AN INVESTIGATION OF TAFE EFFICIENCY

Peter Fieger NCVER

Abstract

This paper summarises the findings of a project investigating the relative efficiency of TAFE institutes across Australia. The paper uses Data Envelopment Analysis (DEA) to derive efficiency scores and regression modelling to identify factors that are predictors of efficiency. The paper finds that overall relative efficiency is quite high. There are economies of scales issues however in that small institutes tend to be relatively less efficient. Modelling of the data suggests that remoteness is the major predictor of efficiency with the analysis finding a strong negative relationship between the two.

Introduction

The purpose of this paper is to investigate the efficiency of individual TAFE institutes and to determine factors that could be predictors of their efficiency.

Over the past several years, the vocational education and training system has been the subject of considerable reform, including opening up the market for training. This by its very nature involves competition and policies such as 'user choice'[1] have been introduced. One of the reasons for this was to increase the efficiency of public vocational education and training organisations, including TAFE Institutes. As such, governments and the public are concerned that these institutes operate in an efficient manner.

Despite this context TAFE institutes are still quite highly regulated and in addition have community welfare obligations. For instance, they are expected to operate in thin markets such as is the case in remote areas. This will necessarily impinge on efficiency. Abbot and Doucouliagos (2000) noted that taxpayers maybe willing to trade-off some level of efficiency for other objectives such as access and equity. Having said this, we might expect that in a public sector context that there would be some consistency in efficiency in institutes due to the funding models in place.

Nevertheless, indicators of efficiency and factors leading to inefficiency can provide useful information for institutions on improving their efficiency. This information could also be used to guide funding models.

In this paper we examine three concepts of efficiency (discussed later). Basically however, efficiency can be defined as maximising a set of desired outputs for a given level of inputs. For this paper we use a technique known as Data Envelopment Analysis (DEA) to derive efficiency scores and then regression analysis to identify factors that may impact on efficiency. Furthermore, we attempt to determine whether there is a minimum size necessary to affect an institutions efficiency.

This paper progresses as follows. First, we comment on previous work done in this area and approaches used. Second, there is a discussion on the approach used for this study and the types of variables that are relevant for our analysis. The third section presents main findings in regards

[1] Under the user choice policy employers choose the training provider that is to deliver the off-the-job component of apprenticeships/traineeships. The purpose of this is to make VET more responsive to the needs of industry. (NCVER online VET glossary)

to efficiency and the sorts of factors that may impact on efficiency. The paper concludes with some observations on the findings.

Previous literature

There has been relatively little research that compares the efficiency of TAFE institutes across Australia. There have been earlier studies that have examined the efficiency of TAFE institutes from a state perspective (Abbot & Doucouliagos, 1998; Abbot & Doucouliagos, 2000), and also one that has compared the efficiency of New Zealand Polytechnics (Abbott & Doucouliagos 1999). There have also been studies that have compared the efficiency of Australian universities (Carrington, Coelli & Rao, 2004; Abbot & Doucouliagos, 2003; Abbot & Doucouliagos, 2009).

The Carrington et al (2004) study is perhaps the most comprehensive of the papers that have examined the efficiency of Australian universities or VET institutions. They found that the university sector is relatively efficient, but they also examined a variety of variables that can be predictors of efficiency. They broke these variables into quality variables (e.g., student satisfaction and proportion full-time employed post study) and environmental variables such as proportion of students who are Indigenous, who come from a low socio-economic background and who come from rural and remote locations. Their analysis found only two variables that were significant - location and proportion of students from rural and remote locations negatively influence efficiency. The approach taken by Carrington et al (2004) provided a good starting point as we are interested in factors that are predictors of efficiency as well as the efficiency scores.

Approach

We use Data Envelopment Analysis (DEA) to derive the relative efficiency scores of TAFE institutes and then regression modelling to test for factors that may be predictors of efficiency. Data Envelopment Analysis (DEA) is a technique that is used to evaluate the performance of a set of what are called Decision Making Units (Charnes et al., 1978; Cooper et al., 2004). It is a non-parametric linear programming technique that constructs a frontier over the data (Coelli et al, 2005).

