motor registry, different input combinations used by registries, and the environment in which services are delivered.

More sophisticated techniques, such as DEA, assess the efficiency of a motor registry by comparing its inputs to produce services, and take into account its scale and its operating environment in examining that efficiency. Better information on the efficiency of motor registries provides the RTA with additional opportunities to free funds for other uses such as road maintenance, or to provide the same services with less reliance on state motor vehicle weight taxes or fuel levies. This paper examines the scope for the RTA to improve the efficiency of its motor registries using DEA.

---

[34] Motor registries may perform up to 150 different types of transactions. Each type of transaction is weighted by the average time taken to perform it.

Related studies

During this study, it was unknown whether there were other studies that measured the efficiency of motor registries using DEA. However, there was a proposal to measure the performance of Queensland Transport using DEA (National Road Transport Commission 1994). The proposal aims to develop performance indicators for four areas of Queensland Transport's operations — road maintenance, road construction, system stewardship and driver licensing and vehicle registration. Customer service centres provide driver licensing and vehicle registration services. The proposal considers that customer services are the output of customer service centres, and that labour, capital and materials are the inputs. Tasman Asia Pacific (1996) prepared a report for the National Road Transport Commission that included an assessment of the efficiency of Queensland customer service centres after the completion of this analysis. The study has one output for customer service centres total, minutes of service (the number and type of transactions weighted by the average time for each type of transaction) and two inputs, total labour costs and other operating costs which exclude rates, rent and capital purchases because there is incomplete information on these costs.

There is a substantial body of literature on financial institutions (banks and credit unions), post offices and pharmacies using DEA. Motor registries operate in an analogous manner to these service providers because they provide counter services and form part of a branch network. Therefore, these studies provided a guide to specifying the outputs and inputs used in this study.

Studies of the efficiency of financial institutions have similar measures for inputs. However, their measures for outputs differ from those used in this study because the monetary transactions that take place in financial institutions are of a different nature from those in motor registries.

Ferrier and Lovell (1990), Aly et al (1990), Fried, Lovell and Eeckhaut (1993), Ferrier et al (1993), and Fried and Lovell (1994) considered the important inputs of financial institutions were labour, raw materials and capital. These inputs were either combined and measured in aggregate operating expenditure or measured individually — for example, labour was measured in staff numbers or the wage bill; raw materials were measured by expenditure on this input, or as non-labour operating expenditure; and capital was measured by rental costs or by its book value. Outputs included a variety of loans and bank deposits which were measured in physical quantities or in dollars.

Doble (1995) presented a model of post offices in the United Kingdom. Labour, measured in full-time equivalent hours by counter clerks, was the only input used in the study. Doble excluded hours of work done by branch managers, arguing that the management of staff did not have a direct effect on the production of counter transactions.

A post office handles approximately 190 different types of transactions. Different types of transactions require different amounts of staff time, so transactions are weighted by the average time to complete each type of transaction. Doble included nine categories of weighted transactions, such as issuing stamps and vehicle licences, as outputs in his study. The large sample of 1281 post offices allowed Doble to include a large number of outputs in the DEA analysis and still obtain sensible results. The study included an output for the quality of service provided by post offices, which was measured by the average time that a customer waits for service.

The Färe, Grosskopf and Roos (1995) study of state–owned Swedish pharmacies included several outputs, such as preparing prescriptions, delivering drugs, and selling special articles and food for people with disabilities (which were measured by the number of services provided by pharmacy staff); selling consumer goods (which was measured by the number of sales, or transactions); and conveying information on drugs (which was measured by the hours spent collecting, preparing and conveying the information). The Swedish Government required the pharmacies to meet certain quality of service standards, so several attributes of the provision of pharmacy services were included in the model: business hours, the percentage of prescriptions filled in one day, and the time that customers wait for service. Inputs included pharmacists and technical staff, measured by the hours they worked during the year, and non-labour operating expenditure.

5.5.2 Provision of motor registry services

Outputs

A motor registry may perform up to 150 different transactions, which include the issue and renewal of driver licences, motor vehicle registration, number plates, firearm licences and driver licence testing. The RTA records all the transactions conducted by motor registry counter staff. The different transactions require similar staff skills but different amounts of time, so the total number of transactions might not reflect the resources used. Thus, the total number of transactions weighted by the average time spent to perform each type of transaction was adopted as a proxy for the services provided by motor registries.

The total number of weighted transactions did not reflect the quality of the service provided. One aspect of the quality of service provided by a motor registry is the time that a person queues for service: the longer a person waits, the more likely they are to be dissatisfied with the quality of service. Twice a year, the RTA measures this waiting time in a motor registry. In this study, waiting time was calculated as the average of the two surveys.

However, waiting time is an output that registry management should minimise, so the inverse of waiting time is used in the analysis. Doble (1995) used an alternative method of inverting waiting time: the average time that a person waits for service in a post office was subtracted from the highest average time that a person waits for service. This indicator of quality now measured the time that a customer did *not* wait for service compared with the maximum time they *could* wait for service.

Inputs

Motor registries use people, capital equipment and raw materials (such as stationery) to provide their services. Labour was measured by the total hours that staff work in the year, capital was measured by the value of computer equipment, and raw materials were measured by the expenditure on these items.

The total hours that staff work in a year included the work of managers and supervisors, permanent and casual staff, and staff on loan to the registry. It excluded recreation, sick and special leave, training away from the registry, staff on loan to other registries, and managers away attending meetings or participating in quality improvement teams to improve the performance of motor registries in a particular region.

The capital of a motor registry included computer terminals, photocopiers, telephones and buildings. However, the RTA had incomplete information on these assets; it suggested that the number of computer terminals was a good proxy for the capital used by a motor registry. However, information on the number of computer terminals actually used to serve customers was not available. The bulk of the computer terminals were installed in registries in 1991, and only minor investment in computer equipment had occurred since that date. Therefore, the value of computers installed in a motor registry in 1991 should have reflected the number of computer terminals it used, provided the number of terminals had not been altered in response to significant changes in the demand for registry services.

The main raw materials used to produce transactions included licences, plates, postage and stationery. Total expenditure on these items was used to measure this input.

In summary, the DEA model for motor registries had weighted transactions and the reciprocal of waiting time as outputs, and the total staff hours, the value of raw materials and the value of computers in 1991 as inputs.

