International Journal of Training Research, 2016
Vol. 14, No. 1, 62–75, http://dx.doi.org/10.1080/14480220.2016.1152030

Efficiency of Australian technical and further education providers

Peter Fieger*, Renato Villano and Ray Cooksey

Business School, University of New England, Armidale, 2351, Australia

(*Received 10 August 2015; accepted 1 December 2015*)

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.

Keywords: Vocational education; productivity; Australian education; production frontiers; performance measurement

Introduction

Increased competition for scarce public funding has highlighted the need for governments to encourage the Technical and Further Education (TAFE) sector to demonstrate improved productivity. Efficiency of TAFE institutes is thus of great interest to policymakers, regulators, consumers and to the institutions themselves. Knowledge about institutional efficiency may be useful to government agencies in allocating funds. Furthermore, a better understanding of the drivers of efficiency can only be based on a better assessment of the relationships between financial and administrative inputs into institutions versus the produced outputs. Such outputs may include hours taught, graduates produced, employment outcomes and the like.

Institutions may use information about their own efficiency to benchmark themselves against other institutions and to make adjustments to their own internal 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 quasi-market mechanisms used in the provision of educational products, knowledge of alternative means to establish benchmarks of efficiency is of importance to all stakeholders in educational institutions. Indicators of institutional efficiency thus represent a potential performance measure (Karmel, Fieger, Blomberg, & Loveder, 2013).

Achieving better measures of efficiency is vital in the Australian context. All levels of education in Australia maintain a balance of government and private delivery, funding and benefit. Increasingly, governments seek to leverage their own public investments through co-investment by individuals, and by improving the efficacy and efficiency of

*Corresponding author. Email: peter.fieger@ncver.edu.au

educational delivery agents. Of the various educational groups, VET has been the most exposed to marketisation and subsequent analysis – with arguably mixed results.

The contemporary approach to the analysis of the efficiency in the form of a production function embodying multiple inputs 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 Schmidt (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 types of production frontier methods has facilitated 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' productivities 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, and 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. The only nationwide study examining the TAFE sector was published by Fieger, Karmel, and Stanwick (2010). This study applied DEA and also added a Tobit model in order to uncover predictors of TAFE efficiency. The primary finding was that remoteness is highly related to inefficiency.

The present study is motivated by this earlier research. Our predominant aims are to employ a slightly modified definition of efficiency by estimating the efficiency of converting institutional resources into teaching loads. To do this, we will employ stochastic frontier analysis. We are thus also aiming to demonstrate the suitability of this parametric method in the context of TAFE efficiency analysis. Finally, we are seeking to determine which additional predictors may be associated with efficiency.

Method of analysis

The present study seeks to build on the research presented in Fieger et al. (2010). We employ the stochastic frontier framework. The foundations for this methodology were laid by Aigner et al. (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.

SFA differs from DEA in that it is a parametric method, e.g. there are specific distributional assumptions in respect to the data that are used in the model. Each of the two methods has distinct advantages and disadvantages. The main advantages of DEA are that this method does not depend on the explicit specification of the form of the production function, and that it can easily deal with multiple inputs and outputs. Main disadvantages of DEA emanate from its non-parametric nature and manifest themselves in the difficulty of conducting statistical hypothesis testing and its sensitivity to outliers defining the production function. On the other hand, SFA is less sensitive to outliers and allows for hypothesis testing, but in its basic form is limited to a single output and requires careful specification of the functional form.

SFA and DEA represent methods to measure essentially the same outcome, e.g. institutional efficiency. Previous research has shown that the empirical results determined by either method do not often substantially differ (see, for instance, Cullinane, Wang, Song, & Ji, 2006 or Hossain, Kamil, Baten, & Mustafa, 2012). Other empirical research, however, has uncovered substantial discrepancies of estimated efficiencies between the two methods. For instance, Fiorentino, Karmann, and Koetter (2006) performed an analysis of the productivity of German banks and compared the efficiency scores produced by both methods. They found noticeably different results, and that the discrepancies in their results could be attributed to the inclusion of outliers and comparing institutes from different fragments of the market. Jacobs (2001) also found inconsistencies between the results of SFA and DEA and ascribed differences to the existence of 'random noise' and outliers.