The DEA approach provides various types of efficiency scores. Firstly, it provides information on technical efficiency. This can be defined as the maximum number of useful outputs obtained for a given set of inputs. Technical efficiency can be further broken down into constant returns to scale and variable returns to scale. Very simply put, constant returns to scale assumes that outputs increase in proportion to inputs. There is an assumption that scale of economies does not change with an increase in the size of the unit (institution) under investigation. This assumption can be naive in a lot of instances. Variable returns to scale does assume that scale of economies change with an increase in the size of the unit (either increasing or decreasing). This a more realistic assumption in many instances, including TAFEs. Using variable returns to scale, institutes of a similar size are compared for efficiency. The other measure of efficiency we use is scale efficiency, which is derived by dividing variable returns to scale and constant returns to scale. This shows is the extent to which an institution can benefit from returns to scale by changing the size of an institution to an 'optimal' size.

Strengths and limitations of the DEA technique

Previous research using data envelopment analysis outlines various strengths and limitations of the technique. These are summarised below.

Strengths

Agasisti and Johnes (2009) summarise some of the main strengths of DEA. Firstly, it is suited to a context where there are multiple inputs and multiple outputs, such as the case in business units or indeed an entire organisation. Secondly, using this method it is not necessary to impose a functional form on the process (i.e. an algebraic relationship between the variables). Thirdly, the method provides various analytical information. Apart from the relative efficiency scores, there is also information on peer groups (efficient units on the same part of the frontier as inefficient units) and slacks, which is information on the amounts of input/output that could be reduced.

Weaknesses

The method will always identify as efficient at least one of the cases under examination. In reality they may all be inefficient (Abbot & Doucougliagos, 2000). The choice and availability of inputs and outputs effect the efficiency scores obtained. Generally, inputs and outputs within the control of the business unit (or in our case institution) are chosen. However, the level of information required is often not readily available, so their can be a reliance on 'proxy' variables. Furthermore, the exclusion of a relevant input or output can bias the results.

Following on from the above, the method is somewhat 'data hungry' meaning that we need to restrict the number of inputs and outputs that are used. This leads to an over simplification of this issue. Coelli et al (2005) also discuss several other limitations to the method including the possibility of measurement error, the possible influence of outliers, and that the results can be misleading if environmental differences (between institutions) are not accounted for. Because efficiency is a relative measure, efficiency scores across different studies cannot be directly compared.

Data

The data we used for the analysis comes mainly from NCVER's national provider collection and NCVER's Student Outcomes Survey[2]. However, institute level financial data was obtained either from the institutes annual reports, the institutes themselves, or from the relevant state training authorities.

Variables used in the analysis

We use four categories of variables for our analysis - input, output, quality and environmental variables. These are discussed in turn.

Input variables

The input measure we use is expenditure on salaries, wages and related expenses, and other expenditure (excluding capital costs[3]). We do not want to use too many input variables given the relatively low number of observations (institutes). In the Carrington et al (2004) study, they used operating costs (academic and general staff salaries, and other expenses) as their main input variable. Abbott and Doucouliagos (2000) used two input measures in their study - total number of teaching hours (as a proxy for labour) and capital expenditures (as a proxy for capital stock). In their study on New Zealand Polytechnics, Abbott and Doucouliagos (1999) used three input measures - number of full-time equivalent teaching staff, number of full-time equivalent non-teaching staff, and value of fixed assets. Note how these variables are all related. In our study we use salaries and other expenditure as inputs, which is directly related to the number of teaching and non-teaching staff.

---

[2] We use the results of the Student Outcomes Survey for graduates of the year under examination (so for 2007 we use the results of the 2008 Student Outcomes Survey which is undertaken six months post course.

[3] We are not including capital costs here as they would distort the efficiencies. Capital costs are not easily linked to outputs for a given year.

