New South Wales
TREASURY

Office of Financial Management

Research & Information Paper

PERFORMANCE MEASUREMENT AND RESOURCE ALLOCATION

TRP 97-9 **November 1997**

Performance Measurement and Resource Allocation

Suthathip Yaisawarng

Union College
Schenectady, NY 12308
U.S.A.

and

NSW Treasury
Governor Macquarie Tower
1 Farrer Place, Sydney 2000
AUSTRALIA

Nara Puthucheary

NSW Treasury
Governor Macquarie Tower
1 Farrer Place, Sydney 2000
AUSTRALIA

Prepared for International Conference on Public Sector Efficiency
27-28 November 1997
The University of New South Wales, Sydney 2052, AUSTRALIA

General inquiries concerning this document should be initially directed to:
Suthathip Yaisawarng (Tel: 9228 4823) or Nara Puhucheary (Tel: 9228 5796) of NSW Treasury.
This publication can be accessed from the Treasury's Office of Financial Management Internet site
[http://www.treasury.nsw.gov.au/]. For printed copies contact the Publications Officer on Tel: 9228 4426.

Contents

Performance Measurement and Resource Allocation

Suthathip Yaisawarng and Nara Puthucheary[*]

1. Introduction

Several studies have documented that public and private sector organisations do not always use resources efficiently. The opportunity cost concerns have inspired numerous research projects addressing the causes of inefficiency and ways to improve efficiency. Suggested remedies range from reforms to the industries such as restructuring and deregulation to promote competition, to reforms to the actual 'firm' such as changing the services provided and management style.

Performance measurement tells us how well an organisation uses resources in production. It is an essential tool for both public and private sector organisations. In a well functioning market, price changes provide signals about the quality of and demand for goods and services. Although firms operating in these markets receive feedback on how they perform through indicators of profitability and market share, performance measurement approaches can provide information on the extent of potential improvement that may be possible. Governments generally provide goods and services where there has been market failure which cannot be adequately addressed by the regulation of markets. The general absence of price signals in such circumstances means that performance indicators are even more important to guide decisions to ensure that services are delivered to the community efficiently.

The NSW Government provides a range of services to the community through Government Trading Enterprises (GTEs)[1] and Budget Sector agencies. The NSW Treasury has initiated several pilot studies in collaboration with selected agencies to develop more comprehensive measures of performance such as total factor productivity (TFP) and data envelopment analysis (DEA) efficiency indices. Pierce and Puthucheary (1997) summarises the TFP and DEA studies and provides a general indication of how performance measurement may be used as a catalyst for change in organisational structure and as a means of improving agencies' accountability.

---

* The views expressed in this paper are the authors' and do not necessary reflect those of NSW Treasury or the NSW Government. The authors thank Kathy Kang for her constructive comments, patience, and encouragement. We also thank Phil Bickerstaff, Frank Jordan, Anthony Holland, and Roger Carrington for their insights. All remaining errors are ours.

[1] Some public services such as electricity and rail are provided by Government Trading Enterprises (GTEs) which receive most of their funding from user charges. The focus of this paper is on Budget Sector agencies rather than GTEs.

Typically, fiscal policy constraints make it difficult or impossible to supply all the services that are demanded. Governments are under increasing pressure to ensure the best use of funds in providing public services. The NSW government implemented program budgeting to this in 1986-87 (NSW Government (1997)).

The next phase in budget reform is the gradual introduction of performance-based budgeting which aims to tighten the relationship between resource allocation, the government's intended policy outcomes and the outputs produced by agencies.

This paper describes how an explicit link can be established between performance indicators and resource allocation, with an aim to enhance productivity. Specifically, the paper addresses two major issues. Firstly, how can information on performance be used to improve resource allocation across the operating units within each agency? This information can also be used to improve resource allocation across agencies. Secondly, what incentives are needed to produce the desired performance improvement?

The proposed mechanism for improving resource allocation developed in this paper, if implemented, would contribute to performance-based budgeting by providing an explicit link between resource allocation and efficiency. The paper describes a mechanism that enables individual service providers to obtain the efficient combination of resources needed to provide a predetermined level of services. The efficient combination of resources is the least-cost input combination, given the available technology and input prices. This may be hard to measure in the public sector, given that the input prices communicated to government agencies may not reflect their opportunity cost, owing to government decisions to subsidise certain public sector uses of capital. Additionally, the government is hampered by imperfect information on the efficient cost of providing the specified level of services. Given these constraints, NSW Treasury has focussed on technical efficiency in the belief that this approach has the scope to produce a better understanding of efficient costs in government service provision[2].

The paper uses DEA (Data Envelopment Analysis) technical efficiency scores as performance indicators. Since DEA is conducted at the individual unit level such as the hospital level, rather than at the agency level, for example, Department of Health, the paper approaches the problem of resource allocation from the bottom up, ie, the resource requirements are determined at the unit level and all requirements are summed to obtain the agency's funding requirement [3]. The focus is on the allocation of resources across units within an agency.

---

[2] A unit is technically efficient if it uses the least input quantities to produce a set level of outputs, given technology, ie., the input combination that is on the production frontier. The unit is allocatively efficient if it uses the frontier input mix that is the cheapest, given input prices. Therefore, if allocative inefficiency exists at the current mix of inputs, technically efficient input combination is not the same as the least-cost input combination. However, technical efficiency is a necessary condition for achieving the least-cost input bundle.