The present TAFE environment is characterised by rapid changes, and also the coexistence of institutes with vastly different characters (small vs. large; rural vs. urban, etc.). In light of findings such as those by Fiorentino et al. (2006) and Jacobs (2001), it appears useful to apply SFA as an alternative method to DEA in estimating institutional efficiencies. While it is not intended that the present study provides a direct comparison of the two methods in respect to the earlier 2010 research by Fieger, Karmel, and Stanwick, (the present study uses more recent data and also different variables), it will be instructive to determine whether a similar pattern of institutional efficiencies emerges.

Stochastic frontier production functions are an extension to the classic Cobb and Douglas (1928) function which can generally be expressed in this form:

$$Y = e^{\beta_0} X_1^{\beta_1} X_2^{\beta_2} \ldots X_n^{\beta_n} e^{\varepsilon}. \tag{1}$$

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

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

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 (1977) 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.

$$\varepsilon_i = v_i - u_i. \tag{3}$$

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

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

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

$$TE_i = e^{-u_i}. \tag{5}$$

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

$$\mu_\varepsilon = \mu_u = -\sigma_u^2 \sqrt{\frac{2}{\pi}} \tag{6}$$

and the variance of error ε as:

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

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.

$$\sigma_\varepsilon^2 = \sigma_v^2 + \sigma_u^2. \tag{8}$$

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:

$$\lambda = \frac{\sigma_u}{\sigma_v} \tag{9}$$

and

$$\gamma = \frac{\sigma_u^2}{\sigma_\varepsilon^2}. \tag{10}$$

These simple relationships represent a convenient means to assess the quality of the results of an SFA. For instance, $\lambda \to 0$ implies that $\sigma_v^2 \to \infty$ and/or $\sigma_u^2 \to 0$ which indicates that the symmetric error dominates the overall error component. Similarly, when $\lambda \to \infty$ then $\sigma_u \to \infty$ or $\sigma_v \to 0$ and therefore deviation from the frontier can be explained by inefficiency. Following from this is that when $\gamma \to 1$, 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 converting financial and administrative resources into teaching outputs in Australian TAFE institutes via SFA and to determine which exogenous variables drive the calculated efficiencies. The potential input variables for such an analysis of the TAFE sector differ substantially from variables that one would consider in a similar analysis of the higher education sector. While in the university environment there are inputs and outputs such as number of full-time staff, staff qualifications, number of graduates, test scores, grades, research outputs such as publications and conference presentations and successful grant applications, similar data are difficult to obtain for TAFEs. There are no uniformly reported research-related inputs and outputs that relate to TAFEs. TAFEs ordinarily 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. When aiming to analyse the efficiency of TAFE providers, it is therefore necessary to rely on a reduced range of data that is utilisable to specify production frontier models. There are additional impediments relating to data availability and quality. These can be categorised as the absence of functional data for the entire sector (e.g. staff qualification data was not reported in a standardised way by institutions), data that are available only for a subset of TAFEs (e.g. certain financial data) and data that is too dissimilar in nature due to the lack of a comprehensive national reporting standard (e.g. assessment beyond competency-based assessment).

Irrespective of the complications outlined above, we have been able to derive and assemble a data-set containing adequate information to undertake the course of research set out in earlier paragraphs. The data used in this study were obtained from various sources, including institutional annual reports, information on institutional websites, personal requests to institutional administrators and state regulators, the Student Outcomes Survey (SOS), and the Students and Courses database at the Australian National Centre for Vocational Education Research (NCVER). The base year chosen for this analysis was 2011 as this year represented the optimum trade-off between the aim to conduct the analysis using fairly recent data and maximising the number of institutes for whom the necessary data was available. Furthermore, we aimed to conduct the analysis for an odd year as those years feature an augmented version of the SOS (The SOS is conducted as small sample size version in even years, featuring a sample size of about 100,000 students and an augmented version in odd years, featuring a sample size of approximately 300,000 students). This enabled the use of the most robust institutional data from this source.