Influence of environmental variables

Some factors beyond the control of management may influence the efficiency of motor registries. The RTA identified two environmental variables which were considered in this study: whether the registry was open on a Saturday, and whether the RTA motor registry did data entry for councils and police stations that operate as agencies of the RTA in remote regions of New South Wales.

To improve customer service, the RTA extended business hours of motor registries by opening selected registries statewide on Saturday mornings. This allowed these registries to use their capital more effectively. These registries faced heavy demand on Saturdays. It was estimated that about two-thirds of Saturday customers came from the Monday to Friday business of the same registry. The remaining third was estimated to come from business of other registries that did not open on Saturday. Given that most customers choose to visit a motor registry rather than use the mail to complete their business, the additional business to Saturday traders lowered the demand for services in other registries that did not open that day. Moreover, motor registry managers had limited scope to adjust labour to reductions in demand for Monday to Friday trade, because they had to employ full-time staff to cover the peak period of demand for their business.[35] Therefore, a motor registry that traded Monday to Friday could appear less efficient, compared with a similar registry that traded Monday to Saturday.

All but one of the agencies did not have access to the central computer system of the RTA, so other registries (either nearby country registries or registries requiring extra work) processed their transactions into the RTA database. Labour and capital were likely to be used more efficiently when processing an agency’s transactions because staff did not serve customers. Further, agency work could be processed during non-peak periods of customer demand, enabling labour and capital to be used more fully. Accordingly, agency work allowed parent registries to process more transactions than non-parent registries, for the same amount of inputs.

---

[35] The RTA is in the process of negotiating for greater flexibility in staffing patterns so registry managers can schedule staff to meet the varying demand for services during the day.

5.5.3 Data and estimation of technical efficiency

Registry managers have no control over the demand for the registry's services. Therefore, their main objective is to handle transactions with the least inputs. This implies minimising the resources used to complete each transaction while maintaining the quality of service. Accordingly, a DEA model with an input-minimising orientation was used to estimate the technical efficiency of motor registries. There were 137 motor registries in NSW in 1994-95, but only 131 registries were included in this study. (Registries that were to close during 1994-95 were excluded, because there was incomplete data for these registries.) Forty registries opened on Saturdays and eighteen performed data entry for agencies of the RTA.

Outlier analysis

DEA is susceptible to outliers, which are observations which are not typical of the rest of the data. The production frontier estimated by DEA is determined by extreme input and output points, so a single outlier can heavily influence the measures of technical efficiency. Outliers may arise for two reasons. Outliers may arise from errors in the data caused by measurement error or reporting error. Alternatively, if the data is viewed to come from a probability distribution then it is possible for the data to have observations with a low probability of occurring. Outliers may reflect important phenomena which would go unnoticed if they were excluded from the analysis (Burgess and Wilson 1993).

The data for motor registries were screened for potential outliers by examining the summary statistics for each output and input, which are presented Table 5.5.1.

Table 5.5.1: Descriptive statistics for NSW motor registry offices[a]

	Mean	*Standard Deviation*	*Minimum*	*Maximum*
Outputs				
Weighted transactions	179 932	118 850	24 080	607 784
Reciprocal of waiting time (1/minutes)	0.43	0.88	0.10	10.00
Inputs				
Labour (hours)	14 029	9 136	2 809	41 906
Raw materials ($)	66 789	47 452	7 379	286 675
Computers ($)	91 060	36 884	41 163	234 048

[a] 131 observations

Furthermore, output–input ratios were calculated for each motor registry and the values were checked using a two-and-a-half standard deviation rule. That is, any observation which was observed to be greater than two-and-a-half standard deviations from the sample mean was considered a potential outlier. Potential outliers were referred to the RTA who checked the data and confirmed there were no obvious measurement or reporting errors. Consequently, no observations were excluded from the analysis.

5.5.4 Results

Summary statistics of the various measures of technical efficiency are presented in Table 5.5.2. The method for calculating the technical efficiency and its components is presented in Appendix A. The results presented suggest that pure technical efficiency was the main source of technical inefficiency rather than scale inefficiency. Pure technical efficiency indicated the possible improvement in the use of inputs to produce the same outputs that the RTA could achieve without altering the scale of its motor registries. This could be called managerial efficiency. On average, it appeared that motor registries may be able to produce the same level of measured outputs with 15 per cent fewer inputs.

Table 5.5.2: Summary statistics: technical efficiency of NSW motor registry offices

	Technical efficiency (scale and pure technical efficiency) [a]	*Pure technical efficiency* [b]	*Scale efficiency* [c]
Mean	0.81	0.85	0.95
Standard deviation	0.10	0.09	0.07
Minimum	0.55	0.63	0.64
Number of efficient motor registries	6	14	6

[a] Constant returns to scale model

[b] Variable returns to scale model

[c] Constant returns to scale model/variable returns to scale model

Motor registries, on average, appeared to be 95 per cent scale efficient. If motor registries could adjust to their optimal scale, then they may be able to further reduce inputs, on average, by 5 per cent. The results in Table 5.5.3 indicate that the majority of registries which appeared to be less efficient were experiencing increasing returns to scale, suggesting that they were too small rather than too big.

There may be social, demographic or geographic reasons for a motor registry being a particular size. If there were no barriers to amalgamation or separation of registries, then information on scale efficiency could assist the RTA in determining the optimal size of their registries. If barriers do exist, information on scale efficiency indicates the costs incurred in maintaining the existing level of service provision in a particular region.

Table 5.5.3: Summary of returns to scale on NSW motor registries

	Increasing returns to scale	*Constant returns to scale*	*Decreasing returns to scale*
Number of motor registries	75	6	50

The mean pure technical efficiency score could overstate the efficiency of motor registries if some inputs are used excessively, beyond that reflected in the efficiency scores. About 70 per cent of motor registries have excessive input use. Table 5.5.4 reveals the scope for further non-radial reductions in inputs (termed as ‘slacks’) once a motor registry operates on the production frontier. Motor registries with such excessive inputs may be able to reduce their labour, on average, by 616 hours, their raw materials by $11 152 and their computer terminals by $11 274. Excessive inputs as a proportion of total inputs accounted for about 9 per cent of raw materials, 5 per cent of computers and less than 1 per cent for labour.

