Efficiency of Australian TAFE and further education providers

Fieger, Peter and Villano, Renato and Cooksey, Ray

University of New England, Armidale, NSW, Australia

5 March 2015

Efficiency of Australian Technical and Further Education Providers

Peter Fieger[1], Renato Villano[1], Ray Cooksey[1]

Abstract

Budgetary constraints on the public purse have led Australian Federal and State governments to focus increasingly on the efficiency of public institutions, including Technical and Further Education (TAFE) institutes. In this study, we define efficiency as the relationship between financial and administrative inputs and educational outputs. We employ stochastic frontier analysis in determining the efficiency of Australian TAFE institutes using data sourced from institutional annual reports, the Student Outcomes Survey and administrative databases. We found significant economies of scale effects and conclude that increasing institutional size for very small institutions may result in increased efficiencies.

Introduction

Increased competition for scarce public funding has highlighted the need for governments to encourage the TAFE sector to demonstrate improved productivity. Efficiency of TAFE institutes is thus of great interest to policy makers, regulators, consumers, and to the institutions themselves. Knowledge about institutional efficiency may be useful to government agencies in allocating funds and in assessing the impact of funding decisions on the relationship between financial and administrative input into institutions versus the produced output, for instance in the form of hours taught, graduates produced, employment outcomes and other outcomes. Institutions may use information about their own efficiency to benchmark themselves against other institutions and to make adjustments to their own resource allocation. Regulators can also use this knowledge to potentially identify areas of high risk in the delivery of vocational education and training (VET). Moreover, due to limited market mechanisms in the provision of educational products, knowledge of alternative means

[1] University of New England, NSW, Australia
Contact: pfieger@myune.edu.au

to establish benchmarks of efficiency is of importance to all stakeholders in educational institutions.

The contemporary approach to the analysis of the efficiency in the form of a production function was pioneered by Farrell. In his seminal paper (Farrell, 1957), he argued that the measurement of efficiency is necessary to ascertain whether additional inputs are needed to increase desired outputs or if such outputs can be increased by raising efficiency alone. He also developed a generalisable production function which enabled the computation of efficiency measurements under multiple input scenarios. In the 1970s two groups of researchers arrived at two different techniques for the specification of production frontiers: Aigner, Lovell, and Schmitt (1977) formulated the first stochastic frontier model, a parametric maximum likelihood technique which overcame the previous limitations of frontier estimation by introducing a new approach to the specification of the error term, namely its separation into a normal 'noise' term and a one sided inefficiency term. Almost at the same time, Charnes, Cooper, and Rhodes (1978) published their work on a non-parametric linear programming method, Data Envelopment Analysis (DEA). This method focuses on the scalar measure of the efficiency of each unit under consideration which is obtained after the determination of weights for the observed data for inputs and outputs.

The introduction of both of production frontier methods has led a growing body of empirical research. One of their features is their utility in multiple input and output scenarios, which makes this form of efficiency analysis particularly useful for non-commercial units (often called Decision Making Units (DMU)). While production frontier methods have been used in the analysis of commercial contexts, one of the main applications has been the efficiency analysis of public institutions and government owned entities. The spectrum of sectors analysed has varied across a wide field of institutional units, ranging from hospitals, public transport, public utilities, and prisons, to numerous applications of educational contexts.

In this study, we employ parametric Stochastic Frontier Analysis (SFA) to determine the efficiency of Australian TAFE institutes. We will proceed in the following manner: First, we review the theoretical underpinnings of the technique used and identify and describe the appropriate variables and data that are going to be used in the analysis. Then, we operationalise the model and discuss the resulting estimates and efficiencies. Finally, we consider the practical relevance of our research results and whether concrete policy implications could emerge from our findings.