In the chosen base year, there were 69 TAFE and TAFE-like institutions (such as Polytechnics, Skills institutes, etc.) in operation. It was, however, not possible to include all institutes in the analysis. 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 53 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 indicates the number of student contact hours by institution. For the inefficiency model we were interested in testing the hypothesis that several additional variables are related to TAFE efficiency. The load pass rate (LPR) was calculated using the Students and Courses database. The derivation of this indicator uses the number of competency achieved (A), recognition of prior learning granted (RPL), competency not achieved (F) and withdrawals (W) via the following formula (ANR, 2011, p. 208):

$$LPR = \frac{A + RPL}{A + RPL + F + W} * 100. \tag{11}$$

A number of items were sourced from the 2011 SOS. These included institutional proportions in terms of sex, students who used a language other than English at home, and institute share of disabled students. Other variables included were the average age of the student body at individual institutions, and a remoteness score derived from the institutional mean of the ABS' ARIA variable. The individual ARIA code for students' place of residence ranges from 1 (major city) to 5 (very remote). We also used the SOS to determine the number of different courses offered by each institution which had at least one student enrolled. Additional variables that where derived from the SOS were the percentage of students enrolled in Certificate III or higher courses and the percentage of apprentices and trainees. Finally, it was of interest whether the course completion rate is related to institutional efficiency. Course completions in vocational education are difficult to ascertain, due to a number of factors (for a more detailed discussion of this issue and a model-based approach to estimate completion rates, see Mark & Karmel, 2010). We applied the methodology described in Mark and Karmel and estimated 2011 completion rates from data drawn from the students and courses database.

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:

$$T = e^{\beta_0} TE^{\beta_1} CT^{\beta_2} e^{\varepsilon} \tag{12}$$

where T denotes the output in teaching hours, TE the institutional total expenditure and CT the number of courses taught by a given TAFE. CT was included as it is an indicator of the complexity of college administration. Taking the natural logarithm of (12) and accounting for the SFA specific error component as shown by Battese and Coelli (1995) resolves to:

$$\ln(T_i) = \beta_0 + \beta_1 \ln(TE_i) + \beta_2 \ln(CT_i) + v_i - u_i. \quad (13)$$

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

In addition to the frontier production function (13) 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:

$$\mu = \delta_0 + \sum_{k=1}^{K} \delta_k z_k. \quad (14)$$

Here, z represents the hypothesised K predictors of efficiency and δ the parameters that needed to be estimated. In our model we hypothesised that predominantly demographic factors influence efficiency, as these factors may require administrative adjustments to TAFE operations. We therefore entered the variables representing institutional indicators for student age, disability, English as a second language, remoteness, percentage of students enrolled in Certificate III or higher, percentage of apprentices and trainees, average load pass rate and course completion rate into our inefficiency model (for descriptive statistics see Table 2).

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 $\lambda = 0$ 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 $\lambda = 0$. This means that negative skewness of the residuals of the ordinary least squares (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, and 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

Table 1. Descriptive statistics SFA model.

Variable	*N*	Mean	*SD*	Minimum	Maximum
Teaching hours	53	5,744,117	4,168,754	473,279	22,346,943
Number of courses offered	53	175.7	81.7	32	439
Total expenditure	53	82,688,314	53,725,118	12,324,312	288,974,000

Table 2. Descriptive statistics inefficiency model.

Variable	*N*	Mean	*SD*	Minimum	Maximum
Student age	53	32.92	2.18	27.57	37.07
Students with disability (%)	53	9.60	2.86	4.74	18.55
English second language (%)	53	16.79	9.82	4.62	40.24
Remoteness (ARIA)	53	2.05	0.97	1.06	4.74
Certificate 3 or higher (%)	53	82.44	7.81	59.84	96.71
Apprentices & Trainees (%)	53	16.90	7.24	4.47	49.48
Load pass rate (%)	53	81.43	6.61	57.03	94.25
Completion rate (dev from mean)	53	−0.92	8.09	−16.40	18.30

Table 3. Estimates for OLS and SFA models.

