A DEA approach for measuring university departments' efficiency

Tzeremes, Nickolaos and Halkos, George

University of Thessaly, Department of Economics

2010

A DEA approach for measuring university departments' efficiency

by

George E. Halkos[*]**, Nickolaos G. Tzeremes, Stavros A. Kourtzidis**

Department of Economics, University of Thessaly, Korai 43, 38333, Volos, Greece

Abstract

This paper uses Data Envelopment Analysis (DEA) in order to determine the performance levels of 16 departments of a public owned university. Particularly, the constant returns to scale (CRS) and variable returns to scale (VRS) models have been applied alongside with bootstrap techniques in order to determine accurate performance estimates. The study illustrates how the recent developments in efficiency analysis and statistical inference can be applied when evaluating institutional performance issues. The results reveal the existence of misallocation of resources or/and inefficient application of departments' policy development.

Keywords: Departments' efficiency; Data Envelopment Analysis; Bootstrap techniques; Kernel density estimation.

JEL Codes: C60, C67, I20, I23

---

[*] Corresponding author. Email: halkos@econ.uth.gr

1. Introduction

A sufficient number of studies have investigated institutions' efficiency and came across with several problems. According to Johnes and Johnes (1993), a critical issue in measuring the efficiency of higher education institutes, is how to aggregate the heterogeneous inputs and outputs, in the absence of market prices. In order to measure the efficiency, performance indicators (PIs) were developed, each of which measures the input or the output of a homogeneous set of products. The most commonly used PI in the case of universities is the number of publications (Moed et al. 1984; Harris 1988; Johnes 1990). However, Glass et al. (2006) argue that PIs focus only on one variable, without being capable of including the multiple inputs and outputs that are necessary in higher education institutes. Also, PIs fail to aggregate multiple inputs and outputs because they are not able to provide objective weights, which could assist to succeed it.

An alternative approach of assessing the efficiency is the econometric method, which defines a production function and assumes that deviations from it are composed of two terms, inefficiency and error. The error term represents randomness and includes the exogenous factors as well as the econometric error, which follows the normal distribution. Important features of the econometric approach are the assumption of production technology and the strict parametric nature (Worthington 2001). The econometric approach have led to the development of the stochastic frontier approach (SFA) and has been applied by several researchers in order to evaluate the performance of higher education institutes (Verry and Layard 1975; Graves, Marchand and Thompson 1982; Hirsch et al. 1984; Johnes 1988, 1997; Cohn, Rhine and Santos 1989; De Groot, McMahon and Volkwein 1991; Glass et al. 1995; Johnes 1996; Izadi et al. 2002).

Another way to measure efficiency is the mathematical approach and its basic tool is Data Envelopment Analysis (DEA), which is a suitable tool for assessing the performance in

higher education (Bougnol and Dula 2006). DEA measures the relative efficiency of an institute and objectivity is the most important advantage provided. The efficiency of each DMU measured as the ratio of weighted outputs to weighted inputs, where the weights are not assigned a priori but are calculated so as to reflect the DMU at its most efficient value relative to the other DMUs (Johnes 2006). In opposition to the previous approach, DEA makes no assumptions regarding the distribution of inefficiencies or the functional form of the production function (Banker, Conrad and Strauss 1986). DEA offers freedom in the selection of the variables, which can be measured in different units. An important advantage is the calculation of shadow prices and slack variables (Stiakakis and Fouliras 2009). Specifically, shadow prices are able to define which efficient Decision Making Unit (DMU) is a benchmark for the inefficient under assessment DMU (Johnes and Johnes 1993).

However, DEA assumes that deviations from the efficient frontier are the result of inefficiency. This could lead to overstatement or understatement of the results while there are no assumptions regarding the exogenous factors or measurement error. Also, its non-stochastic nature does not allow confidence intervals to be calculated. However this has been tackled by Atkinson and Wilson (1995) and Simar and Wilson (1998, 2000) who use a bootstrap methodology, which applies Monte Carlo techniques in order to approach the distribution and to calculate confidence intervals.

The approach of DEA has been used for higher education institutes in many countries such as Australia (Madden et al. 1997, Avrikan 2001 and Abbott and Doucouliagos 2003), China (Ng and Li 2000), Germany (Fandel 2007), United Kingdom (Athanassopoulos and Shale 1997, Sarrico et al. 1997, Flegg et al. 2004) and USA (Colbert et al. 2000).

Our study, by applying the above advances of statistical inference in DEA models, measures the departments' efficiency of a state owned Greek university, the University of Thessaly. Moreover, the paper demonstrates how bootstrap techniques can be applied into

institution efficiency measurement and thus to obtain bias corrected efficiency estimates and confidence intervals, in contrast with the straightforward applications of DEA techniques.

The paper is organized as follows. Section 2 reviews the existing relative literature whereas section 3 presents the various variables used in the formulation of the proposed models. In section 4 the techniques adopted both in theoretical and mathematical formulations are presented. Section 5 discusses the empirical findings of our study. The final section concludes the paper commenting on the derived results and the implied policy implications.

2. Literature Review

Lindsay (1976) argues that a public principal does not measure the value of a product by its market price, but from its characteristics. Public authority can evaluate only the most obvious characteristics and this implies that economic resources are directed towards them. On the contrary, private enterprises evaluate all the characteristics of a product. Sisk (1981) applied Lindsay's theory to academic institutions, however he used only one input and one output. Ahn et al. (1988) extended Sisk's research by adding multiple inputs and outputs and used a DEA model to check the hypothesis that public universities are more efficient than private universities. They used capital and labour as inputs and teaching and research as outputs, measured by the number of full time equivalents separately for undergraduate and postgraduate teaching and the amount of federal grants and contracts respectively.