Table 5.5.4: Slack input variables

Input	*Number of registries with slacks*	*Mean*	*Total slacks as a percentage of total inputs*
Labour (hours)	6	616	less than 1
Raw materials ($)	67	11 152	9
Computers ($)	51	11 274	6

There are several methods for including environmental variables in a DEA analysis. These are discussed in Chapter 2. This study used a two-step procedure to analyse external factors of operating environments which might have explained the variation in technical efficiencies across motor registries. The two-step procedure required that inputs and environmental variables were not correlated (Lovell 1993), and so Saturday trading and agency work were not strongly correlated with the inputs used in this study.

The pure technical efficiency scores were regressed against the environmental variables that indicated whether the registry conducted agency work or traded on Saturdays. Given that the pure technical efficiency scores were truncated at one, and that ordinary least squares estimation of censored data produced biased and inconsistent results, a Tobit procedure was used (Judge et al. 1988).

The algebraic version of the Tobit model is presented in equation (1) and the estimated model is presented in Table 5.5.5.

(1) $TE_i = \beta_0 + \beta_1 AGENCY_i + \beta_2 SAT_i + u_i$ $\quad i = 1, \ldots, 131$

TE_i is the pure technical efficiency score of the i-th registry and u_i is a normally distributed disturbance term. $AGENCY_i$ and SAT_i are binary variables for the i-th registry. A value of one in the *AGENCY* variable indicates that the registry conducts agency work while a zero indicates otherwise. Similarly, a one in the *SAT* variable indicates that it trades on Saturdays while a zero indicates that it is not open on Saturdays.

The sign of the coefficients indicates the direction of influence of the environmental variables, and the ratio of the estimated coefficients to their standard errors (t-ratios) indicates the strength of the relationship between efficiency and each variable. The squared correlation between the observed and expected values indicates how much of the variation in efficiency scores can be explained by the environmental variables (agency work and Saturday trading).

Table 5.5.5: Results of the Tobit regression

	Normalised coefficient	*Standard error*	*t-ratio*	*Regressionn coefficient*
SAT	0.1678	0.1921	0.8736	0.0161
AGENCY	0.1941	0.2577	0.7531	0.0187
CONSTANT	8.7864	0.6001	14.64	0.8445
TE	10.4040	0.7026	14.81	

Log-likelihood function = 90.4220
Mean-square error = 0.0076
Mean error = 0.0019
Squared correlation between observed and expected values = 0.0091

The signs on the estimated coefficients were as expected. Both variables had a positive influence upon the level of pure technical efficiency. However, neither variable was significant at the 5 per cent level of significance. The equation explained about 1 percent of the variation in pure technical efficiency scores. Based on these results, the pure technical efficiency scores were not adjusted for the influence of agency work and Saturday trading.

5.5.5 Conclusions

The analysis of the technical efficiency of motor registries, as measured, suggested, on average, that they may be able to produce the same level of outputs with 15 per cent fewer inputs. Pure technical efficiency appeared to be the main source of inefficiency rather than scale inefficiency. However, if motor registries could achieve optimal scale they could further reduce inputs, on average, by 5 per cent. Care is required in interpreting the results because there were weaknesses with the measure for capital. Nevertheless, the results provided indicative information on the technical efficiency of motor registries.

To improve the measure of capital, the RTA has surveyed each motor registry to obtain the number of computer terminals it uses to process transactions. Further, it is developing weightings for agency work. This will reduce the potential for these transactions to improve the technical efficiency of registries that process agency work. These improvements will be included in future studies that determine the technical efficiency of motor registries.

NSW Treasury has provided the RTA with the technical efficiency scores for individual motor registry offices and associated peer information from this study as a systematic framework for raising and addressing questions about the performance of their motor registries.

APPENDIX A TECHNICAL GUIDE TO DEA

DEA is the term used by Charnes and Cooper (1985) to describe a non-parametric approach to measuring efficiency. Diewert (1993) and Zeitsch and Lawrence (1996) recently reviewed this technique, and the following discussion borrows from their work.

A1 Technical efficiency

There are several different ways to present the technical efficiency linear programming problem for DEA. The simplest presentation for the input-oriented, constant returns to scale version of DEA is:

(A1) Minimise E_n

$w_1,...,w_N,E_n$

Subject to:

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

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

$$w_j \geq 0 \qquad j = 1,...., N .$$

where there are N organisations producing I different outputs y_{in} for $i = 1, \ldots , I$ using K different inputs x_{kn} for $k = 1, \ldots , K$. The w_j are weights applied across the N organisations and the solution, E_n*, is the efficiency score of the nth organisation. The linear program solves for the convex combination of the N data points that can produce at least the observation n output and use at most E_n* times the observation n combination of inputs. To get a full set of efficiency scores, this problem has to be solved for each of the N organisations.

The linear programming problem for the output-oriented, constant returns to scale version of DEA is similar to the above problem, except that it takes the convex combination of observations that uses no more inputs than organisation *n* and produces the maximum amount of outputs. The output-oriented DEA linear programming problem is given by the following equation:

(A2) Maximise F_n

$w_1,\ldots,w_N,F_n$

Subject to:

$$\sum_{j=1}^{N} w_j y_{ij} - F_n y_{in} \geq 0 \qquad i = 1,\ldots., I$$

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

$$w_j \geq 0 \qquad j = 1,\ldots., N .$$

The first constraint indicates that the output of the hypothetical weighted average has to be at least as great as *n*'s output scaled up by the factor F_n. The second set of constraints state that the weighted average of the inputs cannot be any larger than *n*'s input.

Returning to the input-oriented case, the constant returns to scale technical efficiency score can be decomposed into three components — one due to a sub-optimal scale of operations (scale efficiency); a second due to an inability to dispose of 'surplus' inputs (congestion efficiency); and a residual or 'pure' technical efficiency. To form these measures, the DEA linear programs in (A1) need to be re-run under the assumptions of variable returns to scale and variable returns to scale with congestion.

The variable returns to scale DEA linear program is given by:

(A3) Minimise S_n

$w_1,\ldots,w_N,S_n$

Subject to:

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

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

$$\sum_{j=1}^{N} w_j = 1$$

$$w_j \geq 0 \qquad j = 1,\ldots., N.$$

As noted in Chapter 3, the extra constraint that the weights must sum to one has the effect of pulling in the frontier to form a tighter envelope around the data.