Production frontiers and their application in education

Efficiency analysis utilising SFA or DEA has been applied frequently in educational contexts. However, despite the popularity of econometric frontier analysis overseas, the existing published research utilising SFA or DEA in Australian education is somewhat limited. Most of the existing published research has focussed on universities. Avkiran (2001) applied DEA and used 1995 data of Australian universities to determine universities' productivity in respect to the delivery of educational services and fee paying enrolments. Other DEA studies examining cross-sectional university performance were performed by Abbott and Doucouliagos (2003), Carrington, Coelli, & Rao (2005), and Worthington and Lee (2008). Horne and Hu (2008) and Abbott and Doucouliagos (2009) published SFA research of Australian and New Zealand and Australian universities. Finally, only a small number of studies involving Australian TAFEs could be identified. These were notably the research by Abbott and Doucouliagos (2002) who performed DEA analyses utilising data from Victorian institutes only and one nationwide DEA study by Fieger (2010). There has been no previous published efficiency analysis of the Australian TAFE sector which utilised the stochastic frontier approach.

Method of Analysis

We employ the stochastic frontier framework. The foundations for this methodology were laid by Aigner, Lovell, and Schmidt (1977) who formulated the first stochastic frontier model. Their main contribution was the introduction of a new approach to the specification of the error term, namely its separation into a normal 'noise' term and a one sided inefficiency term. Stochastic frontier production functions are an extension to the classic Cobb-Douglas (1928) function which can generally be expressed in this form:

$$Y = e^{\beta_0} X_1^{\beta_1} X_2^{\beta_2} \ldots X_n^{\beta_n} e^{\varepsilon} \qquad (1)$$

This model can then be transformed by taking the log of both sides:

$$ln(Y) = \beta_0 + \beta_1 \ln(X_1) + \beta_2 \ln(X_2) \ldots \beta_n \ln(X_n) + \varepsilon = \beta_0 + \sum_{i=1}^{n} \beta_i \ln(X_i) + \varepsilon \qquad (2)$$

This model is easily recognisable as a variation of the classical multiple regression model in which Y stands for the output, β_o for the intercept, β_i for a vector of inputs and ε for statistical noise. Aigner et al.'s contribution was to postulate that in SFA the error term ε essentially corresponds to two error components, one being the statistical noise portion v, and the other being the non-negative technical efficiency u which is distributed independently from v.

$$\varepsilon_i = v_i - u_i \qquad (3)$$

The original Cobb-Douglas function can thus be re-formulated as

$$ln(Y) = \beta_0 + \sum_{i=1}^{n} \beta_i \ln(X_i) + v_i - u_i \qquad (4)$$

where the technical efficiency TE_i of u_i can then be determined by

$$TE_i = e^{-u_i} \qquad (5)$$

TE_i is meant to be located between 0 and 1 and is ordinarily assumed to be positively half-normally distributed[2]. Aigner et al. determined the mean of ε and u as:

$$\mu_\varepsilon = \mu_u = -\sigma_u \sqrt[2]{\frac{2}{\pi}} \qquad (6)$$

and the variance of error ε as:

$$var(\varepsilon) = var(u) + var(u) = \frac{\pi-2}{\pi} \sigma_u^2 + \sigma_v^2. \qquad (7)$$

where σ_u represents the variance of the normal distribution prior to truncation to 0. The parameterisation above allows for the specification of additional relationships which enable the interpretation of results. The total variance in the error term is given by σ_ε^2.

$$\sigma_\varepsilon^2 = \sigma_v^2 + \sigma_u^2 \qquad (8)$$

[2] In this study we will apply the half normal error distribution assumption. There are, however, other error distribution assumptions possible, such as exponential, half-truncated, gamma etc.

The ratio of the standard deviation of the inefficiency component to the standard deviation of the 'noise' error component is given by λ, and γ is an indicator of the portion of the one sided error component in the overall variance:

$$\lambda = \frac{\sigma_u}{\sigma_v} \qquad (9) \qquad \text{and} \qquad \gamma = \frac{\sigma_u^2}{\sigma_\varepsilon^2} \qquad (10)$$

These simple relationships represent a convenient means to assess the quality of the results of a SFA. For instance, $\lambda \rightarrow 0$ implies that $\sigma_v^2 \rightarrow \infty$ and/or $\sigma_u^2 \rightarrow 0$ which indicates that the symmetric error dominates the overall error component. Similarly, when $\lambda \rightarrow \infty$ then $\sigma_u \rightarrow \infty$ or $\sigma_v \rightarrow 0$ and therefore deviation from the frontier can be explained by inefficiency. Following from this is that when $\gamma \rightarrow 1$ the amount of the explained inefficiency increases over the portion of random noise, that is, the value of γ is the approximate proportion that is attributed to inefficiency.