Tomkins and Green (1988) measured the efficiency of twenty accounting departments of English universities by running six DEA models. Particular interest presents the inclusion of research postgraduate students, as well as the number of publications as a measure for research and the number of academic staff as a measure for teaching. Johnes and Johnes (1993) divided publications into categories: papers in academic journals, letters in academic journals, articles in professional journals, articles in popular journals, authored

books, edited books, published official reports and contributions to edited works. Moreover, an article was identified if it was published in a journal included in Diamond's list (Diamond 1989).

Madden, Savage and Kemp (1997) included as inputs the number of auxiliary staff and administrative staff except from academic staff. Furthermore, the authors argue that the appropriate measure of teaching is the number of graduating students because it incorporates the quality into teaching under the assumption that more graduating students implies higher teaching quality. Flegg et al. (2004) and Johnes and Yu (2008) support that the number of students must be included as an input along with capital and labour.

All researches mentioned so far measure the efficiency among similar departments of different universities. Sinuany-Stern, Mehrez and Barboy (1994) were the first who measured the efficiency among departments of the same university and specifically at Ben-Gurion University. The authors state that although there are many drawbacks in this kind of analysis because DEA assumes that DMUs (departments in our case) are homogeneous, which is hard to be the case, there are many advantages. A simple qualitative analysis may not be enough and suffers for subjective biases, while DEA can provide objective quantitative measures which at least means that it is a valuable complementary approach for the decision maker. The same direction is followed by other researches such as King (1997), Arcelus and Coleman (1997), Sarrico and Dyson (2000), Tauer, Fried and Fry (2007) and Kao and Hung (2008).

As Johnes (2006) states, one of the advantages of DEA, is that it calculates one simple score for each DMU which is easily comprehensible from everyone. The drawback of the basic DEA technique is that it provides no indication whether these simple scores vary statistically significant. Bootstrap techniques have been developed in order to overcome this problem and they are used to estimate 95% confidence intervals for each DMU (Simar and Wilson 1998, 2000). Moreover, these techniques are applied to DEA estimators which are

biased by construction and eliminate this bias. Bootstrap techniques have been used in higher education education by Johnes (2006a, 2006b).

3. Data

As a public institution, university uses multiple inputs to produce multiple outputs. In this study we use as inputs the number of academic staff, the number of auxiliary staff (teaching aide staff, technical and administrative staff), the number of students (undergraduates, postgraduates, doctorate students) and total income (governmental funding). Kao and Hung (2008) used preassigned weights for undergraduate and postgraduate students. Following Kao and Hung (2008) methodology, we pre-assigned weights to the academic staff, auxiliary staff and number of students.

The number of academic staff is used commonly in literature (Tomkins and Green 1988; Johnes and Johnes 1993) and it is constituted only by faculty members. There are four ranks of faculty members (professor, associate professor, assistant professor and lecturer), so we assigne weights to each rank in order to construct a proper aggregated measure of academic staff (Madden, Savage and Kemp 1997). Weights are assigned based on the assumption that a professor is expected to produce more research work than a lecturer. Thus, professors are assigned with 1, associate professors with 0.75, assistant professors with 0.5 and lecturers with 0.25. These weights are chosen so the distance between two ranks is $1/4=0.25$.

The second input, also used by Arcelus and Coleman (1997) and Madden, Savage and Kemp (1997), is the auxiliary staff, which is constituted by teaching aide, technical and administrative staff. This input is used under the assumption that teaching, administrative and technical duties have a negative influence on the research of academic staff because they have an outcome in limiting their available time for research. Therefore, higher auxiliary staff means higher expected research (Johnes 1988). We assigned weights to each

category of auxiliary staff as before. Teaching aide staff was assigned with 1, while technical and administrative staff is assigned with 0.5.

The third input is the number of students, which according to Flegg et al. (2004) and Johnes and Yu (2008) can be included as an input. In contrast with other studies, total number of students is preferred from full-time equivalents (Agasisti and Johnes 2010). Like academic staff, there are three student ranks (undergraduates, postgraduates and doctorate students) so we assigne weights to each one. Thus, doctorate students are assigned with 1, postgraduates with 0.666 and undergraduates with 0.333.

The fourth input is the total income from research which is used by the vast majority of the literature in many forms (Tomkins and Green 1988; Beasley 1990; Sinuany-Stern, Mehrez and Barboy 1994; Athanassopoulos and Shale 1997). Sometimes income can be found as total income or total grants and other times can be found as income from research or from other sources.

As it is widely accepted by in the entire literature, the outputs that are produced by a university are teaching and research. Some researches measure teaching according to the hours a professor teaches, which is a convenient approach giving the fact that it's easy for a researcher to gather this data. However, this measure does not include the quality of teaching. A simple way to accomplish this is to measure the number of graduating students. The hypothesis is that higher number of graduating students means higher quality of teaching (Madden, Savage and Kemp 1997). Once more, we assign weights to each student rank. Thus, postgraduates are assigned with 1 and undergraduates with 0.5.