The DEA linear programming problem under variable returns to scale with congestion is given by:

(A4) Minimise C_n

$w_1,\ldots,w_N,C_n$

Subject to:

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

$$\sum_{j=1}^{N} w_j x_{kj} - C_n x_{kn} = 0 \qquad k = 1,\ldots., K$$

$$\sum_{j=1}^{N} w_j = 1$$

$$w_j \geq 0 \qquad j = 1,\ldots., N.$$

The effect of placing an equality on the input constraint is to allow the frontier to 'bend backwards' as in Figure 3.4. In technical terms, the assumption of 'free disposability' of inputs is removed. This means that an organisation cannot costlessly get rid of inputs to move down to the segment of the frontier that runs parallel to the axes in Figure 3.2.

The three components of technical efficiency can now be defined as follows:

(A5) Scale efficiency = E_n / S_n

(A6) Congestion efficiency = S_n / C_n

(A7) Residual efficiency = C_n

The product of (A5), (A6) and (A7) is simply the constant returns to scale efficiency score, E_n, in the original DEA model (A1).

As noted in Chapter 2, a scale efficiency score of less than one does not indicate whether the organisation is bigger or smaller than its optimal size. To establish this, an additional variant of DEA — one subject to non-increasing returns to scale — must be run. The DEA linear programming problem for the non-increasing returns to scale case is given by:

(A8) Minimise R_n

$w_1,...,w_N,R_n$

Subject to:

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

$$\sum_{j=1}^{N} w_j x_{kj} - R_n x_{kn} \leq 0 \qquad k = 1,...., K$$

$$\sum_{j=1}^{N} w_j \leq 1$$

$$w_j \geq 0 \qquad j = 1,...., N.$$

If the scale efficiency score is less than one and E_n and R_n are equal, then n is subject to increasing returns to scale and would need to increase its size to reach its optimum scale. If E_n is less than R_n , then n is subject to decreasing returns to scale and would need to reduce its size to reach its optimum scale.

The technical efficiency DEA problems outlined in this section only require information on output and input quantities. They do not use any information on output or input prices. As noted in the chapters, the difficulty of allocating prices to human services outputs makes DEA a relatively attractive technique compared with total factor productivity. However, even for human services, information on input prices and costs is often available, allowing an organisation's cost and allocative efficiency to be calculated.

A2 Cost and allocative efficiency

If the input prices for each organisation are known, then the cost efficiency score for each observation can be calculated by solving N linear programs of the form:

(A9) Minimise $\sum_{k=1}^{K} p_{kn} x_{kn}$

$w_1,\ldots w_N, x_{1n},\ldots,x_{Kn}$

Subject to:

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

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

$$w_j \geq 0 \qquad j = 1,\ldots., N$$

where $p_{1n}, \ldots, p_{Kn}$ are the input prices for the K inputs that unit n faces.

This linear program chooses the input quantities that minimise n's total costs subject to a feasibility constraint and assuming that the input prices it faces are fixed. The feasibility constraint requires that the weighted average which forms the hypothetical efficient organisation has outputs at least as great as n's and inputs no greater than n's. The solution vector to (A9) $x^*_{1n},\ldots,x^*_{Kn}$ is n's cost-minimising level of inputs given its input prices and output level.

The technical efficiency scores derived from the linear programming problem (A1) can be combined with the solutions to the cost-minimising linear programming problems (A9) to form measures of the cost and allocative efficiency of each organisation. Specifically, cost efficiency is found by dividing the costs that would be faced by an organisation if it used the cost minimising level of inputs by its actual costs. Thus:

(A10) Cost efficiency for the n-th observation= $\sum_{k=1}^{K} p_{kn} x_{kn}^{*} / \sum_{k=1}^{K} p_{kn} x_{kn}$.

A score of one for this index would indicate that an organisation is cost efficient.

Allocative inefficiency is calculated by dividing costs faced by an organisation assuming it used the cost-minimising level of inputs by costs assuming the organisation used the technically efficient level of inputs. Thus:

(A11) Allocative efficiency $= \sum_{k=1}^{K} p_{kn} x_{kn}^{*} / E_n \sum_{k=1}^{K} p_{kn} x_{kn}$

where E_n is the technical efficiency score derived from the linear programming problem (A1).

From (A11) it can be seen that an organisation's cost efficiency is the product of its allocative efficiency and its technical efficiency.

A3 Accounting for operating environment differences

There is always a trade-off in DEA studies between ensuring that a like-with-like comparison and maximum use is made of the information available to learn how to improve performance. Limiting the study to like-with-like comparisons leads to the comparisons being 'fairer' and perhaps more readily acceptable to managers. However, a diverse range of operating environments may be useful in the study to provide a wider range of ideas and operating styles from which managers could learn.

In most cases, at least some allowance will need to be made for differences in the organisations' operating environments. As noted in Chapter 2, there are several different ways to do this. The simplest way is to restrict the comparison set to other organisations that have similar or less favourable operating environments. However, this selection process can be arbitrary and excludes a lot of the information that might be available. More sophisticated ways of allowing for operating environment differences involve either single or multiple stage adjustment processes.

In the single stage adjustment methods, the DEA linear programming problem directly incorporates the operating environment characteristic that is being adjusted for. Again, there are several ways this can be done. One option is to simply include the characteristic in a manner analogous to the other inputs. This assumes that the characteristic can be radially contracted as can the other inputs, which is unlikely to be realistic in most cases. For instance, in the example of the impact of an area's socioeconomic status on schools, a school cannot change (at least in the short run) the socioeconomic status of its neighbourhood. Rather, it is more appropriate to take this as being fixed. In that case, the characteristic needs to be included as a constraint in the linear program, but not one into which the efficiency score enters. This reflects the fact that the organisation has no control over the characteristic. This is known as including the characteristic as a non-discretionary input. The linear program for this problem is as follows:

(A12) Minimise V_n

$w_1,\ldots,w_N,V_n$

Subject to:

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

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

$$\sum_{j=1}^{N} w_j z_j - z_n \leq 0$$

$$\sum_{j=1}^{N} w_j = 1$$

$$w_j \geq 0 \qquad j = 1,\ldots, N .$$

where z_n is the value of the operating environment characteristic in question for unit n, n=1,…, N. This specification can be extended to include multiple operating environment characteristics but, as with any DEA specification, including more constraints will tend to inflate efficiency scores. This technique also requires the operating environment characteristic to be a continuous variable.