Variables	OLS		MLE			
	Model 1 Est	*P*>|*t*|	Model 2 Est	*P*>|*z*|	Model 3 Est	*P*>|*z*|
Stochastic frontier						
Number of courses offered	0.553	<.001	0.345	<.001	0.117	.077
Total expenditure	0.926	<.001	0.989	<.001	0.981	<.001
Constant	−4.221	<.001	−4.022	<.001	−2.827	<.001
Inefficiency Model						
Student age					0.221	.284
Students with disability					−0.353	<.001
English second language					0.018	.700
Remoteness (ARIA)					1.690	<.001
Certificate 3 or higher					0.069	.073
Apprentices & Trainees					0.087	.076
Load pass rate					−0.213	.001
Completion rate					0.075	.184
Constant					−1.488	.882
*R*squared	0.913					
Wald Chi-sq			385.4	<.001	872.9	<.001
Sigma *v*			0.126	<.001	0.105	<.001
Sigma *u*			0.387	<.001		
Sigma2			0.165	<.001		
Lambda			3.073	<.001		
Gamma			0.904			

(Cullinane & Song, 2006). The *R*-squared estimate of the OLS is, with 0.913, very substantial and indicates that most of the variation in teaching hours can indeed be explained by total expenditure and the number of courses taught by institute. The total expenditure and number of courses offered are highly significant and both exhibit the sign that would be expected, e.g. higher salary 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, we find that in the frontier model both predictors are significant. 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 very high and denotes that 90% of the variability in delivered teaching hours could be attributed to technical inefficiencies. The closeness of γ to 1 indicates 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 whether the units investigated by our Cobb-Douglas model use constant returns to scale technology.

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 'salary expenditure', 'other expenditure', 'number of courses' and 'number of campuses' was calculated as 1.33 and the test for equality to 1 yielded a chi-squared value of 15.76 ($p < .001$). We thus are able to 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 full 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 slightly higher magnitude. The main items of interest in Model 3 are thus the inefficiency effects. We note that the mean student age, proportion of English as a second language and overall course completion rate are not related to institutional inefficiency. The strong significance of remoteness points to inefficiency being a function of remoteness. This result confirms the findings of Fieger et al. (2010), who found remoteness was 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 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, and Crouchley (2002) found some incidental relationship between remoteness and inefficiency. The load pass rate (percentage of successfully completed subjects) is highly significant and indicates that lower subject completion rates are related to institutional inefficiency. We speculate that it is rather efficient TAFE administration that influences lower load pass rates than the other way around. The percentage of students with a disability is negatively related to institutional inefficiency. This is a surprising result and appears to be counterintuitive. Generally, one would expect that institutions' need to accommodate larger numbers of disabled students increases their costs and thus decreases efficiency. It is, however, possible that the institutional costs of providing facilities and support for disabled students are reasonably constant and independent of the number of such students. This appears to be the case here, and it may be factors such as disabled students enrolling in larger numbers in courses with a lower cost base that account for the detected negative relationship between disability and inefficiency.

In Model 3 we find further, albeit weaker, positive associations between the proportion of students enrolled in Certificates III or higher, and the proportions of apprentices and trainees. This finding represents an expected result, as it can be expected that an increase in both categories results in higher costs, which would indicate lower efficiencies if the number of delivery hours is constant.

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

$$TE = \left\{ \frac{1 - \Phi(\sigma_* - \mu_{*i}/\sigma_*)}{1 - \Phi(-\frac{\mu_{*i}}{\sigma_*})} \right\} \exp\left(-\mu_{*i} + \frac{1}{2}\sigma_*^2 \right) \tag{15}$$

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

$$\mu_{*i} = -\epsilon_i \sigma_u^2 / \sigma_s^2 \tag{16}$$

and

$$\sigma_* = \sigma_u \sigma_v / \sigma_s. \tag{17}$$

The calculated efficiencies for Model 3 can be found in Table 4.

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 the marginal cost of additional 'hours taught' tend to fall 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 & 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.

Table 4. Observed institutional efficiencies.