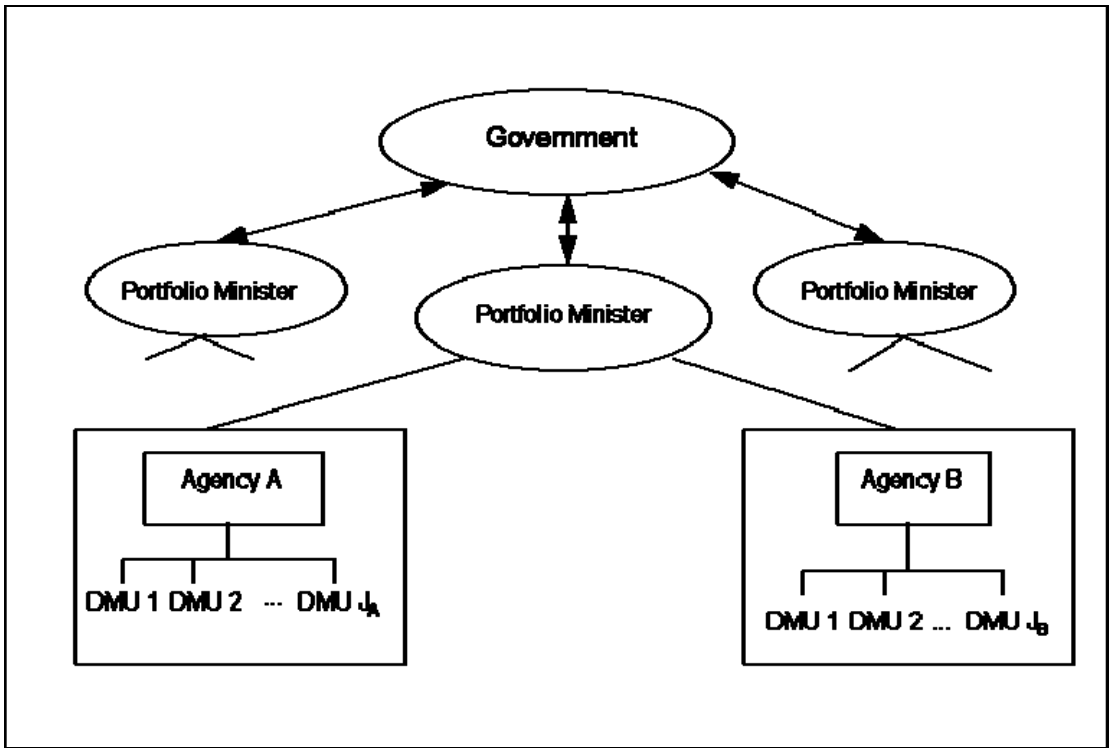
[3] Alternatively, one could approach the problem of resource allocations at the global level from the top down, i.e., the allocation of resources flows from the Government to the Ministers to Chief Executives, etc.

The rest of the paper is organized as follows. Section 2 presents the conceptual framework. Section 3 provides an overview of DEA. Results of DEA applications provide information for managers and for the allocation of resources. These issues are addressed in Sections 4 and 5, respectively. Specifically, Section 4 provides discussion on the use of DEA efficiency scores as an internal management tool. Section 5 provides detail on how DEA efficiency scores can be used to develop funding requirements, to set benchmarking targets for individual units, and to distribute resources. It includes incentive and monitoring mechanisms designed to encourage managers and agencies to achieve benchmarking targets. Section 6 raises some implementation issues.

2. Conceptual Framework

Budget Sector agencies provide a broad array of public services in New South Wales including health care, education, community services such as counselling and protection of children, corrective services, crime prevention, and vehicle registration. NSW Treasury advises the Government on the efficient allocation of resources to such services and monitors the funding of these services. Each agency distributes the allocated funds to different units and programs within the agency, in accordance with the Minister's priorities. Figure 1 is an outline of a typical institutional structure.

Figure 1: A Structure of Budget Sector Agencies



Agencies A and B may be considered as the Department of Health and Health Care Complaints Commission, for example. These agencies are under the Minister for Health. Agency A comprises J_A hospitals, which in DEA we referred to as DMUs (Decision Making Units).

The existing method of resource allocation across Budget Sector agencies[4] is largely a top down approach. In October of each year, the Treasurer provides an initial funding figure to each agency and Portfolio Ministers are invited to submit proposals for recurrent and capital funding allocations for their agencies. The recurrent expenditure proposal may include disputed amounts for maintaining existing programs, works in progress and additional funding for enhancements [5]. Requests for capital funds are for existing and new projects. During the budget process, the Government considers the funding proposals for all agencies and allocates the available resources to the Budget Sector agencies in consultation with the Minister of the agency and Treasury [6]. If total requests are more than the available resources, which is the norm, the Government may impose an across the board cut, request selected agencies to finance their services through internal savings, and/or reduce the services provided. At the end of the budget process, Portfolio Ministers submit a Ministerial Appropriation Summary, Agency Financial Summary and Program Statements to the Government. The Appropriation Summary for each agency includes three-year forward estimates. The forward estimates, with some adjustments, become the starting point for the budget process for the following years. The funds allocated to each agency and each program for the given budget year are published in the Budget Papers.