Output variables

For this study we use number of full year training equivalents as a proxy for training outputs of an institution. These equivalents are split into trades and non-trades to reflect the higher costs associated with teaching the trades. Carrington et al (2004) also used equivalent full-time student units as a way of measuring quantity of teaching. They split their student load by science and non-science students for the same reasons. Abbott and Doucouliagos (1999, 2000) used student contact hours and number of full-time equivalent enrolments respectively as their output measure.

Quality variables

It is hypothesised that the quality of VET provision may relate to efficiency. The argument here is that it may cost more to provide quality provision, thereby decreasing efficiency. Carrington et al (2004) considered three quality measures in their study being student satisfaction with the course, average graduate starting salary, and graduate full-time employment. Abbott and Doucouliagos (2009) in their study of efficiency of Australian and New Zealand universities used data from the Australian Course Experience Questionnaire (on perceptions relating to generic skills) to obtain a rough measure of teaching quality.

For our purposes, this type of information can be obtained from NCVER's Student Outcomes Survey[4]. More specifically, we can derive the proportion of graduates[5] who were satisfied with the training they received and the difference between the proportion of graduates employed (full-time) before the course and after the course. We focus on the difference as many VET students are already in employment. There is a need to be aware however that employment is somewhat reliant on local economic conditions.

In addition we use load pass rate[6], achieved main intention of doing the course, and willingness of the student to recommend the institution of instruction to others as other quality measures. Load pass rate is an indication of the success of students and could be reflective of the teaching effort put in. Likewise, achieving main intention of the course and recommending the institute could also be reflective of teaching effort.

Environmental variables

There are also a variety of other variables that can affect the efficiency of institutes, known here as environmental variables. These can be divided into those relating to students and other variables.

Research shows (e.g. Krause 2005, Mills et al 2009) that certain groups of students, on average, do not fare as well as other students and so the cost of delivery may be higher for these students. Student characteristics that are relevant here include previous level of education, Indigeneity, disability, whether a rural or remote student and English or Non-English speaking background. Carrington et al (2004) used similar student characteristic variables in their study – proportion of Indigenous students, proportion of students from rural and remote areas, location of the institute, and proportion of students from a low socio-economic background. Upon review of our variables indicating percentage of students from remote areas, remote location and percentage of indigenous students we expected that these variables are highly correlated with each other. Closer analysis confirmed this suspicion with observed Pearson correlations in excess of .92. In order to impede the impact of potential multicollinearity in our subsequent regression model, we

[4] This survey, undertaken six months post completion of course, asks graduates and module completers a variety of information about their course and outcomes post course.

[5] We could look at the satisfaction of module completers but this increases the number of variables in our study. Furthermore, it can be argued that graduates could be more discerning in their assessment of satisfaction as they have completed the whole course.

[6] Load pass rate can be defined as all hours successfully completed (including recognition of prior learning) over all hours (including hours that lead to failure or withdrawal).

performed a principal component analysis with the aim of extracting the dominant underlying factor from these three variables. This factor, named 'remoteness indicator', accounts for 90% of the variance of the three analysed variables.

In another study Abbott and Doucouliagos (2009) found that the percentage of overseas students impacted positively on efficiency in Australian universities. We therefore hypothesized that a similar effect may be observed in TAFE institutes and thus included an environmental variable indicating the percentage of overseas students at individual institutions.

Another environmental variable quantifies part-time students. Institutions can incur extra costs with part-time students due to extra administration costs, posting out course work and out of hours teaching times. For our study we use proportion of part-time students in the institution as an explanatory variable.

In addition to these variables, we also use average hours per student. It could be expected that the higher the hours per student the more efficient in terms of resourcing. Finally, we include a variable on proportion of students who are apprentices or trainees. These are expected to be more resource intensive.

Analysis

Data Envelopment Analysis is used to determine technical (variable returns to scale) and scale efficiency (previously defined). This technique will use the input and output data discussed earlier.