Institute	Efficiency	Institute	Efficiency	Institute	Efficiency	Institute	Efficiency
74	0.18	25	0.87	19	0.93	65	0.95
58	0.31	11	0.87	15	0.93	38	0.95
56	0.45	48	0.88	49	0.94	37	0.96
40	0.60	26	0.88	1	0.94	13	0.96
57	0.64	45	0.89	10	0.94	66	0.96
77	0.75	27	0.90	28	0.94	36	0.96
52	0.75	55	0.90	51	0.94	50	0.96
14	0.78	64	0.90	30	0.94	34	0.97
22	0.80	4	0.91	46	0.95	33	0.97
24	0.83	16	0.91	5	0.95	18	0.98
32	0.84	20	0.93	17	0.95	71	0.99
53	0.84	35	0.93	31	0.95		
7	0.86	47	0.93	29	0.95		
23	0.87	43	0.93	44	0.95		

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

$$E = 1 - \frac{3.1^{*}10^{5}}{S} \tag{18}$$

where S indicates size as measured by teaching hours. The resulting fit explains about 88% 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 (18) yields

$$\frac{dE}{ds} = \frac{3.1^{*}10^{5}}{s^{2}}. \tag{19}$$

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

$$T = \sqrt{3.1^{*}10^{5*}TH_{max}} = 2.6^{*}10^{6} \tag{19}$$

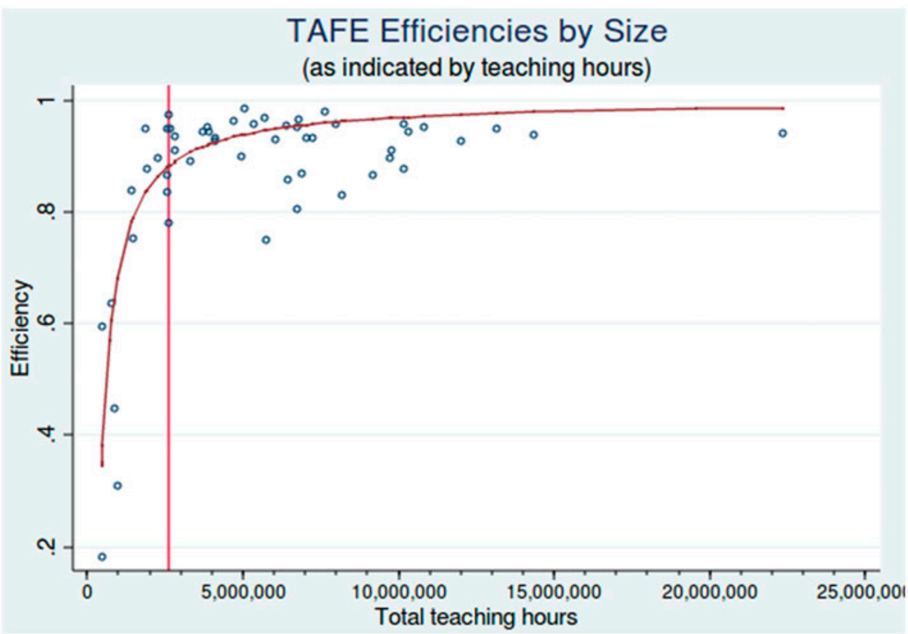

Figure 1. Technical efficiencies as a function of institute size (vertical line indicates 2.6 million delivery hours).

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 2.6 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 and administrative resources to units taught, TAFE institutes should be of a size that corresponds to at least 2.6 million teaching hours. Interestingly, this finding is similar to the results presented in Fieger et al. (2010), where a different methodology, different data and variables, and a different base year were employed.

This similarity may add impetus to the validity of our findings presented here and also, to a certain extent, may point to the comparability between DEA and SFA when employed in the Australian TAFE context. From a policy perspective, however, it does need to be pointed out that there are considerations in addition to efficiency that need to be considered when funding and evaluating TAFEs. These include the provision of educational services such as TAFEs in areas that would otherwise lack such facilities, which in turn may result in considerable negative social, educational and labour market consequences. This limitation could serve as an argument to restrict the drawing of policy implications from this type of efficiency analysis to a setting of comparable institutions in respect to social and geographic environments.