Data characteristics and preparation

The aim of this study is to ascertain the efficiency of Australian TAFE institutes via SFA and to determine which exogenous variables drive the calculated efficiencies. When deciding on an approach to undertake efficiency frontier analysis of TAFE institutes one has to take into account some specific circumstances that are unique to the VET sector. Similar efficiency frontier analyses involving universities or secondary schools can often rely on data such as the number of full time staff, staff qualifications, number of graduates, test scores, grades, research outputs such as publications and conference presentations, successful grant applications, and others. Data comparable to the aforementioned are difficult to obtain for TAFEs. There is obviously a scarcity of research and research related inputs and outputs that relate to TAFEs. Many TAFEs employ a large percentage of part time lecturers, and this proportion differs from institution to institution and reliable data about this proportion is difficult to obtain. Furthermore, TAFEs do not consistently award grades in the same way for some or all of their courses through 'competency based' assessments.

It is therefore clear that there are some circumstances that encumber the specification of frontier efficiency models for TAFE providers. The majority of those circumstances can be categorized into three groups: a) the absence of functional data for the entire sector (e.g. staff qualification data was not reported in a standardised way by institutions), b) partial data only available for a subset of TAFEs (e.g. certain financial data), and c) data that is too dissimilar in nature due to the lack of a comprehensive national reporting standard (e.g. assessment beyond competency based assessment).

Despite the aforementioned difficulties we have been able to assemble and derive a dataset containing adequate information to undertake the course of research set out in earlier paragraphs. The data used in this study came from several sources. These sources included institutional annual reports, information on institutional websites, personal requests to institutional administrators and state regulators, the Student Outcome Survey (SOS), and the Students and Courses database at NCVER. Of significance was the choice of year(s) for which data should be obtained. It was intended to assemble a panel of data comprising a number of years in an effort to a) maximize the number of data points and b) enable analysis of changes in efficiency over a given period. However, data collection was more difficult than anticipated as institutes do not publish financial data in a uniform pattern. Specifically the collecting of several consecutive years of financial data appeared to be difficult. It was thus decided to focus on one particular year with the following stipulation: a) the year had to be as recent as possible, b) it had to be an augmented SOS year[3] to enable the use of the most robust institutional data, and c) the chosen year had to have the maximum of available data points. Taking these considerations into account 2011 was chosen as the year of analysis.

The initial plan was to include all 69 Australian TAFE and TAFE like institutions[4] in this analysis. However, this intention was impeded by a number of factors. In addition to those institutes that did not provide data, some institutions proved to be too specialised to be compared on an equal footing with the majority of TAFE institutes. These were notably the Driver Education Centre of Australia and the National Art School. Some of the TAFE units of universities did not have delineated financial data for their TAFE division available. After considering availability of data for the remaining institutes it was decided to include those units in the final data set that had data for the total expenditure variable in 2011 available. This yielded 56 TAFEs for inclusion in the analysis.

In addition to financial expenditure data the 'teaching hours' variable used in the efficiency analysis was sourced from NCVER's Students and Courses database. This variable

[3] Odd years feature an augmented sample of the SOS, containing about 300,000 questionnaires, of which about one third receives a response. In these years the SOS is designed to enable estimates at an institutional level. In even years the SOS sample contains about 100,000 questionnaires, and the focus of estimates is the state level.
[4] In the context of this study, the term 'TAFE and TAFE like institute' refers to TAFE institutes, TAFE divisions of a university, Skills Institutes and Polytechnics. From here on only referred to as 'TAFE'.