Academic research is the most controversial output. Although it is widely accepted as an output, it can be measured in various ways. The two core ways to measure research is the income from research (Ahn, Charnes and Cooper 1988, Beasley 1990, 1995; Flegg et al. 2004) and the number of publications (Sinuany-Stern, Mehrez and Barboy 1993; Johnes and Johnes 1993; Johnes and Yu 2008). In the first case, the argument is that more significant

research will attract more income. However, this is an indirect measurement, while the number of publications is a direct measurement of academic research and we prefer to use it in our research. Furthermore, income from research does not reflect academic research (papers, conferences etc.) in Greek Universities but income from other research activities, which lead us to treat "income from research" as any other income and use it as an input, while number of publications is used as an output.

A critical question is how many journals will be used in the research. The inclusion of a very small number of journals might bias the result in favour of departments which produce general research against the departments which produce specialized research. On the contrary, the inclusion of too many journals means that an article in an infamous journal has the same value with an article in a famous journal (Johnes 1988). Many researches have used only the articles published in the most reputable journals, but these researches refer to British universities in most cases, whereas academic staff tends to publish in widely recognized journals (Johnes 1988). According to Harris (1988), Australian academics, with a few exceptions, tend to publish in less recognized journals. This proposition stands for Greek academics too. Thus, we followed Harris' research and we included all articles in refereed journals.

According to Carrington, Coelli and Prasada Rao (2005) and Worthington and Lee (2008) "weighted publications" is the most suitable measure of research. Thus, in academic research the following categories with their weights are included. Articles in foreign journals are assigned with 1, articles in Greek journals with 0.75, books, monographs and chapters in books are considered of the same value and are assigned with 0.50 and articles in conferences with 0.25. Along with articles in conferences we measure discussion papers in the same category, as Madden, Savage and Kemp (1997) used in their research.

Dyson et al. (2001) raised some issues that must be examined in a DEA model. In the present paper, we will deal with two of these issues, the homogeneity of DMUs and the

number of variables. In order to be homogeneous, DMUs must have a similar range of activities and produce similar outputs. The activities of all the departments are teaching and research. Teaching is measured by the number of graduating students and research is measured by the number of publications which are both similar for all the departments. However, it would be useful if we could include other forms of research such as laboratorial research (however it is difficult to be measured). Additionally, DMUs must use a similar range of inputs, as it is the fact in our case. Our inputs are the number of academic staff, the number of auxiliary staff, the number of students and the total income, which are all similar for every department. The last hypothesis of homogeneity is that all DMUs operate in a similar environment, which is true because all departments operate under the legal framework that is the same for all the Greek universities. Moreover, departments operate under the framework of the same university.

According to Dyson et al. (2001) the number of DMUs must be at least $2 \times m \times s$ where m is the number of inputs and s the number of outputs. In our case $2 \times 4 \times 2 = 16$ is equal with the number of DMUs under evaluation indicating a "proper" number of inputs/outputs used.

The data for the number of academic and auxiliary staff, the number of undergraduate and postgraduate students, the number of graduating students and total income were collected from the annual internal report of Evaluation Quality Unit of the University of Thessaly, from the Office of Academic Affairs and from the departments' secretariats and they refer to the period 2008-2009. The data for the publication were provided from the departments' official websites and from annual internal report of the Evaluation Quality Unit.

4. Methodology

4.1 Efficiency measurement

Efficiency analysis was dated back to the work of Debreu (1951), Koopmans (1951) and Farrell (1957) who were the first to measure empirically the efficiency of production units. Following the notation by Simar and Wilson (2008) we can imply that the process of production is constrained by the production set Ψ which is the set of physically attainable points (x, y) so that:

$$\Psi = \left\{ (x, y) \in \Re_+^{N+M} \middle| x \quad can \quad produce \quad y \right\} \tag{1}$$

where $x \in \Re_+^N$ is the input vector and $y \in \Re_+^M$ is the output vector. In that respect the efficient boundary of Ψ is the locus of optimal production plans. This boundary is called the production frontier and can be expressed as:

$$\partial\Psi = \left\{ (x, y) \in \Psi \middle| (\theta x, y) \notin \Psi, \forall 0 < \theta < 1, (x, \lambda y) \notin \Psi, \forall \lambda > 1 \right\} \tag{2}.$$

According to Daraio and Simar (2007) the locus of optimal production plans can be either input or output oriented. In the input oriented framework the input requirement set and its efficient boundary aims to reduce the input amounts keeping the present output levels. In contrast the output oriented framework seeks to maximize the output levels keeping the present input levels. The choice between input and output orientation is based on whether the decision maker controls most the inputs or the outputs. This study uses the assumption of output orientation since public universities have greater control of the research produced and the graduates (outputs). In contrast with the inputs which the amounts of are directly controlled by the Greek Ministry of Education, Lifelong Learning and Religious Affairs and

indirectly by the Universities' departments. Therefore, the production set Ψ is characterized by output feasibility sets defined for all $x \in \Re_+^N$ as:

$$Y(x) = \left\{ y \in \Re_+^M \middle| (x, y) \in \Psi \right\} \tag{3},$$

and the output oriented efficiency boundary $\partial Y(x)$ is defined for a given $x \in \Re_+^N$ as:

$$\partial Y(x) = \left\{ y \middle| y \in Y(x), \lambda y \notin Y(x), \forall \lambda > 1 \right\} \tag{4}.$$

Then the Debreu-Farrell output measure of efficiency for a production unit located at $(x, y) \in \Re_+^{N+M}$ is:

$$\lambda(x, y) = \sup \left\{ \lambda \middle| (x, \lambda y) \in \Psi \right\} \tag{5}.$$