Under global budgeting, Portfolio Ministers have discretion to reallocate funds between and within programs and across agencies in their portfolio. In some cases, allotted funds may be transferred between years, subject to the Treasurer's approval.

---

[4] This paragraph is an excerpt from an introduction of the Budget Paper No.3: Budget Estimates 1997-98 (NSW Government (1997)).

[5] That is, provision of a new service or enhancement of an existing service.

[6] Information may also flow from the Government to agencies and vice versa during the negotiation process.

It should be noted that resource allocation across units within an agency is typically the responsibility of the agency itself and the relevant Minister. The current monitoring system by Treasury is at the agency level. Each agency submits to Treasury a monthly report on actual year-to-date expenses and revenues and projections for the full year. Individual units within an agency, for example hospitals, may provide services for a number of programs including ambulances, general community based services, acute health services, and mental health services. The monitoring system by Treasury at the agency level does not directly relate performance with resource allocation.

For Budget Sector agencies such as the Department of Health, the Department of School Education and the Police Service which comprise several units (hospitals, schools, police patrols) performing similar tasks, DEA can be used to analyse the performance of individual units within the agency. This paper suggests a way to use the DEA efficiency score to develop funding requirements for each unit or DMU [7]. The aggregate of funding requirements for all DMUs in an agency would be the amount of funding which the agency would require to provide the services if all units operate at best practice [8]. These funding requirements can be used as a basis for resource allocation between units within an agency and across agencies. The proposed method could be seen as an enhancement of performance-based budgeting which is gradually evolving in the NSW public sector.

---

[7] Although the discussion in Section 5 focuses on the funding requirement at the DMU level, this requirement may also be subdivided into programs within each DMU.

[8] If the funding requirement for each DMU is divided by programs, the agency's requirement will also be divided into programs comparable with the existing budget estimates.

3. An Overview of DEA

This section illustrates the concept of DEA technical efficiency [9] and describes a procedure known as a "tie-breaking technique" to rank technically efficient units [10]. DEA is a linear programming method that constructs a production frontier from the most productive observations in the sample. The constructed frontier represents *actual best practice* in the set of DMUs included in the study, rather than a *theoretically optimum* benchmark. The efficiency score for each observation is then derived with respect to the observed best practice frontier. In effect, DEA compares the performance of different observations performing a similar set of tasks with the best performers in the sample. Each observation under consideration is referred to as a DMU.

---

[9] There is an extensive literature in economics and operations research that discusses the mathematical formulation of DEA models. For a non-technical exposition, see Steering Committee for the Review of Commonwealth/State Service Provision (1997). A more technical discussion can be found in F@ re, Grosskopf, and Lovell (1994), and Lovell (1993).

[10] . There are a number of commercial software packages that can be used to compute the standard DEA efficiency scores. While these commercial packages cannot handle the tie-breaking technique, a computer code can be written to invoke any linear programming (LP) solver to perform the computation.

Consider an agency consisting of six DMUs: A-F. Each DMU produces the same amount of output ***y*** using different combinations of inputs, x_1 and x_2. Figure 2 displays the input combinations of these DMUs.

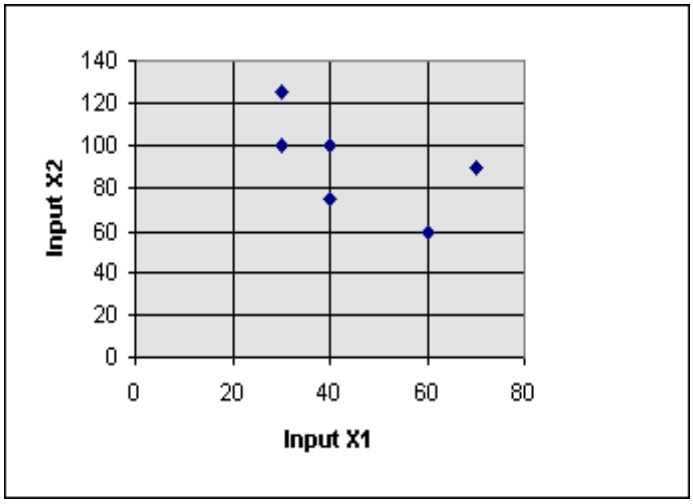


Figure 2: A DEA Frontier

Since all DMUs in this particular example produce the same level of output, a piecewise linear isoquant that envelops all the data points or the DEA frontier is ABCD with the vertical extension from A to the North and the horizontal extension from D to the East. DMUs A, B, C and D are technically efficient since they cannot proportionally (radially) reduce their current inputs (holding the ratio of input x_1 to input x_2 constant) while producing the chosen level of output. Each receives a radial technical efficiency score of one. However, DMU A can reduce input x_2 by the vertical distance AB, holding input x_1 and output y constant. An additional reduction in x_2 of AB is called the "non-radial slack." By eliminating the non-radial slack, the ratio of the two inputs is changed. To achieve best practice, DMU A needs to adjust how the inputs are utilised. DMUs E and F are inefficient, relative to DMUs A, B, C and D. Both have efficiency scores of less than one.