indicates the number of student contact hours by institution. A number of further items were sourced predominantly from the 2011 SOS. These included institutional proportions in terms of sex, student type (module completers/graduates), indigenous students, students who used a language other than English at home, and disabled students. Other variables included were the average age of the student body at individual institutions, and an average institutional remoteness score derived from the ABS's ARIA variable. We also used the SOS to determine the number of different courses offered by each institution which had at least one student enrolled. A categorical variable indicating size was derived from the total expenditure variable. The categories created were 'very large', signifying total expenditure in excess of $120,000,000, large ($70,000,000 to $120,000,000), medium ($45,000,000 to $ 69,999,999), small ($25,000,000 to $44,999,999), and very small with total expenditure of less than $25,000,000.

Empirical model

In this study we aimed to evaluate the technical production efficiency of a number of TAFE institutes. Our interest was in determining institutional efficiency based on basic financial expenditure and administrative input and the produced output as measured by teaching contact hours. The starting point to operationalise our efficiency model was in the form of a production function as expressed by a Cobb-Douglas equation:

$$T = e^{\beta_0} E^{\beta_1} C^{\beta_2} e^{\varepsilon} \qquad (11)$$

where *T* denotes the output in teaching hours, *E* the total expenditure, and *C* the number of courses offered by a given TAFE. C was included as it is an indicator of the complexity of college administration. Taking the natural logarithm of (11) and accounting for the SFA specific error component as shown by Battese and Coelli (1995) resolves to:

$$\ln(T_i) = \beta_0 + \beta_1 \ln(E_i) + \beta_2 \ln(C_i) + v_i - u_i \qquad (12)$$

Descriptive statistics for variables used in estimating this model can be found in Table 1.

Table 1 Descriptive statistics SFA model

Variable	N	Mean	StdDev	Minimum	Maximum
Teaching Hours	56	5,521,177.5	4,174,682.5	473,279	22,346,943
Total Expenditure	56	79,966,968.0	53,563,163.2	12,324,312	288,974,000
Number of courses offered	56	172.6	83.3	32	439

In addition to the frontier production function (12) we intended to investigate which exogenous variables may be influencing technical efficiency. We therefore specified a second component in which we included some variables which were hypothesised to influence efficiency:

$$\mu = \delta_0 + \sum_{k=1}^{K} \delta_k z_k \qquad (13)$$

Here, z represents the hypothesised K predictors of efficiency and δ the parameters that needed to be estimated. In our model we hypothesized that predominantly demographic factors influence efficiency, as these factors may require administrative adjustments to TAFE operations. We therefore entered the variables with institutional indicators for English as a second language, disability, remoteness, age and sex, into our efficiency model (for descriptive statistics see Table 2).

Table 2 Descriptive statistics inefficiency model

Variable	N	Mean	StdDev	Minimum	Maximum
English second language	56	16.3	9.8	4.6	40.2
Students with disability	56	9.4	2.9	4.4	18.5
Remoteness (ARIA)	56	2.1	1.0	1.1	4.7
Student age	56	33.0	2.2	27.6	37.1
Proportion of males	56	57.2	10.7	32.8	96.6

This two component scenario would have originally been estimated in a two step approach, where the first step specifies the stochastic production frontier and leads to the estimation of efficiency scores and the second step is to estimate the relationship between efficiency scores and efficiency predictors. Wang and Schmidt (2002) have demonstrated that this two step procedure is biased and that instead stochastic frontier models and the way in which efficiency u_I depends on predictors can and should be estimated in one single step using maximum likelihood estimation.

