# Chapter 3

## 3.0 DEVELOPING DEA MODELS FOR THE EVALUATION OF EFFICIENCY IN MALAYSIAN PUBLIC UNIVERSITIES

### 3.1. Introduction

This chapter will begin with a discussion on some of the common issues pertaining to the development of DEA models. This includes the appropriate selection of input and output variables so as to reflect the prevailing production technology of DMUs. Other issues discussed are the conceptual problems regarding teaching and research input/output measurement, and the quality differences in resources used as well as the outputs produced by different DMUs. Despite the fact that there are conceptual problems in measuring public sector efficiency which includes measuring output, describing the production technology and recognizing quality differences in service provision (Madden & Savage, 1997); developing a DEA-based framework model for evaluating the efficiency of Malaysian universities is still worthy due to the fact this method of study is the first of its kind undertaken in the Malaysian context. In this chapter, these issues will be addressed with close references to the previous empirical studies to justify for the selection of input/output variables of the DEA model for measuring the efficiency of the Malaysian universities.

### 3.2. Some underlying issues

In many empirical studies, controversy often centers on the areas of model development. The first issue underlying the DEA approach concerns its basic assumptions on measuring. The DMUs are expected to be relatively homogeneous and to employ a common technology to convert inputs to outputs (McMillan *et al*, 1998). In short, the DMUs should do similar activities and have a similar physical attributes.

The second issue concerns how to develop a set of outputs and inputs for which the data can be collected for institutions. The choice of inputs and outputs of the institutions must reflect the production process. In other words, the model must capture the key inputs used and the outputs produced within the production process.

The third issue focuses on the number of inputs and outputs. DEA can be sensitive to the selection of inputs and outputs as indicated by McMillan *et al*, (1998). This is because as the number of variables increases, the number of DMUs deemed efficient and the inefficiency scores of inefficient units will typically increase. According to Avkiran, (2001) it is best to keep the number of variables to be less than or equal to one-third of the number of observations.

The final issue is the quality of input/output variables that measure the two major activities of teaching and research. Researchers differ in their views on the selection of input and output variables as they consider some variables to incorporate better form of quality than others. For example, some argue that the quality of research output can be incorporated through publications. According to Johnes&Johnes (1995), publication counts do approximate the 'impact' (thus measures) the 'quality' of a piece of research.

McMillan *et al* (1998), though, suggested that the amount of federal grant support on research awarded to each university is a proxy which incorporates the quality in research output. Flegg *et al* (2004) argues for the likeliness of research income to reflect the perceived quality and provides a more up-to-date picture of research output compared to publications.

Salerno's (2003) observation is that those who favor journal articles as research output would argue that research income neglects the quality aspect of research while those who favor research expenditure claim not all research output is in the form of journal articles.

### 3.3. Modeling of the University Production Process

Public universities in the sample are to be relatively homogeneous and employ common technology to convert inputs into outputs. The universities use human and physical resources to conduct teaching and produce research. They operate in a similar setting where basic educational facilities like classrooms, libraries, computer labs, etc. are made available.

In modeling the universities production process, the identification of inputs and outputs of the 'production process' is of great importance. Beasley, (1995) stated that the selection of inputs and outputs of a university requires much concern to be put on a conceptual view of what are the inputs and outputs of a university. The specification of input and outputs measures should be comprehensive as they should fully measure the activities of the organizations under evaluation and should be operationally meaningful in the sense that they should be commonly used (Bowlin, 1999). Two main activities have been commonly and universally recognized as the underpinning aspects or dimension of higher educational quality assessment i.e. teaching and research activities (Green, 1994).

Hence, the choice of inputs and outputs should reflect the production process of these two major activities. Some examples of inputs to a university are operating expenditures, capital expenditures, and academic and supporting staff. While outputs can take the form of student enrolments or graduates, publication counts, citations, or other proxies like the amount of research income or research expenditure.

To develop a DEA model for evaluating the efficiency of Malaysian public universities, we select the inputs and outputs of the production process based on the previous empirical studies. Consideration is also given to the availability of data, an issue raised by Beasley (1995). These forms the rationale for choosing the variables.

### 3.3.1. Input Variable

Two most commonly selected inputs for DEA university efficiency models are the number of academic staff and operating expenditures. Out of nine university efficiency studies, listed in Appendix 2.1, four have specifically employed the staff number (a combination of either academic, non-academic and/or researchers) as the input measure to their efficiency model (Johnes & Johnes, 1993 and 1995, Avkiran, 2001, Madden *et al*, 1997, Abott *et al*, 2003). Avkiran (2001), for example, argued that universities employ people to produce enrolments, and generate research output, thus he utilised both academic and non-academic staff as input measures.