Two-stage adjustment procedures allow more flexibility than the above procedure. These techniques typically carry out the initial DEA calculation without referring to operating environment characteristics. Then, they regress the efficiency scores from the DEA problem against a set of operating environment characteristics using the Tobit regression technique to allow for the truncated distribution of the efficiency scores. The DEA scores can be adjusted to ‘standardise’ for the particular characteristic. This produces a set of efficiency scores assuming that all organisations were operating with the same degree of this characteristic. This approach has the advantage of not requiring the direction of the characteristic’s impact on the efficiency scores to be specified in advance. It also means that statistical tests can be carried out on the strength of the relationship between the characteristic and efficiency levels.

Some practitioners have extended the two-stage adjustment procedure to three and four stages to allow for slacks as well as radial inefficiency. Fried, Schmidt and Yaisawarng (1995) described one such approach in detail, as well as reviewing different approaches to adjusting for operating environment differences.

APPENDIX B PROGRAMS FOR THE APPLICATION OF DEA

There are a number of software options for running DEA. These can be categorised as specialist DEA software and other software which has the capacity to conduct linear programming and which can be customised to perform DEA. Some examples and contact points are listed below.

B 1 Specialist DEA software

This software can be purchased from the following contact points.

- DEAP — Tim Coelli, University of New England, Armidale, Australia
  Fax 067 73 3607, Email: tcoelli@metz.une.edu.au
- Frontier Analyst — Bernard Petter, Strategic Retail Directions
  Fax 03 9574 8882, Web site: http://www.scotnet.co.uk/banxia/famain.html
- Frontier Analyst — Marjory Sweeney, Banxia Software, Glasgow
  Fax: +44 141 552 5765, Web site: http://www.scotnet.co.uk/banxia
- IDEAS — Shirley Shmering, 1 Consulting (US), Fax: + 413 256 1211
- PASS — Dr C.T. Clark, Productivity Assessment Consultants, Educational Productivity Council, University of Texas, U.S.A.
  Fax +1 512 301 1931
- Warwick DEA — Antreas Athanassopoulos, Warwick University (UK),
  Fax +44 203 52 3719
  Web site: http://www.csv.warwick.ac.uk/~bsrlu/dea/deas/deas1.htm

B2 Linear programming software

This software has been developed to run linear programming problems specifically.

- General Algebraic Modelling System (GAMS)
  Web site: http://www.gams.com/docs/intro.htm
- LINDO — Web Site: http://www.lindo.com/products.html

  This software has a linear programming option.
- Microsoft Excel (using the solver tool)

- SAS — Web site: http://www.sas.com/www-bin/jump.pl
- SHAZAM — Web Site: http://shazam.econ.ubc.ca/

Example of programming for DEA using Shazam

This section is a brief review of how to program the variable returns to scale DEA problem in the Shazam econometrics package. The programs were run using version 7 of Shazam and the data for the twenty hospitals listed in Table 3.1.

Using matrix notation, the linear programming module of Shazam is set up to solve problems of the form:

> Maximise $\boldsymbol{c'x}$
> Subject to the constraints:
> $\boldsymbol{Ax} \leq \boldsymbol{b}$
> $\boldsymbol{x} \geq \mathbf{0}$

The x vector contains the coefficients for which Shazam solves (the w_i's and the E_n in the terminology of Chapter 3) and the term $c'x$ is the objective function of the linear program. Because the objective function only contains the E_n term, all the coefficients in the c vector will be zero except for the last one which is equal to one. An option within the linear program command allows the problem to be changed from maximising the objective function to minimising it.

The problem can be thought of in terms analogous to equation system (1) in Chapter 3. The **A** matrix contains the output values, input values and the summation coefficients corresponding to all the terms in the constraints of (1), up to but not including the last term before the inequality sign. The last terms before the inequality sign in the output and input constraints change for each DEA problem being run, because they contain the information specific to the organisation being examined in the individual run.

To automate the process so all twenty DEA problems are solved in the one Shazam run, a number of do loops are introduced. The input file for the Shazam run is:

```
FI 22 A20.TXT
FI 23 B20.TXT
FI 24 C20.TXT
FI 25 D20.TXT
READ (22) A / ROWS=6 COLS=20
READ (23) B / ROWS=6 COLS=1
READ (24) C / ROWS=21 COLS=1
```

```
READ (25) D / ROWS=6 COLS=1
DO #=1,20
COPY A:# B / FROW=1;2 TROW=1;2
COPY A:# D / FROW=3;4 TROW=3;4
MATRIX D=-1@D
MATRIX AD=A|D
?LP C AD B / MIN ITER=100 PRIMAL=S#
ENDO
SMPL 1 21
GENR T=TIME(0)
GENR SCORE=0.0
DO #=1,20
GENR SCORE(#)=S#(21)
ENDO
FORMAT(11F7.4)
PRINT T S1-S10 / FORMAT
PRINT T S11-S20 / FORMAT
PRINT T SCORE / FORMAT
STOP
```

The first commands nominate the files which contain the data. The data are then read in matrix form. The **A** matrix contains the output and input data for the first twenty terms in the constraints in Chapter 3's equation system (1). The first do loop is then formed, which enters the information specific to each organisation found in the terms immediately before system (1)'s inequality signs. The first copy command moves the organisation's output data into the first element of the *b* vector (this is slightly different to the way in which (1) is set out, but it has an equivalent effect). Shazam only allows less than or equal to constraints, so all the output quantities are multiplied by –1 before being entered into the **A** matrix. This makes the less than or equal to constraint on the negatives equivalent to a greater than or equal to constraint on the positive output quantities. Similarly, because Shazam does not explicitly have an equal to constraint, constraints for the sum of the weights — one less than or equal to and the other greater than or equal to — are included. The only way that both constraints can be satisfied is by equality.

The second copy command and the following two matrix commands move the input information for each organisation into the last terms before the inequality signs in (1). This gives the 21 x 6 matrix **AD**.

The DEA linear program runs are done and the solution to each run in an ***s*** vector is saved, the first twenty elements of which contain the weights for the run and the twenty-first element of which is the efficiency score for that organisation. A new score variable, which just contains all the efficiency scores is formed, and finally the ***s*** and score vectors are printed. The ***s*** vectors identify the peer group (those organisations with non-zero weights) and their relative contribution to forming the hypothetical best practice target for the organisation in question.

The following tables reproduce the **A**, ***b***, and ***d*** matrices read into the program. The **c** matrix was described above.