DMU E's radial technical efficiency is OE'/OE which is less than one, say, 0.88. This means that if DMU E utilises its inputs in the most efficient manner, given the existing technology, it could produce the same level of output with 88 per cent of its current inputs. Alternatively, if DMU E were productively efficient it would use 12 per cent less of both inputs. This, in effect, is the input combination at E' in Figure 2. Note that the ratio of input x_1 to input x_2 at E is the same as at E', implying that DMU E needs to focus on total input use rather than the allocation or mix of inputs to improve its efficiency. To become productively efficient, ie., to move toward E', DMU E has two peers which it should seek to imitate; namely, B and C. E' is a linear combination of DMUs B and C. DMU F's radially productive efficiency score is OF'/OF which is less than one. The target point for DMU F is F', a linear combination of C and D. Hence, DMUs C and D are peers for DMU F. Since F' lies closer to D than C, DMU F should put more weight on imitating DMU D than DMU C.

The standard DEA method produces an efficiency score bounded between zero and one for each DMU, where one indicates that the DMU is technically efficient. The method can be used to identify potential input savings via elimination of technical inefficiency. It also offers insights for managers who wish to improve the performance of their unit. However, a straight application of this method does not allow us to differentiate efficient DMUs, which is necessary for an effective performance-reward program. (Andersen and Petersen (1993) introduce a tie-breaking technique to rank efficient units, which Bogetoft (1995) uses to construct a workable performance-reward scheme.) Recall from Figure 2 that DMU A is technically efficient with a positive non-radial slack in input x_2, while DMUs B, C and D are technically efficient and cannot further reduce their current inputs. It is possible to conclude that DMU A does not perform as well as the remaining three. However, standard DEA efficiency scores do not distinguish DMU B relative to C or D.

To differentiate and rank technically efficient DMUs as identified by the standard DEA model, this paper applies the tie-breaking technique to compute radial technical efficiency scores for "super efficient" DMUs [11]. The tie-breaking technique computes a radial efficiency score for a DMU when the DMU is excluded from the sample. If the removed DMU is technically efficient in the standard DEA model, it will lie on or below the tie-breaking DEA frontier. Hence its efficiency score is at least one. The tie-breaking efficiency score tells the proportion of inputs by which the super efficient DMU could increase without being dominated by a linear combination of the other efficient DMUs [12]. The higher the efficiency score, the more efficient the DMU is.

---

[11] The tie-breaking technique is commonly referred to as modified DEA (MDEA). Lovell, Walters and Wood (1994) apply the tie-breaking technique to a study of US secondary education production. They compute unbounded efficiency scores and use them as a dependent variable in regression analysis. Wilson (1995) uses a similar approach to develop an outlier indicator. Readers are referred to these papers for mathematical formulations of LP problems to compute MDEA scores.

[12] If the removed DMU is technically *inefficient* based on the standard DEA model, there is no change in the frontier when a tie-breaking technique is used. Efficiency scores for these inefficient DMUs remain the same and are less than one.

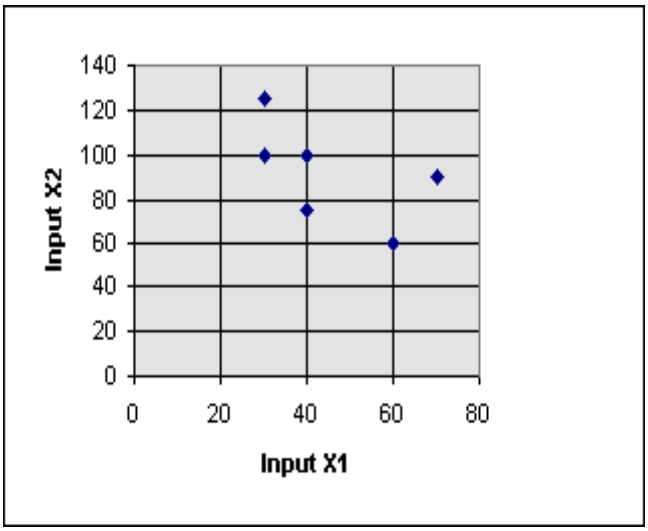


Figure 3: The tie-breaking DEA frontier for DMU B

Figure 3 illustrates the tie-breaking method for DMU B. Under the tie-breaking DEA, DMU B is excluded from the construction of the frontier. In this case, the frontier is formed by ACD, its vertical extension from A and the horizontal extension from D. DMU B is compared to B' which is the linear combination of A and C. The tie-breaking efficiency score for B is OB'/OB, say 1.1. This means that DMU B could increase its current inputs by 10 per cent to produce the pre-specified level of output ***y*** had DMU B used the linear combination of the next best technologies available. In that situation, it would lean more towards DMU A than C. DMU B is 10 per cent super efficient.

To compute the tie-breaking efficiency score for DMU C, the frontier is created from a sample of five DMUs; namely, A, B, D, E, and F. Efficiency score for DMU C is then compared to the frontier formed by ABD and their extensions. Repeating this for the remaining DMUs, one can rank all DMUs from the most efficient to the least. Note that the tie-breaking technique is an analytical tool for differentiating DEA efficient DMUs. The frontier DMU like B' in Figure 3 is less efficient than DMU B itself. Therefore, the result of the tie-breaking technique should be used strictly for the purpose of ranking DEA efficient DMUs and standard DEA scores should be used as a managerial tool and to establish benchmarking targets.