Conclusion

In this study we have applied a stochastic frontier model to estimate the efficiencies of Australian TAFE institutes, focussing on the relationship between financial and administrative inputs and teaching output. We have observed some clear inefficiencies. These were mainly related to the degree of remoteness; the institutional load pass rate and proportions of disabled students, apprentices and trainees; or students enrolled in courses of Certificate III or higher. Specifically, the uncovering of a relationship between increased inefficiency and the proportion of apprentices and trainees and enrolees in Certificates III and higher represents a new and interesting finding. The research presented here could also confirm earlier findings of a strong relationship between remoteness and inefficiency.

This study found significant economies of scale effects in the Australian TAFE system. These effects diminish once institutions exceed a certain minimum threshold in size. In this research, we determined that there are impediments for efficiency in institutes with less than 2.6 million teaching hours. While this finding has some limitations, it may be of value for policymaking decisions that deal with the restructuring of institutes that are not rural.

Finally, in this analysis we have shown that stochastic frontier analysis represents a valuable and alternative tool for the estimation of efficiency in educational institutions. While often producing similar results to data development analysis, stochastic frontier analysis can offer additional insights into efficiency analysis, such as more robust indicators concerning the quality of the achieved results, lower sensitivity to outliers and the option of statistical hypothesis testing.

Disclosure statement

The authors report no potential conflict of interest.

References

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

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

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

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

ANR. (2011). Annual National Report of the Australian Vocational Education and Training System 2011. Department of Industry, Innovation, Climate Change, Science, Research and Tertiary Education, Canberra. Retrieved October 2013, from http://www.innovation.gov.au/skills/ResourcesAndPublications/Documents/ANRVET2011.pdf

Avkiran, N. K. (2001). Investigating technical and scale efficiencies of Australian universities through data envelopment analysis. *Socio-Economic Planning Sciences, 35*, 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*, 325–332.

Carrington, R., Coelli, T., & Rao, P. (2005). The performance of Australian Universities: Conceptual issues, and preliminary results. *Economic Papers, 24*, 145–163.

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

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

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

Cullinane, K., Wang, T. F., Song, D. W., & Ji, P. (2006). The technical efficiency of container ports: Comparing data envelopment analysis and stochastic frontier analysis. *Transportation Research Part A: Policy and Practice, 40*, 354–374.

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

Fieger, P., Karmel, T., & Stanwick, J. (2010). *An investigation of TAFE efficiency*. Adelaide: National Centre for Vocational Education Research.

Fiorentino, E., Karmann, A., & Koetter, M. 2006, October. The cost efficiency of German Banks: A comparison of SFA and DEA. Retrieved from SSRN: http://ssrn.com/abstract=947340 or http://dx.doi.org/10.2139/ssrn.947340

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

Horne, J., & Hu, B. (2008). Estimation of cost efficiency of Australian universities. *Mathematics and Computers in Simulation*, *78*, 266–75.

Hossain, M. K., Kamil, A. A., Baten, M. A., & Mustafa, A. (2012). Stochastic frontier approach and data envelopment analysis to total factor productivity and efficiency measurement of Bangladeshi rice. *PloS one*, 7(10), e46081.

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*, 63–71.

Jacobs, R. (2001). Alternative methods to examine hospital efficiency: Data envelopment analysis and stochastic frontier analysis. *Health Care Management Science, 4*, 103–115.

Karmel, T., Fieger, P., Blomberg, D., & Loveder, P. (2013). *Performance indicators in the VET sector.*, Discussion paper, National Centre for Vocational Education Research, Adelaide

Mark, K., & Karmel, T. (2010). *The likelihood of completing a VET qualification: A model based approach*. Technical paper. Adelaide: NCVER.

Parsons, L. J. (2004). *Measuring performance using stochastic frontier analysis: An industrial salesforce illustration*. ISBM Report, 6 –2004.

Waldman, D. M. (1982). A stationary point for the stochastic frontier likelihood. *Journal of Econometrics, 18*, 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*, 129–144.

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