Table B1: Example of twenty hospitals — A matrix

	Column 1	*Column 2*	*Column 3*	*Column 4*	*Column 5*	*Column 6*	*Column 7*	*Column 8*	*Column 9*	*Column 10*
Row 1	–150	–225	–90	–160	–50	–75	–200	–350	–400	–250
Row 2	–50	–75	–10	–40	–50	–75	–50	–100	–90	–300
Row 3	200	600	200	600	500	320	375	400	550	900
Row 4	600	1200	200	300	200	150	450	320	480	660
Row 5	1	1	1	1	1	1	1	1	1	1
Row 6	–1	–1	–1	–1	–1	–1	–1	–1	–1	–1

	Column 11	*Column 12*	*Column 13*	*Column 14*	*Column 15*	*Column 16*	*Column 17*	*Column 18*	*Column 19*	*Column 20*
Row 1	–350	–350	–275	–220	–300	–320	–375	–230	–290	–360
Row 2	–350	–400	–375	–40	–10	–275	–230	–50	–90	–70
Row 3	850	720	900	250	115	600	550	200	450	415
Row 4	720	940	850	370	250	590	710	280	410	575
Row 5	1	1	1	1	1	1	1	1	1	1
Row 6	–1	–1	–1	–1	–1	–1	–1	–1	–1	–1

Table B2: Example of twenty hospitals — b and d vectors

	b vector	*d vector*
Row 1	0	0
Row 2	0	0
Row 3	0	0
Row 4	0	0
Row 5	1	0
Row 6	–1	0

REFERENCES AND FURTHER READING

General references

Ali, A.I, and Seiford, L.M. 1993, 'The mathematical programming approach to efficiency analysis', in Fried, H.O., Lovell, C.A.K, and Schmidt, S.S. (eds) 1993, The Measurement of Productive Efficiency, Oxford University Press, New York.

Aly, H.Y., Grabowski, R., Pasurka, C., and Rangan, N. 1990, 'Technical, scale and allocative efficiency in US banking: an empirical investigation', *Review of Economics and Statistics,* LXXII, pp. 211–18.

Anderson, T. 1996, 'A data envelopment analysis (DEA) home page', *Internet: http://www.emp.pdx.edu/dea/homedea.html*

BIE (Bureau of Industry Economics) 1993, *Data Envelopment Analysis: An Explanation*, AGPS, Canberra.

—— 1994a, *International Performance Indicators: Electricity Update 1994*, Research Report 54, AGPS, Canberra.

—— 1994b, *International Performance Indicators: Gas Supply*, Research Report 62, AGPS, Canberra.

—— 1995, *International Benchmarking: Overview 1995*, Research Report 95/20, AGPS, Canberra.

Bogetoft, P. 1995, 'Incentives and productivity measurements', *International Journal of Production Economics*, vol. 39, pp.67–81.

Charnes, A. and Cooper, W.W. 1985, 'Preface to topics in data envelopment analysis', *Annals of Operations Research,* vol 2, pp. 59–94.

Charnes, A., Cooper, W.W., and Rhodes, E. 1978, 'Measuring the efficiency of decision making units', *European Journal of Operational Research*, vol. 2, pp. 429–44.

Coelli, T. 1996, 'DEAP - a data envelopment analysis (computer) program', University of New England Centre for Efficiency and Productivity Analysis, Internet: http://www.une.edu.au/econometrics/deap.htm

Diewert, W.E. 1993, 'Data envelopment analysis: a practical alternative?', in Swan Consultants (eds), *Measuring the Economic Performance of Government Enterprises*, Proceedings of a Conference held at the Sheraton International Airport Hotel, 12 February.

Dyson, R.G., Thanassoulis, E. and Boussofiane, A. 1995, 'A DEA (data envelopment analysis) tutorial, Internet: http://www.csv.warwick.ac.uk/~bsrlu/dea/deat/deat1.htm

Emrouznejad, A. 1995, 'Ali Emrouznejad's DEA home page' Internet: http://www.csv.warwick.ac.uk/~bsrlu/

Färe, R., Grosskopf, S. and Lovell, C.A.K. 1983, 'The structure of technical efficiency', *Scandinavian Journal of Economics*, vol.85, pp. 181–90.

——, ——, Norris, M., and Zhang, Z. 1994, 'Productivity growth, technical progress, and efficiency changes in industrialised countries', *American Economic Review* vol. 84, pp. 66–83.

Fried, H.O. and Lovell, C.A.K. 1994, 'Measuring efficiency and evaluating performance', Workshop at Regent Hotel, Sydney.

——, and Schmidt, S.S. 1993, *The Measurement of Productive Efficiency*, Oxford University Press, New York.

——, Schmidt, S.S. and Yaisawarng, S. 1995, Incorporating the operating environment into a measure of technical efficiency', Paper presented to the Bureau of Industry Economics seminar, Canberra.

Ganley, J.A. and Cubbin, J.S. 1992, *Public Sector Efficiency Measurement: Applications of Data Envelopment Analysis*, North-Holland, Amsterdam.

Grosskopf, S. 1993, 'Efficiency and productivity', in Fried, H.O., Lovell, C.A.K, and Schmidt, S.S. (eds), *The Measurement of Productive Efficiency*, Oxford University Press.

Innovation Network 1993, 'Applying benchmarking to higher education: part I definitions', Internet: http://www.innovnet.com/ben1.htm

Judge, C.G., Hill, R.C., Griffiths, W.E., Luthepohl, H. and Lee, T. 1988, *Introduction to the Theory and Practice of Econometrics*, John Wiley and Sons, New York.

Klein, R. and Carter, N. 1989, 'Performance measurement: a review of concepts and issues', in Public Finance Foundation (ed), *Performance Measurement: Getting the Concepts Right*, London.

Koopmans, T.C. 1951, 'An analysis of production as an efficient combination of activities,' in Koopmans, T.C. (ed), *Activity Analysis of Production and Allocation*, Cowles Commission for Research in Economics, monograph no. 13, John Wiley and Sons, New York.

Lawrence, D. 1993, ‘What performance measures should we use?’, in Swan Consultants (eds), *Measuring the Economic Performance of Government Enterprises*, Proceedings of a Conference held at the Sheraton International Airport Hotel, 12 February.