4. Use of DEA as an Internal Management Tool

The concept of the DEA efficiency measure discussed in the previous section can be extended to a multiple input and multiple output case using a linear programming approach. The DEA efficiency index is a composite performance measure which takes all inputs and outputs in the model into consideration. It is an internal management tool that could potentially reduce managers' probability of making inappropriate decisions (Norton (1994)). It should be viewed as a complementary tool to the existing partial measures which focus on the productivity of a specific input such as output per worker.

Recall that a DEA efficiency score indicates a proportion of current inputs that would be used if a DMU was productively efficient. It also suggests whether an input can be further reduced without reducing other inputs, given a current output level, and provides a set of peer weights used to form a target point for the inefficient DMU. As an internal management tool, information from DEA results could be used to set benchmarking targets for the inefficient DMU and to provide directions to improve efficiency.

For example, consider the inefficient DMU F in Figure 2. If DMU F becomes efficient, keeping the same input ratio, it will be at F'. Hence, F' is the target point for DMU F and the output-input ratios (ie, y/x_1 and y/x_2) at F' are the benchmarking targets for DMU F. For F to achieve the benchmarking targets, it may consult its peers, namely, DMUs D and E. A comparison of DMU F's output-input ratios with those of its peers reveals that DMU F should focus on the utilisation of resources at the current input ratio since both peers have higher respective output-input ratios. In other words, both DMUs D and E produce more output y per unit of input x_1 and per unit of input x_2.

A manager of an inefficient DMU should use the DEA results as a guide in developing a strategic plan to improve efficiency. The procedure begins with an internal investigation of the DMU to identify possible explanations for excessive use of inputs. This may identify situations specific to the DMU that are beyond the control of the manager but not captured by the DEA model. Excessive inputs due to these unique circumstances should be removed on a case-by-case basis from the derivation of the benchmarking targets [13]. Other cases of excessive use of inputs which reflect the abilities of the manager should be documented. The manager also consults the peer units to identify possible factors that might affect the DMU's efficiency.

---

[13] This process may be time-consuming, and the adjustments may be disputed. However, the result may still be worth the effort.

After investigating DMU F and its peers, it is important to compare managerial practice and principles across the organisations. In developing a performance improvement strategy for the inefficient DMU, DEA peer weights may be used to place emphasis on the qualitative and quantitative results from an investigation of each peer. Stoner, Collins and Yetton (1985) suggest four general methods for productivity improvement. They are product and process improvements, work and job improvement, employee motivation methods, and organisational change. If, for example, the investigation reveals that the inefficient DMU has a high turnover rate and that inexperienced staff appear to have low productivity per hour worked, the manager may consider a leadership style that enhances staff morale and motivates the employees to perform to their potential. The manager may also consider training programs to improve staff skills.

Box 1 following summarises the procedures outlined in this Section.

	Box 1 **Procedures for developing a strategic plan for performance improvement**
1	Use DEA efficiency scores as an indication to identify inefficient DMUs, their magnitude for the potential improvement, peers, and specific areas for investigation.
2	Conduct an internal investigation of the inefficient DMU to pinpoint causes for excessive use of inputs.
3	Consult peers on how they utilise their inputs and managerial practices.
4	Analyse the qualitative and quantitative information from the investigation taking into account the DEA peer weights.
5	Formulate a strategic plan for the inefficient DMU to implement. This may require a restructure of the organisation and a change in the culture

5. Use of DEA in Resource Allocation

Suppose that the Government has a sum of \$X billion [14] to allocate between ***R*** Budget Sector agencies. Agency ***r, r = 1, 2, ..., R***, consists of J_r DMUs. The \$X billion fund is broadly divided into two parts: $\$X_B$ billion for the provision of services and the remaining $\$X_W$ billion for monetary rewards.

DEA efficiency scores are computed for each DMU relative to others in the given agency. Since DEA does not provide a performance indicator at the agency level, it is not possible to allocate resources across the agencies based on agency performance. The DEA performance indicator is used to derive the funding requirement at the DMU level and the total requirements for all J_r DMUs could be used as the funding requirement for agency ***r***. In NSW, the budget negotiation process for the financial year, say 1998-99, begins in the preceeding October (1997) and therefore the performance indicators used in the negotiation would be for 1996-97 or earlier. [15]

---

[14] Several factors play important roles in the determination of the total resources available for the Government service provision including the Government's ability to raise revenue, its priorities, and borrowing. The process to determine the \$X billion is beyond the scope of this paper.

[15] The proposed funding requirement can be used in the adjustment process to reach the next year's budget and also to develop forward estimates in the future budgetary process.

The remainder of this section consists of three parts: Section 5.1 discusses a development of a DMU's funding requirement and the relevant benchmarking targets. The targets include both partial and composite performance measures. The proposals for all DMUs may be aggregated to derive the agency's funding requirements. Section 5.2 proposes a way to distribute monetary rewards based on the tie-breaking DEA scores. Section 5.3 presents a mechanism for monitoring the performance of each DMU.