The ordinary least square method of regression analysis is used to regress the quality and environmental variables in table 1 onto the variable returns to scale technical efficiency scores (derived from the Data Envelopment Analysis technique), in order to determine if any of these are predictors of efficiency. Given the small number of cases and relatively large number of variables we will use a p value of less than 0.1 as the cut-off for significance.

Findings

This section shows graphically the efficiency scores obtained by the Data Envelopment Analysis technique, and following this the results of the regression analysis. The efficiency results presented below are for 2007. We also calculated efficiency scores for 2008. Although there were a number of amalgamations of institutes in that year, the results of the efficiency analysis were roughly comparable to 2007 results, indicating a certain comparability of year on year efficiencies.

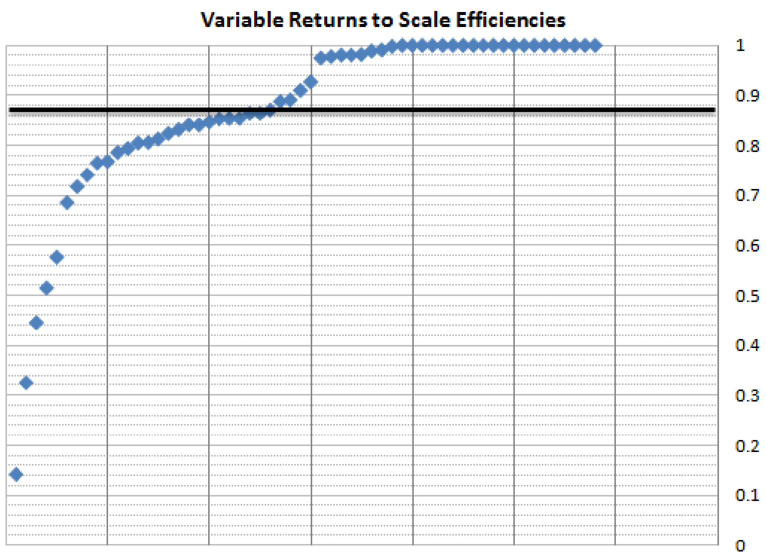
In our regression analysis we focus attention on the variable returns to scale efficiency scores because VRS efficiencies are not influenced by scale inefficiencies. Additionally, when investigating economy of scale issues we us constant returns to scale efficiencies, as we are here attempt to examine issues of institution size in relation to efficiency.

Efficiency scores

We calculated the mean constant returns to scale technical efficiencies for 2007 as 0.815, and found that 10 out of the 58 institutes could be considered technically efficient. In contrast, the lowest technical efficiency was 0.14. We notice that most institutes (36; 62%) fall within the 0.8 to 1 efficiency range. There are however a few institutes where the relative efficiency is quite low.

As mentioned above, the constant returns to scale efficiency calculations assume that TAFE institutes operate on an optimal scale. The history and geography of an institute may cause it to be not operating at an optimal scale. For this reason, variable returns to scale scores have been calculated to abstract from scale effects (see Figure 1). This results in a higher mean efficiency across the institutes (0.871 as compared to 0.815).

Figure 1



There are also twice as many institutes that are technically efficient as compared to the constant returns to scale assumption (20 as compared to 10). If we look at the example of one of the institutes in our study, the technical efficiency score was 0.62 under the constant returns to scale assumption, however, the institute was technically efficient (1.0) under the variable returns to scale assumption.

If there is a difference between variable and constant returns to scale efficiencies in any one institution, that indicates the presence of scale inefficiencies, and that inefficiency can be calculated by dividing constant returns to scale efficiency by variable returns to scale efficiency (Coelli, 1996). This scale efficiency indicates the potential gain from achieving the optimal size of that institution. As an example, one of the institutes has a constant returns to scale score of .785 and a variable returns to scale score of .864 resulting in a scale efficiency score of .908 - this is the efficiency of the institute if they were of optimal size.