Lovell, C.A.K., 1993a, ‘Production frontiers and productive efficiency’, in Fried, H.O., Lovell, C.A.K. and Schmidt, S.S. (eds) 1993, *The Measurement of Productive Efficiency*, Oxford University Press, New York.

——, 1993b, ‘Linear programming approaches to the measurement and analysis of productive efficiency’, Working Paper 94–393E, College of Business Administration, University of Georgia, Athens, Georgia.

Nakamura, A.O. and Warburton, W.P. 1996, ‘Developing high performance public sector performance measures’, [mimeo].

Nunamaker, T.R. 1985, ‘Using data envelopment analysis to measure the efficiency of non-profit organisations: a critical evaluation’, *Managerial and Decision Economics*, vol.6, no. 1, pp. 50–8.

Pestieau, P. and Tulkens, H. 1993, ‘Assessing and explaining the performance of public enterprises’, *Finanz Archiv*, vol.50, no. 3, pp. 293–323.

Pierce, J. 1995, Public sector performance measurement – A policy makers’ perspective. Presented to the Symposium on Public Sector Performance Measurement, University of New England Conference on Efficiency and Productivity, Armidale, November.

——, and Puthucheary, N. 1997, Using performance measures to drive change within the public sector: the NSW experience’, Presented at Leading for Outcomes 4, Kooralbyn Hotel Resort, Brisbane.

Rouse, P. 1996, *Alternative Approaches to the Treatment of Environmental Factors in DEA: an Evaluation*’, Department of Accounting and Finance, University of Auckland.

Salarian, J. and Jomini, P. 1994, ‘Total Factor Productivity, Annual User Charges, Capital Stocks and Cost Allocation: A Synthesis’, *The Australian Economic Review*, 3rd Quarter.

SCNPMGTE (Steering Committee on National Performance Monitoring of Government Trading Enterprises) 1992, *Measuring the Total Factor Productivity of Government Trading Enterprises*, Canberra.

—— 1996, *Government Trading Enterprises Performance Indicators: 1990-91 to 1994-95*, AGPS, Canberra.

SCRCSSP (Steering Committee for Review of Commonwealth/State Service Provision) 1995, *Report on Government Service Provision 1995*, AGPS, Canberra.

—— 1997, *Report on Government Service Provision 1997,* Vols. 1 & 2, AGPS, Canberra.

Seiford, L.M. and Thrall, R.M. 1990, 'Recent developments in DEA: the mathematical programming approach to frontier analysis', *Journal of Econometrics,* vol. 46, pp. 7–38.

Senge, P. 1990, *The Fifth Discipline: The Art and Practice of the Learning Organization*, Currency Doubleday, New York.

Smith, P. 1990, 'The use of performance indicators in the public sector', *Journal of the Royal Statistical Society,* vol. 153, no.1, pp. 53–72.

Tasman Asia Pacific 1996, Measuring the Productivity Performance of State Road Authorities: A Case Study of Queensland Transport, Report Prepared for the National Road Transport Commission.

White, K. 1993, *Shazam Econometrics Computer Program version 7.0*, McGraw Hill, Internet: http://shazam.econ.ubc.ca/

Zeitsch, J. and Lawrence, D. 1996, 'Decomposing economic inefficiency in base-load power plants', *Journal of Productivity Analysis,* vol.7, pp. 359–78.

——, ——, Salerian, J. 1994, 'Comparing like with like in productivity studies - apples, oranges and electricity', *Economic Record,* vol.70, no. 209, pp.162–70.

Victorian hospitals case study

Banker, R., Das S. and Datar S. 1989, 'Analysis of cost variances for management control in hospitals', *Research in Government Non-Profit Accounting*, vol. 5, pp. 269–91.

Burgess, J. and Wilson, P. 1993, 'Technical efficiency in veterans administration hospitals', in Fried, H., Lovell, K. and Schmidt, S. 1993, *The Measurement of Productive Efficiency — Techniques and Applications*, University Press, New York, pp. 335–351.

Eagar, K. and Hindle, D. 1994, *A Short Introduction to Casemix,* National Casemix Education Series no. 1, February, Department of Human Services and Health, Canberra.

HCS (Health and Community Services) 1994, *Casemix Funding for Public Hospitals 1994-95*, Victorian Department of Health and Community Services, Melbourne.

NHDC (National Health Data Committee) 1995, *National Health Data Dictionary, Version 4.0,* Canberra: Australian Institute of Health and Welfare.

NSW Treasury 1994, 'Applicability of data envelopment analysis for measuring service provision performance in Australia', Report to Steering Committee on Commonwealth/State Service Provision, 16 November [unpublished mimeo] Industry Commission, Melbourne.

SCRCSSP (Steering Committee for the Review of Commonwealth/State Service Provision) 1995, *Report on Government Service Provision 1995*, Industry Commission, Melbourne.

Seiford, L. 1994 'A DEA bibliography (1978–92)' in Charnes, A., Cooper, W., Lewin, A.Y., and Seiford, L.M. (eds) 1994, *Data Envelopment Analysis: Theory, Methodology and Applications,* Kluwer Academic Publishers, London.

Valdmanis, V. 1992, 'Sensitivity analysis for DEA models' *Journal of Public Economics,* vol. 48, pp. 185–205.

Queensland oral health case study

Banker, R. D., Das, S. and Datar, S. M. 1989, 'Analysis of cost variances for management control in hospitals', *Research in Governmental and Nonprofit Accounting*, vol 5, pp 268–91.

Charnes, A., Cooper, W., Lewin, A. Y. and Seiford, L. M. 1994, *Data Envelopment Analysis: Theory, Methodology and Applications*, Kluwer Academic Publishers, United States.

Chilingerian, J. A. and Sherman, H. D. 1990, 'Managing physician efficiency and effectiveness in providing hospital services', *Health Services Management Resources*, vol 3, no. 1, pp 3–15.

NSW correctional services case study

ABS 1995a, *Government financial statistics*, Unpublished data AGPS, Canberra.

— 1995b, *National Accounts*, cat. no. 5242.0, AGPS, Canberra.

— 1996, *National Accounts*, cat. no. 5242.0, AGPS, Canberra.

Farrell, M.J. 1957, 'The measurement of productive efficiency', *Journal of the Royal Statistical Society*, A CXX, Part 3, pp. 253–90.

Ganley, J.A. and Cubbin, J.S. 1992, *Public Sector Efficiency Measurement: Applications in Data Envelopment Analysis*, North Holland, Amsterdam.