5.1 Development of a DMU's Funding Requirement and its Benchmarking Targets

The DEA efficiency score may be used in developing a funding requirement for each DMU. If input prices do not reflect the opportunity cost of production due to government subsidy, technical efficiency should be used as the performance indicator. [16]

Consider a three-year model. Suppose that the information on performance [17] for year 1 becomes available in year 2 prior to the beginning of the budgetary process for year 3, which takes place in year 2. First, suppose that the level of service provision for year 3 is expected to remain at year 1 level, where efficiency scores for all DMUs are available. To deliver the services at the efficient level, each DMU may be allowed to request funding equivalent to the cost of the technically efficient or the best practice input combination. For example, DMU A has a budget of $300,000 and receives a technical efficiency score of 0.92 in year 1. In year 3, the year for which the funding requirement is being developed, DMU A would be expected to use the best practice input combination in providing the services and therefore would be given a budget of 0.92 x $300,000 or $276,000. [18, 19]

---

[16] If input prices are available, a DEA model for cost minimisation can be solved and the least-cost input combination for each DMU, given its input prices and the existing technology, can be identified.

[17] Information on performance for the most recent year available should be used.

[18] An implicit assumption here is that input prices, whatever level they may be, remain constant as in the previous period. Inflation which raises the general price level can be incorporated if desired. However, the principle of allocating resources remains unaffected.

[19] This may not be the minimal budget if it is possible to further reduce some inputs through elimination of positive non-radial slacks in those inputs. Since a non-radial input slack is measured in the same unit as the unit of input, incorporating it in the resource allocation process requires that the input price be known. When DMU-specific input prices are not available, the average price may be used to estimate the potential cost savings and to adjust the funding requirement.

Next, we allow the demand for existing services to change. The additional services should be provided using best practice input levels. For example, suppose that the demand for DMU A's existing services is expected to increase by 10 per cent. DMU A should receive additional resources of $27,600 to accommodate growing demand. If there is demand for new services, this would be classed as an enhancement under the current NSW Budget System. Its estimated cost would be included in the funding requirement. For illustration purpose, suppose that the estimated cost for the new services is $35,000. In this example, DMU A needs $338,600 ($276,000 + $27,600 + $ 35,000). Adding the funding requirements for all $\boldsymbol{J_r}$ DMUs results in total resources needed by the agency ***r***. This is the *least* amount of funds that the agency will bid for a share of $\$X_B$ billion available resources.

The Government allocates available resources to all Budget Sector agencies based on their proposals and taking into account the Government's priorities. If the total funds requested are less than the $\$X_B$ billion available, the Government may invite Ministers to submit proposals for new programs, new investment, or technology development. Alternatively, the Government may decide to increase the incentive pool (ie., $\$X_W$ billion), repay debt or reduce taxes.

When the total funds requested exceeds the available $\$X_B$ billion, which is usually the case, the Government may request agencies to revise their proposals or prioritise their requests. Each agency may request all DMUs under its responsibility to revise their requests, accounting for their priorities and the community's needs. The process continues until total funding requests match the available funds (which may exceed the initial amount of $\$X_B$ billion if the Government has the flexibility to increase the fund for public service provision). This process is iterative, and

the Government makes the final allocation, taking the community demand in each location and constituency into account.

Alternatively, the Government may impose an across-the-board cut, or ask agencies to increase efficiency. Some agencies will receive less than their request and some may experience larger cuts than others. In this case, agency ***r*** faced with the cutback in its funds, may apply an across the board cut to all J_r DMUs. Since this paper suggests the initial funding requirement be prepared according to best practice performance, each DMU receiving funds less than the requirement will need to revise its level of service provision. The services may be curtailed and the implementation of new services may be delayed until funding is available. This paper suggests that the adjustments in the level of services at the DMU level be approved by the responsible agency and the total adjustments for the agency be approved by the Government. [20] The adjustments may take into consideration the potential benefits from capital investment such as technology advancement, which may be realised during the budget year. The proposed process may create pressure for DMUs to attempt to use the available resources efficiently since they are discouraged from deliberately reducing their outputs. [21]

Once the level and nature of service provision are determined, a set of benchmarking targets is developed for each DMU. The benchmarking targets should include both partial and composite measures. [22] The partial measures refer to various output-input ratios based on the best practice input combination. These partial measures could be used to monitor the performance of the agency and DMUs during the budget year. If the DMU's performance falls short of the target ratios, there may be enough time for the DMU to make necessary adjustments so as to achieve the target at the end of the budget year or for the agency and Treasury to be aware of the need for extra funding. A composite performance measure such as DEA takes multiple aspects of the DMU's activities into consideration. Unlike the partial measures that permit managers to direct their attention to increase the productivity of specific inputs in compliance with benchmarking targets, it is difficult to predict or manipulate the DEA efficiency scores. A DEA efficiency score depends on the performance of other DMUs within the agency and is computed at the end of the budget year.

---

[20]. The agency may first classify all DMUs into several groups according to their performance and impose across the board cut to each group. DMUs in the most efficient group receive the least cut and those in the least efficient group receive the most cut. Each DMU may then propose changes in its level of services subject to the given resources.

[21]. However, budget cuts often result in degradation of service quality, eg, longer waiting time for service.

[22]. Partial benchmarking targets serve as a cross control mechanism to minimise the possibility of game playing between agencies and the Government. If the output-input ratios at the frontier for year 3 are less than the targets which are based on technology in year 1, it may suggest a possibility of a conspiracy to push the frontier "inward" and a need for investigation. However, it is unlikely since it involves sustaining collusive behaviour in an environment where there are rewards for DMUs who do not collude. (An incentive reward is discussed in Section 5.2.)