Further to this analysis, we can determine whether an institution exhibits increasing (irs) or decreasing (drs) returns to scale properties. Increasing returns to scale can be interpreted as outputs increasing disproportionately more when inputs increase, and decreasing returns to scale mean that outputs increase at a smaller rate than increasing inputs. In our 2007 analysis of Australian TAFEs, 28 institutes display increasing returns to scale, 10 institutes have constant returns to scale and 20 institutes exhibit decreasing returns to scale properties.

The data envelopment analysis technique also provides information on the effective peers of inefficient institutes. These peers are technically efficient (using variable returns to scale) institutes on the same part of the frontier as the inefficient institute. Theoretically, these peers can be looked at in terms of how they go about their management and teaching so as to achieve improvements in efficiency.

Predictors of efficiency

The variable returns to scale technical efficiency scores determined by DEA were then regressed against quality and environmental variables in order to seek explanations for differences in efficiencies among institutions. The results are presented in table 1:

Table 1: Predictors of efficiency - Regression results

Variable	Parameter Estimate	Standard Error	t Value	Pr>|t|	Standardized Estimate
% Remoteness indicator	-0.1233	0.0276	-4.46	<.0001	-0.6959
% Overseas students	-0.0009	0.0083	0.11	0.9115	0.0161
% Achieved main goal	0.0043	0.0039	1.11	0.2732	0.1708
% Employed after v before	-0.0051	0.0077	-0.66	0.5095	-0.0876
% Satisfied with training	0.0067	0.0082	0.82	0.4173	0.0902
% Would recommend institution	0.0175	0.0130	1.34	0.1870	0.1997
% English second language	0.0027	0.0029	0.91	0.3697	0.1444
% No post school quals and no yr 12	0.0015	0.0020	0.75	0.4569	0.1017
% with disability	-0.0142	0.0099	-1.44	0.1573	-0.1892
% Part time students	0.0066	0.0117	0.57	0.5744	0.2396
% Apprentices	0.0002	0.0033	-0.09	0.9311	-0.0118
Load pass rate	0.0045	0.0035	1.30	0.2015	0.1593
Average hours	0.0012	0.0010	1.18	0.2425	0.4932
Intercept	-2.8933	2.1497	-1.35	0.1852	

The adjusted R^2 statistic for the regression model was 0.55 indicating that 55% of the variance in the technical efficiency scores is explained by these variables.

Only one of the variables is significant at the 10% level, being the remoteness indicator which is a negative predictor of efficiency. Carrington et al. (2004) in their study of Australian universities found location and students from rural and remote locations to be the only significant variables in their analysis. Our remoteness indicator is more comprehensive than that of Carrington et al. as it contains the shared variance of the remoteness, students from remote areas, and percentage indigenous students.

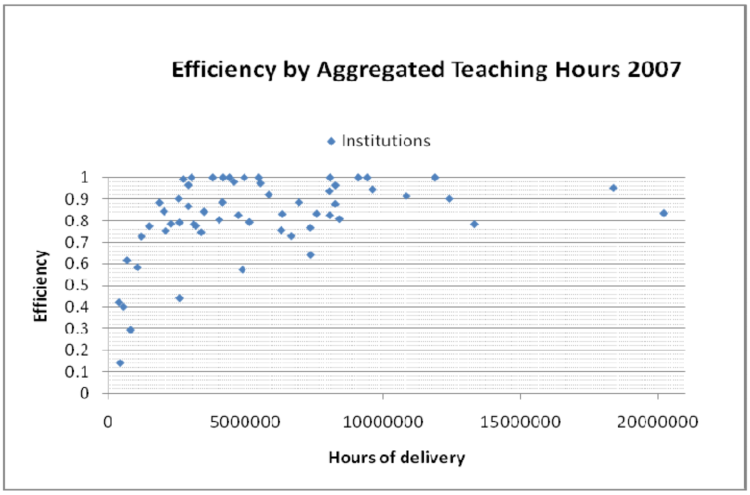
We hypothesised earlier that proportion of overseas students could be a predictor of efficiency. This was prompted by the result Abbott and Doucougliagos' study (2009), which found the percentage of overseas students to be a positive predictor of efficiency in Australian universities. However, this is not the case in our analysis of TAFE institutes and may be due to the lower overall percentage of overseas students at TAFE compared with universities.