Analysis by Waldman (1982) has shown that for the specification of a stochastic frontier model it is beneficial to examine the third moments of the least squares residual. If this quantity is positive, then the least squares slope estimates and λ=0 represent a local maximum of the likelihood. Conversely, if the third moment is negative, the likelihood has a greater value at some other point where λ=0. This means that negative skewness of the residuals of the OLS regression indicates that maximum likelihood estimation is indeed the appropriate procedure to estimate the production frontier. We thus began our analysis with the formulation of a linear regression model identical to our proposed SFA model. The results can be seen in table 3 (model 1). The third moment based of the OLS residuals was estimated to be -0.63, thus indicating to be a satisfactory prerequisite for the maximum likelihood estimation of the stochastic frontier. While the estimates of the OLS model only have limited usefulness, they provide a meaningful starting point for the maximum likelihood estimation (Cullinane & Song 2006). The R-squared estimate of the OLS is with 0.91 fairly substantial and indicates that most of the variation in teaching hours can be explained by total expenditure and number of courses offered by institute. The two independent variables themself are highly significant and both exhibit the sign that would be expected, e.g. higher expenditure and increasing number of courses tend to be associated with a rise in teaching hours.

We could then estimate our basic stochastic frontier model, using the same variables (Table 3, model 2). While coefficients and intercept have the same sign as in OLS regression, along with similar magnitude and strong significance, the real interest here is in the estimated variance parameters. The strong significance of the Wald test indicates that the coefficient(s) are significantly different from zero and thus confirms the model's explanatory power. σ_u and σ_v are both significant. This suggests the statistical significance of the random error and inefficiency component of the model. The significance of λ confirms the presence of inherent statistical inefficiency in the data. The estimate for γ at 0.9 is quite high and denotes that 90% of the variability in delivered teaching hours could be attributed to technical inefficiencies. The closeness of γ to 1 points towards the existence of a deterministic production frontier (Parsons, 2004). The significance of γ and λ affirm the preponderance of inefficiency in the composite error term and also validate SFA as the appropriate tool for this specific analysis (Chen, 2007). Additionally a test was performed to determine wether the units investigated by our Cobb Douglas model use constant returns to scale technology.

Table 3 Estimates for OLS and SFA models

Variables	OLS		MLE			
	Model1		Model2		Model 3	
	Est	P>|t|	Est	P>|z|	Est	P>|z|
Stochastic FrontierModel						
Constant	-4.221	<.001	-4.022	<.001	-2.730	<.001
Total Expenditure	0.926	<.001	0.989	<.001	0.968	<.001
Number of courses offered	0.553	<.001	0.345	<.001	0.134	0.025
Inefficiency Model						
Constant					-17.631	0.001
English second language					0.129	0.027
Students with disability					0.053	0.726
Remoteness (ARIA)					2.708	<.001
Student age					-0.074	0.768
Proportion of males					0.112	0.048
R-squared	0.913					
Wald Chi-sq			385.4	<.001	983.5	<.001
Sigma v			0.126	<.001	0.127	<.001
Sigma u			0.387	<.001		
Lambda			3.073	<.001		
Gamma			0.904			

The test of this hypothesis determines whether the sum of the coefficients in the model is statistically different from 1. The sum of the coefficients for 'total expenditure' and 'number of courses' was calculated as 1.33 and the test for equality to 1 yielded a chi squared value of 6.54 (p=0.0106), so that we could reject the hypothesis of constant returns to scale technology and assume an increasing returns to scale setting. In the scenario considered, this means that outputs will increase disproportionally when inputs are increased.

Having gained insights into the characteristics of our basic frontier model we could proceed to specify the SFA model that included explanatory variables for the technical inefficiency variance function (Table 3, model 3). First we note that parameters and significance of the frontier function are comparable to the model without the inefficiency terms. The Wald chi-squared value and the variance component of the random error term of the whole model were also significant and of similar magnitude. The main items of interest in model three are thus the inefficiency effects. We note that the proportion of students with a disability and the institutional mean age of the student body are not related to institutional efficiency. The strong significance of remoteness points to inefficiency being a function of remoteness. This result confirms the findings of Fieger (2010), who found remoteness being the key variable associated with inefficiency. This finding may be partially attributed to