Antanassapoulos and Shale (1997), in their outcome-related efficiency model used the number of academic staff, research income and expenditure as the resource variables to reflect the level of support given by individual universities towards teaching and research activities. The above studies calculated technical efficiencies of universities. McMillan *et al*, (1998) conducted a cost efficiency study which used total expenditures in the input specification. They included two separate expenditure categories, one categorized as the total expenditures less faculty salaries and, another category of total expenditure plus sponsored research expenditures. Some cost-based measures also use library expenditures or physical investment expenditures as the capital inputs (Salerno, 2003).

In our study, we will use operating expenditure as the single input variable. Operating expenditure is defined as the expenditure incurred in running the education operation. Capital expenditure is not employed as one of the input variables due to lack of data. Although capital expenditure is not employed, operating expenditure do provide a comprehensive input measure as it generally captures all the key resources used by the universities. The operating expenditure data employed for this study consists of emoluments, supply & services, allowance & fixed charges and other expenditures (MOEb, 2001).

The other reason for choosing cost efficiency measure is to see the potential reduction in operating expenditure. In addition, there is a need to reduce the public funding on higher educations, which is of paramount importance. This issue has been discussed in Chapter 1 page 6. Therefore, the DEA model employed in this study is a cost-efficiency model.

A cost efficiency measures is opposed to technical efficiency measures which routinely employ physical inputs as in the case of number of academic staff. Hence, operating expenditure is the more appropriate cost-based variable to be used for the input orientated model. Orientation of the model will be discussed in great detail in the section 3.3.4 of this chapter.

### 3.3.2. Output Variables

In terms of the output measures, a university's output should be defined primarily according to the services it provides in terms of teaching, research, consultancy and other educational services (Flegg *et al* 2004). Thus, aspects of a university's activities are generally captured via teaching and research output.

#### 3.3.2.1. Teaching Output

It is evident in nearly all empirical studies of higher education, *teaching* output is almost exclusively proxied by the physical headcounts of full time equivalent enrolments or the number of degrees. Avkiran (2001), who adhered to the view that universities employ people to produce enrolments and generate research output, thus, chose student enrolment and research quantum as output in his performance model. Anthanassapoulos & Shale (1995) used the number of successful leavers as the output variable as such measure would give an insight to into how effective the universities are, with the given resource allocation and the abilities of their students to achieve the outcomes as graduates. This is equivalent to the number of graduating students (Madden & Savage, 1997) or the number of degrees awarded (Calhoun, 2003).

However, McMillan *et all* (1998) argues against the number of degrees awarded or number of graduating students. According to them, degrees awarded measure completions and level of accomplishment or extent of learning but they neglect the education of those who attend but do not graduate and do not recognize differences in the length of degree programs. In contrast, Madden & Savage(1997), support the use of graduating student number as one of the output specification by stating that it is a simple way of incorporating quality into teaching as such higher quality teaching will produce a larger number of graduating students.

In this study, we use the number of student enrolments as the output of teaching variable as the data for the number of graduating students was incomplete. This is because the newly established universities i.e. KUIM and KUTKM do not have students who have completed their studies in 2001. The undergraduate student variables include the number of students who enroll in the Diploma, Advanced Diploma and Bachelor studies, whereas the graduate student variables comprise of the number of enrolments for Master's and Doctoral programs.

### 3.3.2.2. Research Output

Based on previous literature, publication counts and research income are considered to be the output of research (Van de Panne(1991), Sinuanu-Stern et al (1994) and Tomkins & Green (1988) (Tomkins & Green, 1988; Ahn et al, 1988; Beasley, 1995). However, there are some quality and conceptual problems with these outputs. McMillan *et al*, 1998 pointed out that publication count is the best indicator but it is not easily obtainable and in some cases, maybe unreliable. In addition, the publication count analysis is commonly fraught with difficulty, particularly, due to time lags, (Johnes & Johnes, 1995).

Publications, also pose difficulties in terms of dealing with different types of publications (for instance, articles, books, proceeding, translations, etc). Salerno (2003) stated that the inability to exercise adequate quality control gives rise to another issue to be considered when evaluating the efficiency estimates for institutional comparability as quality control is likely to distort estimates of institutional efficiency.

Another point highlighted by Johnes & Johnes (1995) is that publication counts and research income does not capture the vast differences in resources used by different universities in producing the research output. For example, some universities may have more PCs in their staff offices, good library facilities, availability of research assistants and low teaching load, while others may enjoy less of these luxuries. These factors would relate to the question of quality of the produced research output.