We need to be careful in interpreting the results however. The relatively small number of TAFE institutes together with a relatively large number of explanatory variables can contribute to the insignificance of some variables. Insignificance of variables can also result from variables being highly correlated, as can sometimes be observed in models where highly correlated variables do lead to multicollinearity which makes it difficult to asses the impact of those correlated variables on the model.

Economies of scale issues

One of the prime issues that affects the efficiency of an institution is size. It can be reasonably expected that larger institutions are generally more efficient than smaller institutions. We have created a scatter plot (Figure 2) displaying constant returns to scale technical efficiency (CRSTE) versus teaching hours (as a measure of institution size), which demonstrates that this expectation is valid. It is clear that very small institutions are disproportionately found among institutions with low efficiency and it seems reasonable to assume that their small size prevents them from being more efficient.

Figure 2



It is thus of interest to determine what the minimum size of an institution is from where on its size is no longer an impediment to efficiency. Considering the graph in Figure 2 again, it appears that the pattern of efficiency follows a reciprocal function, with efficiency rising quickly with only a modest increase in size, and then as size continues to increase, assumes a slower growth asymptotically approaching an efficiency of one.

We converted the scale for size (as indicated by 'aggregated teaching hours') by dividing individual institutions' teaching hours by the teaching hours of the largest institution, thus making the scales for efficiency and size comparable. A reciprocal function asymptotically approaching +1 can be described in this form:

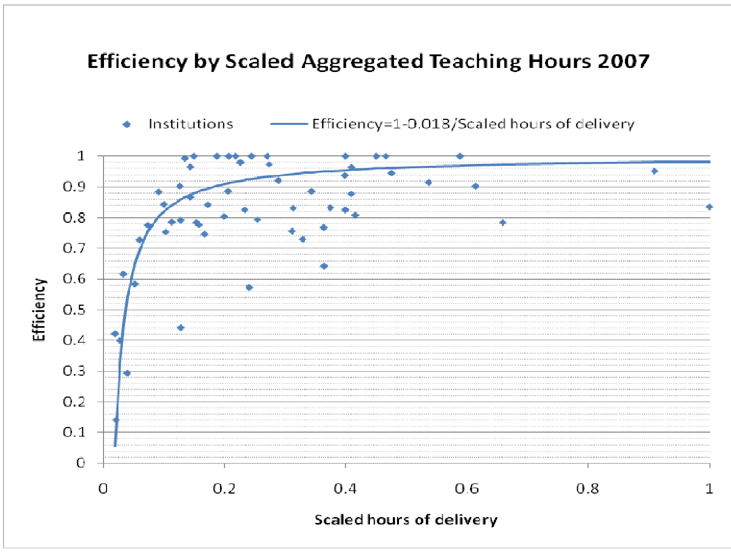
$$f(\text{size}) = 1 - a/\text{size}$$

where a is a constant which defines the tightness of the curve.

We applied an iterative process to determine the numerator a, aiming to maximise the predictive power of the function f. This yielded an a of roughly 0.018, which equated to an explained variance of 58%. The resulting graph is displayed in Figure 3.

One way of defining the point at which institution size becomes instrumental to efficiency would be where the tangent of the function described above is 45 degrees, e.g. at a gradient of one.

Figure 3



This point can be calculated by solving the derivative for size when set to one, thus

$$f'(\text{size}) = a/\text{size}^2 = 1,$$

and therefore size=$\sqrt{a}$, so that in our present case, with 0.018 substituted for a, and then multiplied by the scaling factor the minimum size of an institution can be established as 2.7 million teaching hours, from which point on (normalised) size should not be an impediment to efficiency, relative to the most efficient institutions. It may be advantageous for policy makers to consider this minimum size when contemplating new institutions, amalgamations or the separating of existing TAFEs.