Australia's unique geography and related issues of infrastructure and demographics, however, it must also be noted that 'remoteness' acts also as a proxy for institution size as many urban institutes tend to be significantly larger than rural institutes. Internationally, remoteness is rarely identified as driver of inefficiency, although Izadi, Johnes, Oskrochi, & Crouchley (2002) found some incidental relationship between remoteness and inefficiency. In model 3 we find further, albeit weaker, positive associations between the proportion of males and inefficiency, and the proportion of students with English as a second language and inefficiency. Possible explanations here may be that males tend to be engaged at higher rates in apprenticeships, which require larger administrative and financial efforts on the part of the institution. An assessment of the correlation between the proportion of males and the proportion of apprentices and trainees in 2011 revealed an overall correlation of 0.44 ($p<0.001$), thus supporting this explanation. Greater financial, educational and administrative efforts may also be at play when considering the relationship between increasing inefficiency and higher rates of non-native English speakers. Larger proportions of students with English as a second language may necessitate more intensive teaching modes, such as lower teacher/student ratios, which may in turn explain some variation in institutional inefficiency in respect to the percentage of non-native English speakers.

After verifying the suitability of our model and discussing the interpretation of model statistics and coefficients we were interested in the actual estimated efficiencies of individual institutions. The efficiencies follow from (5) and specifically for the half-normal production model are derived by

$$TE = \left\{\frac{1-\Phi(\sigma_* - \mu_{*i})}{1-\Phi\left(-\frac{\mu_{*i}}{\sigma_*}\right)}\right\}\exp\left(-\mu_{*i} + \frac{1}{2}\sigma_*^2\right) \qquad (14)$$

where Φ signifies the cumulative distribution of the normal distribution and μ_{*i} and $\sigma_*$are defined as

$$\mu_{*i} = -\epsilon_i \sigma_u^2 / \sigma_s^2 \qquad (15) \qquad \text{and} \qquad \sigma_* = \sigma_u \sigma_v / \sigma_s \qquad (16)$$

The calculated efficiencies for model 3 can be found in table 4.

Table 4 Observed institutional efficiencies

Technical efficiency							
Institute	**Efficiency**	**Institute**	**Efficiency**	**Institute**	**Efficiency**	**Institute**	**Efficiency**
1	0.984	**22**	0.862	**36**	0.986	**53**	0.840
4	0.977	**23**	0.921	**37**	0.979	**55**	0.967
5	0.973	**24**	0.964	**38**	0.991	**56**	0.474
7	0.932	**25**	0.968	**40**	0.621	**57**	0.723
10	0.953	**26**	0.908	**43**	0.960	**58**	0.327
11	0.943	**27**	0.985	**44**	0.946	**60**	0.389
13	0.971	**28**	0.973	**45**	0.893	**64**	0.948
14	0.953	**29**	0.959	**46**	0.980	**65**	0.977
15	0.978	**30**	0.987	**47**	0.916	**66**	0.979
16	0.966	**31**	0.967	**48**	0.927	**70**	0.918
17	0.953	**32**	0.866	**49**	0.992	**71**	0.978
18	0.986	**33**	0.982	**50**	0.972	**74**	0.198
19	0.960	**34**	0.996	**51**	0.981	**77**	0.983
20	0.963	**35**	0.920	**52**	0.739	**110**	0.423
Mean: 0.888							
SD: 0.182							

Economies of scale effects can always be suspected where comparable units produce variable quantities of similar goods. The obvious reason for this in the setting under consideration is that increasing 'hours taught' costs decrease on a per hour basis as operational fixed costs can be shared over more hours. In the higher education sector such economies of scale have been well documented (see, for instance, Hashimoto and Cohn, 1997), albeit mostly in the university context. In the Australian TAFE sector, one could reasonably expect that larger institutes exhibit higher efficiency. We were therefore interested in patterns of efficiency in respect to institute size. Figure 1 displays the institutional efficiency in respect to institute size, as measured by teaching hours.