For this efficiency study, we employ publication counts and research income as the output variables. Publication count is the total number of publication in the international journals like AGRICOLA, Biological Abstracts, CAB Abstracts, Chemical Abstracts, IEEE, Medline, Science Citation Index, etc. (MASTIC, 2004). However, a drawback identified is that the publication counts omit publications in the arts, humanities, and social sciences. Because there is a time-lag in obtaining publications, the publication counts fluctuate from one year to another. We smooth the data by taking an average over three years i.e. 1999-2002.

Research income for this study is made up from research grants from the federal, state and other local and foreign sources as well as the universities' own contribution. Federal research grants are the Intensification of Research in Priority Areas (IRPA)[i], fundamental and short term grants.

### 3.3.3. DEA Models Employed

The input and output data variables to be employed in this study are listed in Table 3.1 as shown below. We specify four models that will be used in evaluating the efficiency of Malaysian public universities i.e. DEA1 to DEA4. The specification of these models begins with the modification of the basic model (DEA1) and disaggregating the output variables whichever applies. The changes to the basic model are introduced individually or in combinations to establish whether there was any consistency between different DEA results depending on the input and outputs used. All the models measure cost efficiency as they specify operating expenditure as the single input measure.

**Table 3.1**

**Input and Output Variables for the DEA Models**

| MODEL | DEA1 | DEA2 | DEA3 | DEA4 |
|---|---|---|---|---|
| **Input** | | | | |
| Operating Expenditure(RM) | x | x | x | x |
| **Output** | | | | |
| Number of Undergraduate Enrolments | | | x | |
| Number of Graduate Enrolments | | | x | |
| Total Number of Student Enrolments | x | x | | |
| Research Income (RM) | | x | x | x |
| Publication Count | x | x | x | x |

**DEA1**

This model is the most parsimonious. It has a single cost-based input variable and two output variables. The output variables which represent both the teaching and research activities of the universities are publication counts and the aggregate number of student enrolments. The single variable for total number of student enrolment is the total figure of undergraduate and graduate enrolments.

**DEA2**

As for DEA1, DEA2 applies the same input and output variable plus research funding as an additional research output. This model assesses the universities' efficiency level based on two different research output measures, publication counts and research funding simultenously. This comparison is necessary because presumably, there are some DMUs that are not doing well in publications but sufficiently productive in research activities. Hence, they receive a large amount of research funding.

**DEA3**

DEA3 is similar to DEA2 as such both employ Research Funding and Publication Counts as the measure of output. The difference is that DEA2 specifies a single number of student enrolments as the input variable, while DEA3 employs the disaggregated figures. This is to reveal if the disaggregated enrolment variables will produce different efficiency level from the model utilizing a single enrolment variable.

**DEA4**

The final model specification attempts to examine the efficiency of research activities of the public universities. This model will assess how efficient these universities had been in producing research publications and obtaining funds.

### 3.3.4. Model Orientation and scale assumption

We chose the input-orientation for all of the models because the study is aimed at measuring the cost efficiency of the universities and the potential for cost reductions. As stated by Anthanasapoulos & Shale, (1997), input oriented model measures the justification of expenditure in producing the outputs. All of these models assume the CRS condition because quite a number of studies assume a condition of CRS in their studies. (Ying, C.N. and Sung, K.L, 2000; Madden & Savage, 1997; Johnes & Johnes, 1993). Moreover, at present, we do not have any evidence of the relationship between performance and size within educational institutions to assume a condition of VRS. In this sense, we, therefore, stick to our assumption of a CRS condition for all models.

### 3.4. Conclusion

There are a few issues that one has to consider in developing DEA models for higher education institutions. In selecting the input and output variables for this study, the conceptual and quality issues has been taken into consideration. The issue on quality of research output was resolved by employing the number of publication counts and research income. The conceptual issue of publication counts was handled by taking the average number of publication over three year period. In deciding the appropriate number of input and output to be used, the rule of thumb applied for all DEA models, is the sum of variables should not exceed one third of the sample size (Avkiran, 2001). Hence the sample size is 15 and the total number of variables in the most comprehensive model i.e. DEA3 is five.

Finally, of utmost important in developing DEA models for higher education institutions, one has to consider that the selected data variables must actually be available to be employed. The non-availability of complete and reliable data oftentimes hinders such analysis to be undertaken and this is exactly the point to raise in the concluding chapter under the policy implication of this study. But, we will first evaluate the efficiency of the Malaysian public universities which is outlined in the next chapter.

[1] IRPA stands for the Intensification of Research in Priority Areas Program under the Ministry of Science, Technology and Innovation (MOSTE), which is a scheme designed to encourage research institutions to make proposals for projects in preset priority areas (Alexander G, 1997).