The DEA estimator was first operationalized as linear programming estimators by Charnes, Cooper and Rhodes (1978) assuming the free disposability and the convexity of the production set Ψ. It involves measurement for a given unit (x, y) relative to the convex hull of $X_n = \left\{ (x_i, y_i), i = 1, ..., n \right\}$ and it assumes constant returns to scale (CRS):

$$\hat{\lambda}_{CRS}(x, y) = \sup \left\{ \begin{array}{l} \lambda \middle| \lambda y \leq \sum_{i=1}^n \gamma_i y_i; x \geq \sum_{i=1}^n \gamma_i x_i \quad for \quad (\gamma_1, ..., \gamma_n) \\ such \quad that \quad \gamma_i \geq 0, i = 1, ..., n \end{array} \right\} \tag{6}.$$

Later, Banker, Charnes and Cooper (1984) developed a DEA estimator allowing for variable returns to scale (VRS) as:

$$\hat{\lambda}_{VRS}(x, y) = \sup \left\{ \begin{array}{l} \lambda \middle| \lambda y \leq \sum_{i=1}^n \gamma_i y_i; x \geq \sum_{i=1}^n \gamma_i x_i \quad for \quad (\gamma_1, ..., \gamma_n) \\ such \quad that \quad \sum_{i=1}^n \gamma_i = 1; \ \gamma_i \geq 0, i = 1, ..., n \end{array} \right\} \tag{7}.$$

4.2 Efficiency bias correction and confidence internals construction

DEA estimators are biased by construction and thus biased correction techniques need to be adopted for the improvement of the efficiency scores obtained. Following Simar and Wilson (1998, 2000) we perform the bootstrap procedure for the DEA estimators in order to obtain biased corrected results. The bootstrap procedure is a data-based simulation method for statistical inference (Daraio and Simar 2007, p.52). Some of its main applications[1] are the correction for the bias and construction of confidence intervals of the efficiency estimators (Simar and Wilson, 1998; 2000), applications to Malmquist indices (Simar and Wilson, 1999), statistical procedures for comparing the efficiency means of several groups (Simar and Wilson 2008), test procedures to assess returns to scale (Simar and Wilson, 2002) and criterion for bandwidth selection (Simar and Wilson, 2002; 2008).

The bootstrap bias estimate for the original DEA estimator $\hat{\lambda}_{DEA}(x, y)$ can be calculated as:

$$\hat{BIAS}_B\left(\hat{\lambda}_{DEA}(x, y)\right) = B^{-1}\sum_{b=1}^{B}\hat{\lambda}^*_{DEA,b}(x, y) - \hat{\lambda}_{DEA}(x, y) \tag{8}$$

Furthermore, $\hat{\lambda}^*_{DEA,b}(x, y)$ are the bootstrap values and B is the number of bootstrap replications. Then a biased corrected estimator of $\lambda(x, y)$ can be calculated as:

$$\hat{\hat{\lambda}}_{DEA}(x, y) = \hat{\lambda}_{DEA}(x, y) - \hat{BIAS}_B\left(\hat{\lambda}_{DEA}(x, y)\right) = 2\hat{\lambda}_{DEA}(x, y) - B^{-1}\sum_{b=1}^{B}\hat{\lambda}^*_{DEA,b}(x, y) \tag{9}$$

However, according to Simar and Wilson (2008) this bias correction can create an additional noise and the sample variance of the bootstrap values $\hat{\lambda}^*_{DEA,b}(x, y)$ need to be calculated. The calculation of the variance of the bootstrap values is illustrated below:

$$\hat{\sigma}^2 = B^{-1}\sum_{b=1}^{B}\left[\hat{\lambda}^*_{DEA,b}(x, y) - B^{-1}\sum_{b=1}^{B}\hat{\lambda}^*_{DEA,b}(x, y)\right]^2 \tag{10}$$

In addition we need to avoid the bias correction illustrated in (9) unless:

$$\frac{\left| B I \hat{A} S_B (\hat{\lambda}_{DEA}(x,y)) \right|}{\hat{\sigma}} > \frac{1}{\sqrt{3}} \tag{11}.$$

By expressing the output oriented efficiency in terms of the Shephard (1970) output distance

function we can construct bootstrap confidence intervals for $\hat{\delta}_{DEA}(x,y)$ as:

$$\left[\hat{\delta}_{DEA}(x,y) - \hat{\alpha}_{1-a/2}, \hat{\delta}_{DEA}(x,y) - \hat{\alpha}_{a/2} \right] \tag{12}$$

4.3 A bootstrap test for choosing CCR or BCC model

In order to choose between the adoption of the results obtained by the CCR

(Charnes, Cooper and Rhodes 1978) and BCC (Banker, Charnes and Cooper 1984) models

in terms of the consistency of our results obtained we adopt the method introduced by Simar

and Wilson (2002). Therefore, we compute the DEA efficiency scores under the CRS and

VRS assumption and by using the bootstrap algorithm described previously we test for the

CRS results against the VRS results obtained such as:

$$H_o : \Psi^\vartheta \quad is \quad CRS \quad against \quad H_1 : \Psi^\vartheta \quad is \quad VRS \tag{13}$$

The test statistic can be computed as:

$$T(X_n) = \frac{1}{n} \sum_{i=1}^{n} \frac{\hat{\lambda} crs, n(X_i, Y_i)}{\hat{\lambda} vrs, n(X_i, Y_i)} \tag{14}$$

Then the p-value of the null hypotheses can be approximated by the proportion of bootstrap

samples as:

$$p - value = \sum_{b=1}^{B} \frac{I(T^{*,b} \leq T_{obs})}{B} \tag{15}$$

where B is 2000 bootstrap replications, I is the indicator function and $T^{*,b}$ is the bootstrap samples. Finally, the original observed values are denoted by T_{obs}.