In this graph blue dots identify individual institutes and their location indicates the relationship between efficiency and size. As was hypothesised, smaller institutes appear to exhibit significantly lower efficiency than larger institutes. This graph should be of interest to regulators and policy makers, as it shows a striking change in efficiency over only a small portion of size increase on the far left of the chart. We fitted a curve over the data in order to be able to mathematically define the point at which further increases in size cease to translate into significant gains in efficiency. Practically, this point should define the minimum size for

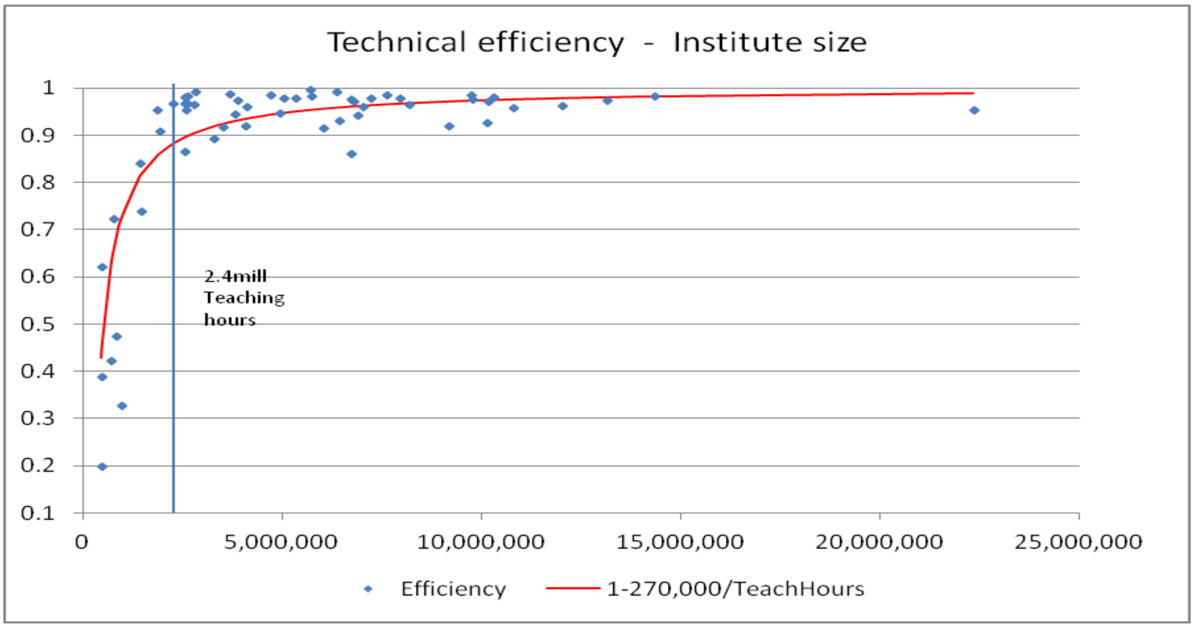

Figure 1 Technical efficiencies as a function of institute size

a TAFE to operate efficiently. The curve fitted defines the relationship efficiency as a function of size as

$$E = 1 - \frac{2.7*10^5}{S} \qquad (17)$$

where S indicates size as measured by teaching hours. The resulting fit explains about 88 percent of the variance in efficiency and is thus a reasonable representation of the data. We then defined the turning point of this function as the point where the strong increase in efficiency in respect to teaching hours eases. The derivative of (17) yields

$$\frac{dE}{ds} = \frac{2.7*10^5}{s^2} \qquad (18)$$

Solving (18) for a slope of 1 and accounting for the different scale of y and x axis yields

$$T = \sqrt{2.7 * 10^5 * TH_{max}} = 2.4 * 10^6 \qquad (19)$$

where TH_{max} represents the teaching hours of the largest institution. It can thus be stated that, based on the above derivation, when institutional size is equal or greater to about 2.4 million teaching hours, size is no longer an impediment to efficiency. Alternatively it can be concluded that, in order to be efficient in the transformation from financial resources to units taught, TAFE institutes should be of a size that corresponds at least 2.4 million teaching hours. Interestingly, this finding is similar to the results presented in Fieger (2010), where a different methodology, different data and variables, and a different base year were employed. This may add impetus to the validity of our findings presented here.