Lovell, C.A.K. 1993, ‘Production frontiers and productive efficiency’, in Fried, H.O., Lovell, C.A.K. and Schimdt, S.S. (eds), *The Measurement of Productive Efficiency*, Oxford University Press, New York.

SCRCSSP (Steering Committee for the Review of Commonwealth/State Service Provision) 1995, *Report on Government Service Provision 1995*, Industry Commission, Melbourne.

Yaisawarng, S. and Vogel, S. 1992, ‘Input technical efficiency of local correctional agencies in New York State’, Unpublished manuscript, Union College, New York.

NSW police patrols case study

Afriat, S. 1972, ‘Efficiency estimation of production functions’, *International Economic Review*, vol.13, pp. 563–98.

Burgess, Jr, J.F. and Wilson, P.W. 1993, ‘Technical efficiency in veterans administrative hospitals’, Fried, H.O., Lovell, C.A.K. and Schmidt, S.S. (eds), *The Measurement of Productive Efficiency*, Oxford University Press, New York, pp. 335–51.

Carrington, R., Connolly, P. and Puthucheary, N. 1996, ‘Improving the performance of NSW government agencies’, Paper presented to the 25th Annual Conference of Economists, Australian National University, Canberra, 22–26 September.

Coelli, T. 1995, ‘A guide to DEAP version 1.0: a data envelopment analysis program’, [Mimeo], Department of Econometrics, University of New England, Armidale, New South Wales.

Darrough, M.N. and Heineke, J.M. 1979, ‘Law enforcement agencies as multiproduct firms: an econometric investigation of production costs’, *Public Finance* 34, pp. 176–95.

F④re, R., Grosskopf, S. and Lovell, C.A.K. 1985, *The Measurement of Efficiency of Production*, Kluwer-Nijhoff Publishing, Boston.

Grosskopf, S. 1986, ‘The role of reference technology in measuring productive efficiency’, *The Economic Journal*, vol. 96, pp. 499–513.

Gyapong, A.O. and Gyimah-Brempong, K., 1988, ‘Factor substitution, price elasticity of factor demand and returns to scale in police production: evidence from Michigan’, *The Economic Journal* vol. 54, pp. 863–78.

Gyimah-Brempong, K., 1987, 'Economies of scale in municipal police departments: the case of Florida', *The Review of Economics and Statistics,* vol. 69, pp. 352–56.

—— 1989, "Production of public safety: are socioeconomic characteristics of local communities important factors?", *Journal of Applied Econometrics,* vol. 4, pp. 57–71.

Judge, G.G., Hill, R.C., Griffiths, W.E., Luthepohl, H. and Lee, T., 1988, *Introduction to the Theory and Practice of Econometrics.* John Wiley and Sons, New York.

Levitt, M.S. and Joyce, M.A.S. 1987, *The Growth and Efficiency of Public Spending*, Cambridge: University Press, Cambridge.

McCarty, T.A. and Yaisawarng, S. 1993, 'Technical efficiency in New Jersey school districts', In Fried H.O., Lovell, C.A.K. and Schmidt, S.S. (eds), *The Measurement of Productive Efficiency*, Oxford University Press, New York , pp. 271–287.

NSW Government 1995, *Financial Statement*, Sydney.

NSW Treasury 1996, *Performance of NSW Government Businesses,* Sydney.

Pierce, J. 1995, 'Public sector performance management - a policy maker's perspective', Paper presented to the New England Conference on Efficiency and Productivity, University of New England, Armidale, New South Wales, 23–24 November.

SCRCSSP (Steering Committee for the Review of Commonwealth/State Service Provision) 1995, *Report on Government Service Provision 1995*, Industry Commission, Melbourne.

Wilson, P.W. 1995, 'Detecting influential observations in data envelopment analysis', *The Journal of Productivity Analysis,* vol. 6, pp. 27-45.

NSW motor registries case study

Aly, H.Y., Grabowski, R., Pasurka, C. and Rangan, N. 1990 'Technical, scale and allocative efficiencies in US banking: an empirical investigation', *Review of Economics and Statistics,* vol. 72, pp. 221–18.

Burgess Jr, J.F. and Wilson, P.W. 1993 'Technical efficiency in veterans administrative hospitals' in Fried, H.O., Lovell, C.A.K. and. Schmidt, S.S. (eds), *The Measurement of Productive Efficiency,* Oxford University Press, New York.

Doble, M. 1995 ‘Measuring and improving technical efficiency in UK post office counters using data envelopment analysis’, *Annals of Public and Cooperative Economics*, vol. 66, pp. 31–64.

Färe, R., Grosskopf, S. and Roos, P. 1995 ‘Productivity and quality changes in Swedish pharmacies’, *International Journal of Production Economics,* vol. 39, pp. 137–47.

Ferrier, G.D., Grosskopf, S., Hayes, K.J., and Yaisawarng, S. 1993, ‘Economies of diversification in the banking industry’, *Journal of Monetary Economics,* vol. 31,. pp. 229–249.

—— and Lovell, C.A.K. 1990 ‘Measuring cost efficiency in banking: econometric and linear programming evidence’, *Journal of Econometrics.* vol. 46, pp. 229–245.

Fried, H.O. and Lovell, C.A.K. 1994, *Credit Union Service,* Filene Research Institute, Madison.

——, ——, and Eeckhaut, P.V. 1993, ‘Evaluating the performance of US credit unions’, *Journal of Banking and Finance,* vol. 17, pp. 251-65.

Judge, C.G., Hill, R.C., Griffiths, W.E., Luthepohl, H. and Lee, T. 1988, *Introduction to the Theory and Practice of Econometrics*, John Wiley and Sons, New York.

Lovell, C.A.K., 1993a, ‘Production frontiers and productive efficiency’, in Fried, H.O., Lovell, C.A.K. and Schmidt, S.S. (eds) 1993, *The Measurement of Productive Efficiency*, Oxford University Press, New York.

National Road Transport Commission 1994, *Aggregate Performance Measures: A Case Study of Queensland Transport. Data Assessment and Proposed Methodology. Interim Report*, Working Paper 17.

Tasman Asia Pacific 1996, *Measuring the Productivity Performance of State Road Authorities: A Case Study of Queensland Transport*, Report Prepared for the National Road Transport Commission, Tasman Asia Pacific, Melbourne.