5. Empirical Results

Firstly we test for the existence of constant or variables returns to scale (equations 13-15) and by approximating the p-value by using the bootstrap algorithm described previously we obtained for this test a p-value of 0.98> 0.05 (with B=2000) hence, we cannot reject the null hypothesis of constant returns to scales and thus the CCR model need to be adopted in our analysis[2]. Table 1 report the results obtained under the hypothesis of constant returns to scale (however, the VRS estimators are very similar to the CRS estimators). As can be realised the departments of: Primary Education, Medical School, Veterinary Science, Physical Education & Sport Science and the department of Economics are reported to be efficient (efficiency score =1). Whereas, the lowest performances are reported for the Departments of Special Education (0.558) and the department of Computer & Communication Engineering (0.637). In addition the department of Biochemistry & Biotechnology (0.939) and the department of Ichthyology & Aquatic Environment (0.925) are reported to have high efficiency scores. When we apply the bootstrap algorithm on the efficiency scores obtained we calculate the biased corrected efficiency scores (CRS BC) along side with the estimated bias (Bias) and its standard deviation (std). As can be realized under the bias correction the efficiency scores have changed significantly however the departments with lowest performance are reported to be the same, these are the departments of Special Education (0.49) and Computer & Communication Engineering (0.549).

The biased corrected results indicate that the departments of: Primary Education, Medical School, Veterinary Science, Physical Education and Sport Science and Economics are reported to have the highest efficiency scores. But a closer look is needed on the lower (LB) and upper (UB) bounds before any conclusions can be made. Indeed the departments

of Economics and Medical School have winder bounds compared to the other departments indicating that the biased efficiency scores may have higher values compared to the other university departments. Similarly the departments of Primary Education, Veterinary Science, Physical Education and Sport Science, Biochemistry and Biotechnology and Ichthyology and Aquatic Environment have greater ranges of biased corrected efficiency scores. This variation indicates the different resource allocation and research policies among the universities departments implying greater variability in their estimated efficiencies scores.

Table 1: Estimated efficiency scores, estimated bias and estimated bias' standard deviations.

a/a	Departments	CRS	CRS (BC)	Bias	std	LB	UB
1	Mechanical Engineering	0.701	0.631	0.071	0.002	0.587	0.696
2	Urban Planning and Regional Development	0.870	0.750	0.119	0.006	0.696	0.861
3	Civil Engineering	0.730	0.608	0.122	0.005	0.581	0.724
4	Architecture	0.739	0.574	0.165	0.015	0.536	0.732
5	Computer & Communication Engineering	0.637	0.549	0.088	0.003	0.508	0.632
6	Primary Education	1.000	0.770	0.230	0.025	0.737	0.990
7	Preschool Education	0.692	0.598	0.094	0.003	0.563	0.684
8	Special Education	0.558	0.490	0.067	0.001	0.466	0.552
9	History, Archaeology and Social Anthropology	0.861	0.745	0.115	0.004	0.701	0.854
10	Agriculture Crop, Production and Rural Environment	0.899	0.804	0.095	0.003	0.755	0.892
11	Ichthyology and Aquatic Environment	0.925	0.692	0.233	0.036	0.645	0.916
12	Medical School	1.000	0.748	0.252	0.042	0.697	0.992
13	Veterinary Science	1.000	0.752	0.248	0.039	0.706	0.991
14	Biochemistry and Biotechnology	0.939	0.698	0.241	0.040	0.652	0.931
15	Physical Education and Sport Science	1.000	0.794	0.206	0.017	0.763	0.992
16	Economics	1.000	0.749	0.251	0.042	0.700	0.993

Figure 1 presents the density estimates of the original and the biased corrected efficiency estimates (CRS) alongside with the lower and upper bounds of the efficiency scores. For the calculation of the density estimates we have used the "normal reference rule-of-thumb" approach bandwidth selection (Silverman 1986) and a second order Gaussian kernel. It appears that the original CRS are leptokurtic and almost identical with the upper bound of the biased corrected efficiency scores whereas the bias corrected efficiency scores appear to be leptokurtic and quite similar with lower bounds estimates. The leptokurtic distributions indicate that there is a rapid fall-off in the density as we move away from the

mean. Furthermore, the pickedness of the distribution suggests a clustering around the mean with rapid fall around it. The density estimates appear to support graphically the previous findings which indicate that among the departments in the University of Thessaly there are different resource allocation policies and inefficiencies in the application of University's general development policy. In addition it appears that the outputs used (research and graduates) are being part of different policy perspectives among the university's departments.