Some caveats associated with the suggested framework should be acknowledged. Firstly, a DEA model may not capture all aspects of DMUs' activities due to data unavailability or a mispecification of the model. Secondly, DEA does not allow for noise or errors in the data set. Additionally, some DMUs may be exceptionally productive in the given year because of luck, whereas others may experience hardship which prevents them from performing at their potential. These unexpected circumstances, in particular the latter case, should be taken into account in developing funding proposals and benchmark targets. Thirdly, the framework assumes that inefficient DMUs can improve their efficiency during the budget year. In reality, it may take more time for DMUs to adjust and learn how to produce at best practice. Fourthly, uncertainty and incomplete information exist at the time the demand for services is forecast. If the actual service outputs exceed the forecast and the DMU is able to produce at that level without additional resources, it will not only meet the benchmarking partial measures but exceed them.

Alternatively, the DMU may need additional funds. In calculating the amount of excess funds, allowance should be made for costs that are not variable in the short run.

Keeping these caveats in mind, this paper suggests that the benchmarking targets be specified as a range, instead of a point. For example, suppose that DMU A's efficiency score in year 1 is 87.5%. If the agency uses a 5 per cent band, the target efficiency score for DMU A in year 3 should be set between 0.83 and 0.92. At DMU A's best practice input combination (after adjusting for non-radial input slacks), various output-input ratios are computed. These figures may be used to construct ranges for the partial benchmarking targets. Alternatively, they may be used as lower bound for target ratios. Recall that DEA gives a list of best practice peers for each inefficient DMU. The corresponding output-input ratios for each peer can be computed. The maximum of each output-input ratio across the peers may be used as the upper bound for DMU A's benchmarking target. [23]

If the DMU's output-input ratios meet the benchmarking targets, the DMU will have a positive balance in its budget. If the DMU is permitted to keep the remaining funds in its budget to be used at its discretion, [24] it may have a stronger incentive to improve its performance. In addition, a performance-based incentive plan discussed below may be put in place to encourage efficient behaviour.

5.2 An Incentive Plan

Bogetoft (1995) presents a theoretical model for productivity improvement that incorporates incentive issues. He derives a set of production plans designed to motivate agencies to increase performance and measures the costs associated with these plans. He concludes that standard DEA scores provide an incentive for DMUs to do just as well as others, whereas the tie-breaking DEA scores [25] motivate the DMUs to do better than others. The DMU's rewards or penalties could be linked to the tie-breaking efficiency scores in a linear fashion.

If there is a monetary incentive payment (reward) to be given for benchmarking performance, it should be allocated based on the tie-breaking DEA scores. Let ME_k be the tie-breaking DEA efficiency score for DMU k, k = 1, 2, ..., N, where N is the number of DMUs that are productively efficient based on standard DEA model, and w_k is a size proxy for DMU k. DMU k's monetary reward is computed as follows: [26]

$$\text{DMU}_k\text{'s reward} = \frac{\text{Total monetary rewards}}{\sum_{i=1}^{N} w_i ME_i} \times w_k ME_k$$

[23] There is a trade-off between each pair of inputs which affects the output-input ratios. The output-input ratio beyond the minimum benchmarking target can be achieved if the DMU adjusts its input mix and uses its inputs to the fullest extent, or if there is a technological advancement. Improvement in technology permits the DMU to produce more output with fewer inputs.

[24] The remaining funds fall into the "Funding Transfers" category of the existing budget procedure.

[25] Bogetoft (1995) refers to tie-breaking DEA scores as modified DEA scores.

[26] Other linear formulae for a performance linked rewards are possible and the formula may vary across agencies.

Where total monetary rewards is $\$X_W$ billion discussed at the beginning of Section 5. Each DMU may use the incentive payment to provide some services that would not otherwise be funded during times of tight fiscal restraint. It may also use the reward to support training programs and staff workshops. If government policy permits, the incentive payment could be used to finance wage increases for managers and staff of the relevant DMUs.

Under the proposed method, the incentive payment is part of the total budget for a given DMU. Although the DMUs should have some discretion in the use of the reward, the use of the additional money must be included in the performance agreement to be approved by the agency and the Government. This is to ensure the reward is used to motivate staff to improve efficiency.

5.3. A Monitoring Mechanism

To monitor the DMUs' performance, the agency may require DMUs to submit interim reports several times throughout the agreement period and a final report at the end of the budget year. The interim report should contain target partial measures and actual partial measures to date, explanations for any discrepancies, and proposed actions for remedy. A final report may contain similar information to the interim report and may also include data necessary for conducting DEA efficiency analysis. A summary of these reports may be included in the agency's annual report currently submitted to the Treasury. This would enable Treasury [27] to monitor the use of resources to deliver services.

This paper suggests that the agency take responsibility in designing the format of interim and final reports. The standardised reports may require less time to complete which will encourage managers to complete them on time. Further, it will be both salutary and relatively easy for the agency to prepare summary reports for Treasury. It is recommended that deadlines be set and strictly adhered to so as not to delay the evaluation process. The DEA study may be conducted by agencies in collaboration with Treasury. [28]

---

[27] The existing monitoring system by Treasury is at the agency level and concerns whether agencies will be able to stay within their budget.[28] Over the past four years, Treasury has developed DEA studies for a number of agencies and provided guidance on how to use the results of the studies.