Conclusion

In this study we have applied a stochastic frontier model to estimate the efficiency of Australian TAFE institutes, focussing on the relationship between financial and administrative inputs and teaching output. We have observed some clear inefficiencies. These are mainly related to the degree of remoteness and student characteristics. The least efficient TAFE institutes are more likely to be found in remote locations, have a higher percentage of males, and a larger proportion of individuals from non English speaking backgrounds. We speculate these inefficiencies are driven by a combination of interrelated factors, including geographic location, available infrastructure and the absence of occupational diversity of graduates. Significant economies of scale effects were observed. These effects disappear once institutions exceed a certain minimum threshold in size. We conclude that increasing institutional size for very small institutions (that is those that produce less than 2.4million teaching hours) may result in increased efficiencies.

References

Abbott, M., & Doucouliagos, C. (2002). A data envelopment analysis of the efficiency of Victorian TAFE institutes. *Australian Economic Review*, *35*(1), 55-69.

Abbott, M., & Doucouliagos, C. (2003). The efficiency of Australian universities: a data envelopment analysis. *Economics of Education Review*, *22*(1), 89-97.

Abbott, M., & Doucouliagos, C. (2009). Competition and efficiency: overseas students and technical efficiency in Australian and New Zealand universities. *Education Economics*, *17*(1), 31-57.

Aigner, D., Lovell, C. A., & Schmidt, P. (1977). Formulation and estimation of stochastic frontier production function models. *Journal of Econometrics*, *6*(1), 21-37.

Avkiran, N. K. (2001). Investigating technical and scale efficiencies of Australian universities through data envelopment analysis. *Socio-Economic Planning Sciences*, *35*(1), 57-80.

Battese, G. E., & Coelli, T. J. (1995). A model for technical inefficiency effects in a stochastic frontier production function for panel data. *Empirical Economics*, *20*(2), 325-332.

Carrington, R., Coelli T. & Rao, P. (2005), 'The performance of Australian Universities: Conceptual issues, and preliminary results', *Economic Papers*, 24 (2),145 – 163.

Charnes, A., Cooper, W. W., & Rhodes, E. (1978). Measuring the efficiency of decision making units. *European Journal of Operational Research*, *2*(6), 429-444.

Chen, C. F. (2007). Applying the stochastic frontier approach to measure hotel managerial efficiency in Taiwan. *Tourism Management*, *28*(3), 696-702.

Cobb, C. W., & Douglas, P. H. (1928). A theory of production. *The American Economic Review*, *18*(1), 139-165.

Cullinane, K., & Song, D. W. (2006). Estimating the relative efficiency of European container ports: a stochastic frontier analysis. *Research in Transportation Economics*, *16*, 85-115.

Farrell, M. J. (1957). The measurement of productive efficiency. *Journal of the Royal Statistical Society. Series A (General)*, *120*(3), 253-290.

Fieger, P. (2010). An investigation of TAFE efficiency. 2010 AVETRA Conference; published conference presentation, Surfers Paradise.

Hashimoto, K., & Cohn, E. (1997). Economies of scale and scope in Japanese private universities. *Education Economics*, *5*(2), 107-115.

Horne, J., & Hu, B. (2008). Estimation of cost efficiency of Australian universities. *Mathematics and Computers in Simulation*, *78*(2), 266-275.

Izadi, H., Johnes, G., Oskrochi, R., & Crouchley, R. (2002). Stochastic frontier estimation of a CES cost function: the case of higher education in Britain. *Economics of Education Review*, *21*(1), 63-71.

Waldman, D. M. (1982). A stationary point for the stochastic frontier likelihood. *Journal of Econometrics*, *18*(2), 275-279.

Wang, H. J., & Schmidt, P. (2002). One-step and two-step estimation of the effects of exogenous variables on technical efficiency levels. *Journal of Productivity Analysis*, *18*(2), 129-144.

Worthington, A. C., & Lee, B. L. (2008). Efficiency, technology and productivity change in Australian universities, 1998–2003. *Economics of Education Review*, *27*(3), 285-298.