Figure 1: Kernel density functions of CRS efficiency estimates using Gaussian Kernel and the appropriate bandwidth (normal reference rule-of-thumb).
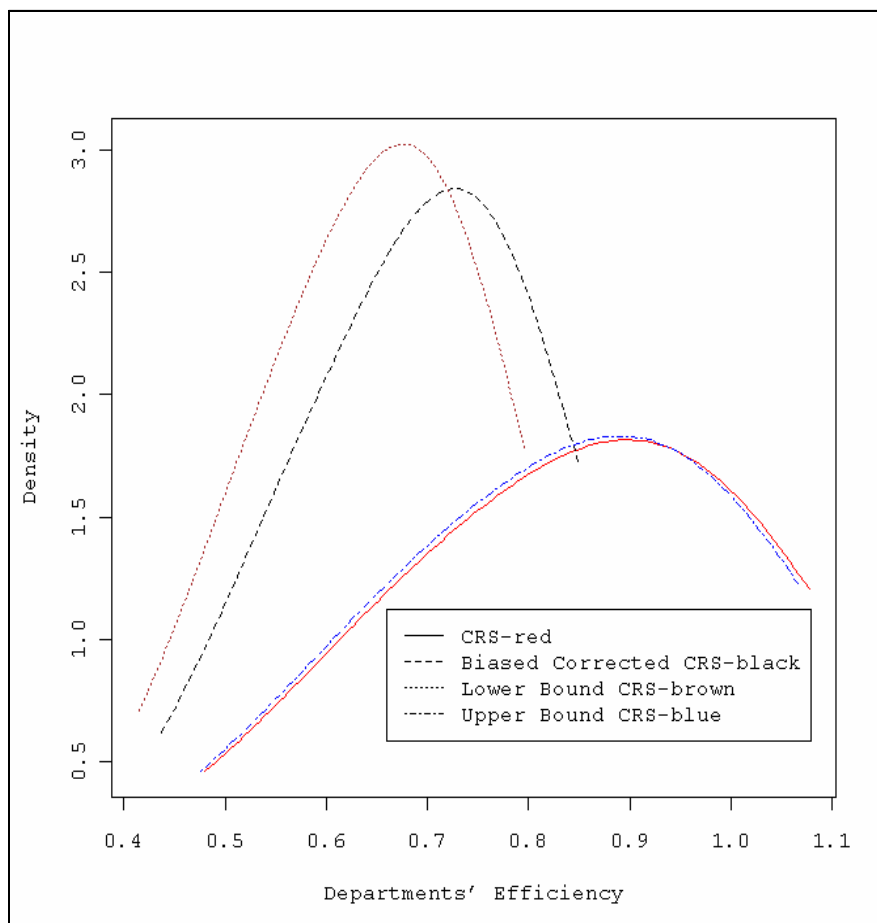
Following Banker (1984) we use the optimal values of $\sum_{i=1}^{n} \gamma_i$ which are given by the efficient departments in order to calculate the most productive scale size (MPSS) of the

inefficient departments. Table 2 provides the scale sizes that departments should operate in order to be efficient. For instance, the department of Agriculture Crop, Production and Rural Environment in order to operate at its MPSS needs to increase the research levels and the graduates' levels by 68.7%. The benchmarks (or the reference set) for the department of Agriculture Crop, Production and Rural Environment are given by the department of Primary Education and the department of Physical Education and Sport Science. It seems difficult to compare these three departments to its thematic and scientific nature however the two reference sets are more closely in terms of the amounts of inputs/outputs to the department of Agriculture Crop, Production and Rural Environment than other departments within the university and therefore they show (by providing coefficients γ_i) how inputs can be decreased and outputs increased in order to make the department under evaluation efficient.

Furthermore, Table 2 provides the relation between the proportional change of inputs and the resulting proportional change in outputs (returns to scales- RTS). As such constant returns to scale arise when a department produces n per cent increase in output by an n per cent rise in all inputs. However if output rises by a larger percentage than inputs, there are increasing return to scales (IRS). Whereas, if outputs increase by a smaller percentage than inputs, there are decreasing returns to scale (DRS). As can be realized only the department of Urban Planning and Regional Development and the department of Computer & Communication Engineering report DRS.

Table 2: Scale efficient targets and MPSS of the departments

a/a	DEPARTMENTS	Efficient Output Target (%)			$\sum \gamma_i$	RTS
		Research	Graduates	Benchmarks		
1	Mechanical Engineering	98.38	98.38	6,15	0.72	Increasing
2	Urban Planning and Regional Development	8.92	8.92	6,15	1.06	Decreasing
3	Civil Engineering	90.10	90.10	6,15	0.72	Increasing
4	Architecture	259.32	259.32	6,15	0.38	Increasing
5	Computer & Communication Engineering	43.81	43.81	6,15,16	1.09	Decreasing
6	Primary Education	0.00	0.00		1.00	Constant
7	Preschool Education	236.28	76.34	6,16	0.82	Increasing
8	Special Education	110.98	110.98	6,15,16	0.85	Increasing
9	History, Archaeology and Social Anthropology	151.79	57.53	6,16	0.74	Increasing
10	Agriculture Crop, Production and Rural Environment	68.71	68.71	6,15	0.66	Increasing
11	Ichthyology and Aquatic Environment	112.17	112.17	6,15	0.51	Increasing
12	Medical School	0.00	0.00		1.00	Constant
13	Veterinary Science	0.00	0.00		1.00	Constant
14	Biochemistry and Biotechnology	215.48	215.48	12,15,16	0.34	Increasing
15	Physical Education and Sport Science	0.00	0.00		1.00	Constant
16	Economics	0.00	0.00		1.00	Constant

6. Conclusions

This paper applies an efficiency analysis in all the departments of University of Thessaly. By applying inferential approach on DEA efficiency scores the paper measures the efficiency of 16 university departments. The majority of the existing studies similar to ours (Sinuany-Stern, Mehrez and Barboy 1994; King 1997; Arcelus and Coleman 1997; Sarrico and Dyson 2000) evaluate the performance of university departments however it is the first time (to our knowledge) that bootstrap techniques are used in DEA formulation measuring university departments' performance. Furthermore, the bootstrap techniques have provided consistency to the original biased CRS results (Simar and Wilson 1998, 2000).