[28] Over the past four years, Treasury has developed DEA studies for a number of agencies and provided guidance on how to use the results of the studies.

6. Implementation Issues

A possible new procedure for resource allocation across units within an agency using DEA is introduced in this paper. The suggested procedure modifies the existing procedure by explicitly using a DEA efficiency index in resource allocation, enhancing productivity and efficient use of resources. This method can also be used to determine the funds required by an agency. This procedure is not part of NSW Government policy.

Further consideration of the procedure must take into account some caveats associated with DEA models, imperfect foresight of demand, changes in technology, and asymmetry of information between the parties involved as discussed in Section 5.1. The procedure may be implemented on a trial basis at selected agencies, may be adopted by some agencies on their own accord, or may involve selected agencies submitting a detailed report on the allocation of resources across their units and the resulting performance impact. The agency's report may be used as additional supporting documentation for current funding bids. The close monitoring system may be used to correct unforseeable problems associated with implementation. The method may be gradually implemented over a number of years, and continuing improvement in data collection and refining DEA models could be done in a parallel fashion.

The procedure is not applicable for agencies which do not have several units performing similar tasks. Such agencies may need to develop other performance indicators for use in performance-based budgeting.

This paper suggests that each DMU's funding requirement be consistent with the best practice input combination for that DMU, given existing technology. The total funding requirement for all DMUs in a given agency may be used as the agency's total funding requirement. It should be noted that incentives exist for DMUs and agencies to bid for funds in excess of requirements. Hence, the funding requirements and bids are unlikely to be identical. It should also be noted that setting benchmarking targets for selected services may increase the risk of goal displacement. For example, managers may devote their efforts to achieve the benchmarking targets (defined in terms of outputs) and ignore the government's desired outcomes. Intended outcomes should also be monitored. [29] All available resources may also come to be used to provide the current targeted outputs consistent with short-term rather than long-term policy goals. A shift in the use of resources within a given unit, from investment in advanced technology and human capital to the provision of routine outputs, may come to impede growth and development in the industry. Further, physical resources may not flow into needed areas unless some support programs are implemented. [30]

The proposed procedure may promote effectiveness and efficiency in delivering services. However, Stewart (1993) points out that establishing benchmarking targets requires that the quantity and nature of services be specified *a priori*. This may limit flexibility in responding to a change in the environment. To overcome this potential problem, flexibility must be built into the budgetary process by permitting each party to request changes. To avoid an unnecessarily lengthy process, the maximum duration for negotiations and possible actions to be taken if the process cannot be completed may need to be specified.

---

[28] Outcomes are already considered during the NSW budget process. Note that several factors outside the control of the agency may affect outcomes. For example, whether a released prisoner reoffends is affected by public housing policy, hospitals, community services as well as the performance of corrective services (prisons).

[30] Macdonald (1996) discusses the redistribution of resources in the health care industry. He stresses that physicians and nurses may not be willing to accept jobs in the growth areas unless the problems associated with medical culture and staffing are solved.

References

Andersen, P. and N.C. Petersen (1993) "A Procedure for Ranking Efficient Units in Data Envelopment Analysis," *Management Sciences,* 39, 1261-1264.

Bogetoft, P. (1995) "Incentives and Productivity Measurements," *International Journal of Production Economics*, 39, 67-77.

Fä re, R., S. Grosskopf, and C.A.K. Lovell (1994) *Production Frontiers*, Cambridge University Press.

Lovell, C.A.K. (1993) "Production Frontiers and Productive Efficiency," in *The Measurement of Productive Efficiency,* H.O. Fried, C.A.K. Lovell and S.S. Schmidt, eds., New York: Oxford University Press.

Lovell, C.A.K., L.C. Walters, and L.L. Wood (1994) "Stratified Models of Education Production using Modified DEA and Regression Analysis," in *Data Envelopment Analysis: Theory, Methodology, and Application*, A. Charnes, W.W. Cooper, A.Y. Lewin and L.M. Seiford, eds., Boston: Kluwer Academic Publishers.

Macdonald, G. (1996) "There are no resources to redistribute," *Sydney Morning Herald*, September 10.

New South Wales Government (1997) *Budget Estimates 1997-98*, Budget Paper No.3, Volume 1.

Norton, R. (1994) "Economics for Managers: Which Offices or Stores Really Perform Best? A New Tool Tells," *Fortune*, October 31, 38.

Pierce, J. and N. Puthucheary (1997) "Using Performance Measures to Drive Change Within the Public Sector," New South Wales Treasury Research and Information Paper no. TRP 97-3.

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.

Stewart, J. (1993) "New Management Relationships in the Public Sector: The Limitations of Government by Contract," *Public Money and Management*, July-September, 7-12.

Stoner, J., R. Collins and P. Yetton (1985) *Management in Australia,* Sydney: Prentice-Hall of Australia Pty Ltd.

Wilson, P. (1995) "Detecting Influential Observations in Data Envelopment Analysis," *Journal of Productivity Analysis*, 6(1), 27-45.