Conclusions

This paper has used Data Envelopment Analysis to examine the technical efficiency of TAFE Institutes in Australia for 2007. This is a technique that has been used before in the context of teaching institutes (TAFE and university). The paper also examined possible predictors of efficiency.

We found that while institutes have relatively high efficiency scores overall there is some variation. Variation arises from two main sources - the size of the institute and another factor we call a remoteness indicator. This includes the proportion of students from remote areas in an institute but this factor is also related to the Indigenous status of students as well as institute size.

There is an underlying assumption about quality which we have not covered in this paper. We purely looked at technical efficiency and some of the factors that may predict this. In our analysis of technical efficiency we also need to be aware that TAFEs' have other obligations (such as community service obligations) that can effect technical efficiency.

Nevertheless, the approach used in this paper can provide useful information to institutes regarding efficiency. The information may also be useful in terms of guiding funding models. For instance, the cost of delivery is higher where there are large proportions of students come from remote areas.

Acknowledgements

NCVER is grateful for the sponsorship and interest shown in this project by the Department of Further Education, Employment, Science and Technology in South Australia.

References

Abbott M & Doucouliagos C 1998 *A DEA analysis of the efficiency of Victorian Technical and Further Education institutes*, Working paper series no. 9802, Faculty Of Business Law, Deakin University, Ballarat.

Abbott M & Doucouliagos C 1999 *Technical and scale efficiency of vocational education and training institutions: The case of the New Zealand Polytechnics*, Working paper no. 22, Centre for the Economics of Education and Training, Monash University, Melbourne.

Abbott M & Doucouliagos C 2000 Amalgamations and the efficiency of Queensland TAFE institutes, *Australian and New Zealand journal of vocational education research*, v. 8, no. 2, pp. 1-19.

Abbot M & Doucouliagos C 2003 The efficiency of Australian universities: a data envelopment analysis, *Economics of education review*, vol 22, no. 1, pp89-98.

Abbot M & Doucouliagos C 2009 Competition and efficiency: overseas students and technical efficiency in Australian and New Zealand universities, *Education Economics*, vol. 17, no. 1, pp. 31-58.

Agasisti T & Johnes G 2009 Beyond Frontiers: Comparing the Efficiency of Higher Education Decision-Making Units Across More than One Country, *Education Economics*, Vol. 17, n. 1, pp. 59-79.

Carrington R, Coelli T & Rao P 2004 *Measuring the performance of Australian universities: conceptual issues and initial results*, The Asia-Pacific Productivity Conference 2004, University of Queensland, Brisbane, 14-16 July.

Charnes A, Cooper W.W., Rhodes, E. 1978 Measuring the efficiency of decision making units, *European Journal of Operational Research*, vol. 2 (1978) pp. 429- 444.

Coelli TJ, Prasada Rao DS, O'Donnell CJ & Batteses GE 2005 *An introduction to efficiency and productivity analysis*, 2nd edition, Springer, New York.

Coelli, TJ 1996 A Guide to DEAP Version 2.1: *A Data Envelopment Analysis (Computer) Program;* Centre for Efficiency and Productivity Analysis (CEPA) Working Papers, No 8/96.

Cooper W, Seiford L & Zhu J 2004 Data envelopment analysis: Models and interpretations, in W Cooper, L Seiford & J Zhu (Eds), *Handbook on data envelopment analysis*, Kluwer Academic Publisher, Boston.

Krause, K., Hartley, R., James, R., & McInnis, C. 2005 *The first year experience in Australian Universities: Findings from a decade of National studies.* Melbourne: Centre for the Study of Higher Education, University of Melbourne.

Mills, C, Heyworth J., Rosenwax L., Carr S. & Rosenberg M 2009 Factors associated with the academic success of first year Health Science students, Advances in Health Sciences Education, vol 14 pp. 205-217