Moreover, by applying the inferential approach and bootstrapped procedures we derived the general conclusion that there are strong inefficiencies among the departments, indicating misallocation of resources or/and inefficient application of departments policy developments. Additionally, the paper provides output target values for policy implications and evaluation among the departments of the University of Thessaly. Finally, this study provides evidence of how the advances and recent developments in efficiency analysis can

be applied for an effective evaluation of performance issues in public owned universities overcoming traditional DEA related problems.

Acknowledgements

Notes

1. See Halkos and Tzeremes (2010) for application of bootstrap techniques on SMEs data.

2. The results under the VRS assumption are available upon request.

References

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

Agasisti, T., and G. Johnes. 2010. Heterogeneity and the evaluation of efficiency: the case of Italian universities. *Applied Economics* 42:1365-1375.

Ahn, T., A. Charnes, and W.W. Cooper. 1988. Some statistical and DEA evaluations of relative efficiencies of public and private institutions of higher learning. *Socio-Economic Planning Sciences* 22: 259-269.

Arcelus, F.J., and D.F. Coleman. 1997. An efficiency review of university departments. *International Journal of System Science* 28: 721-729.

Athanassopoulos, A., and E. Shale. 1997. Assessing the comparative efficiency of higher education institutions in the UK by means of data envelopment analysis. *Education Economics* 5: 117-133.

Atkinson, S.E., and P.W. Wilson. 1995. Comparing mean efficiency and productivity scores from small samples: A bootstrap methodology. *The Journal of Productivity Analysis* 6: 137-152.

Avrikan, N. 2001. Investigating technical and scale efficiencies of Australian Universities through data envelopment analysis. *Socio-Economic Planning Sciencies* 35: 57-80.

Banker, R.D., R.F. Conrad, and R.P. Strauss. 1986. A comparative application of data envelopment analysis and translog methods: An illustrative study of hospital production. *Management Science* 32: 30-44.

Banker, R.D. 1984. Estimating Most Productive Scale Size Using Data Envelopment Analysis, *European Journal of Operational Research* 17: 35-44.

Banker, R.D., A. Charnes, and W.W. Cooper. 1984. Some Models for Estimating Technical and scale inefficiencies in DEA. *Management Science* 32: 1613-1627.

Beasley, J.E. 1990. Comparing university departments. *OMEGA* 18: 171-183.

Beasley, J.E. 1995. Determining teaching and research efficiencies. *The Journal of the Operational Research Society* 46: 441-452.

Bougnol, M.L., and J.H. Dula. 2006. Validating DEA as a ranking tool: An application of DEA to assess performance in higher education. *Annals of Operations Research* 145: 339-365.

Caballero, R., T. Galache, T. Gomez, J. Molina, and A. Torrico. 2004. Budgetary allocations and efficiency in the human resources policy of a university following multiple criteria. *Economics of Education Review* 23: 67-74.

Carrington, R., T. Coelli, and D.S. Prasada Rao. 2005. The performance of Australian universities: Conceptual issues and preliminary results. *Economic Papers* 24: 145-163.

Carrico, CS, S.M. Hogan, R.G. Dyson, and A.D. Athanassopoulos. 1997. Data envelopment analysis and university selection. *The Journal of the Operational Research Society* 48: 1163-1177.

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

Cohn, E., S. Rhine, and M. Santos. 1989. Institutions of higher education as multi-product firms: Economies of scale and scope. *The Review of Economics and Statistics* 71: 284-290.

Colbert, A., R.R. Levary, and M.C. Shaner. 2000 Determining the relative efficiency of MBA programs using DEA. *European Journal of Operational Research* 125: 656-669.

Daraio, C., and L. Simar. 2007. Advanced robust and nonparametric methods in efficiency analysis. Springer Science, New York.

De Groot, H., W. McMahon, and F. Volkwein. 1991. The cost structure of American research universities. *The Review of Economics and Statistics* 73: 424-431.

Debreu, G. 1951. The Coefficient of Resource Utilization. *Econometrica* 19: 273-292.

Diamond, A. 1989. The core journal of economics. *Current Contents* 21: 4-11.

Dyson, R.G., R. Allen, A.S. Camanho, V.V. Podinovski, C.S. Sarrico, and E.A. Shale. 2001. Pitfalls and protocols in DEA. European Journal of Operational Research 132: 245-259.

Fandel, G.2007 On the performance of universities in North Rhine-Westphalia, Germany: Government's redistribution of funds judged using DEA efficiency measures. *European Journal of Operational Research* 176:521-533.

Farrel, M.J. 1957. The measurement of productive efficiency. *Journal of Royal Statistical Society A* 120: 253-290.

Flegg, A.T., D.O. Allen, K. Field, and T.W. Thurlow. 2004. Measuring the efficiency of British universities: A multi-period data envelopment analysis. *Education Economics* 12: 231-249.

Glass, J.C., G. McCallion, D.G. McKillop, S. Rasaratnam, and K.S. Stringer. 2006 Implications of variant efficiency measures for policy evaluations in UK higher education. *Socio-Economic Planning Sciences* 40: 119-142.

Glass, J.C., D.G. McKillop,. and N. Hyndman. 1995. Efficiency in the provision of university teaching and research: An empirical analysis of UK universities. *Journal of Applied Econometrics* 10: 61-72.

Graves, P., R. Marchand, and R. Thompson. 1982. Economics departmental rankings: Research incentives, constraints and efficiency. *The American Economic Review* 72: 1131-1141.

Halkos, G.E., and N.G. Tzeremes. 2010 The effect of foreign ownership on SMEs performance: An efficiency analysis perspective. The *Journal of Productivity Analysis*. doi: 10.1007/s11123-010-0174-2.

Harris, G.T. 1988. Research output in Australian university economics departments, 1974-83*. *Australian Economic Papers* 27: 102-110.

Hirsch, B., R. Austin, J. Brooks, and B. Moore. 1984. Economics departmental rankings: comment. *The American Economic Review* 74: 822-826.

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

Johnes, G. 1988. Determinants of research output in economics departments in British universities. *Research Policy* 17: 171-178.

Johnes, G. 1990. Measures of research output: University departments of economics in the UK, 1984-8. *The Economic Journal* 100: 556-560.

Johnes, G. 1997. Costs and industrial structure in contemporary British higher education. *The Economic Journal* 107:727-737.

Johnes, G., and J. Johnes. 1993. Measuring the research performance of UK economics departments: An application of data envelopment analysis. *Oxford Economic Papers* 45: 332-347.

Johnes, J. 1996. Performance assessment in higher education in Britain. *European Journal of Operational Research* 89:18-33.

Johnes, J. 2006a. Data envelopment analysis and its application to the measuring of efficiency in higher education. *Economics of Education Review* 25: 273-288.

Johnes, J. 2006b. Measuring efficiency: A comparison of multilevel modelling and data envelopment analysis in the context of higher education. *Bulletin of Economic Research* 58: 75-104.

Johnes J., and L. Yu. 2008. Measuring the research performance of Chinese higher education institutions using data envelopment analysis. *China Economic Review* 19: 679-696.

Kao, C., and H.T. Hung. 2008. Efficiency analysis of university departments: An empirical study. *Omega* 36: 653-664.

King, W.D. 1997. Input and output substitution in higher education. *Economic Letters* 57: 107-111.

Koopmans, T.C. 1951. Analysis of production as an efficient combination of activities, in T.C. Koopmans, eds. Activity Analysis of Production and Allocation, Wiley, New York.

Lindsay, C. 1976. A theory of government enterprise. *The Journal of Political Economy* 84: 1061-1077.

Madden, G., S. Savage, and S. Kemp. 1997. Measuring public sector efficiency: A study of economics departments at Australian Universities. *Education Economics* 5: 153-168.

Moed, H.F., W.J.M. Burger, J.G. Frankfort, and A.F.J. Van Raan. 1985. The use of bibliometric data for the measurement of university research performance. *Research Policy* 14: 131-149.

Ng, Y.C., and S.K. Li. 2000. Measuring the research performance of Chinese higher education institutions: An application of data envelopment analysis. *Education Economics* 8: 139-156.

Sarrico, C.S., and R.G. Dyson. 2000. Using DEA for planning in UK universities-An institutional perspective. *The Journal of the Operational Research Society* 51: 789-800.

Sarrico, C.S., S.M. Hogan, R.G. Dyson, and A.D. Athanassopoulos. 1997. Data envelopment analysis and university selection. *Journal of the Operational Research Society* 48: 1163-1177.

Shephard, R.W. 1970. Theory of Cost and Production Function. Princeton University Press, Princeton, NJ.

Silverman, B.W. 1986. Density estimation for statistics and data analysis. Chapman and Hall, London.

Simar, L., and P. Wilson. 2008. Statistical interference in nonparametric frontier models: recent developments and perspectives. In: Fried, H., C.A.K. Lovell, S. Schmidt, eds. The measurement of productive efficiency and productivity change. Oxford University Press, New York.

Simar, L., and P.W. Wilson. 1998. Sensitivity analysis of efficiency scores: how to bootstrap in nonparametric frontier models. *Management Science* 44: 49–61.

Simar, L., and P.W. Wilson. 1999. Estimating and Bootstrapping Malmquist Indices. *European Journal of Operational Research* 115: 459-471.

Simar, L., and P.W. Wilson. 2000. A general methodology for bootstrapping in non-parametric frontier models. *Journal of Applied Statistics* 27: 779-802.

Simar, L., and P.W. Wilson. 2002. Nonparametric tests of returns to scale. *European Journal of Operational Research* 139: 115-132.

Sinuany-Stern, Z., A. Mehrez, and A. Barboy. 1994. Academic departments efficiency via dea. *Computers & Operations Research* 21: 543-556.

Sisk, D. 1981. A theory of government enterprise: University Ph. D. Production. *Public Choice* 37: 357-363.

Stiakakis, E., and P. Fouliras. 2009. The impact of environmental practices on firms' efficiency: the case of ICT-producing sectors. *Operational Research* 9: 311-328.

Tauer, L.W., H.O. Fried, and W.E. Fry. 2007. Measuring of academic departments within a college. *Education Economics* 15: 473-489.

Tomkins, C., and R. Green. 1988. An experiment in the use of data envelopment analysis for evaluating the efficiency of UK university departments of accounting. *Financial Accountability & Management* 4: 147-164.

Verry, D.W., and P.R.G. Layard. 1975. Cost functions for university teaching and research. *The Economic Journal* 85: 55-74.

Worthington, A. 2001. An empirical survey of frontier efficiency measurement techniques in education. *Education Economics* 9: 245-268.

Worthington, A., and B.L. Lee. 2008. Efficiency, technology and productivity change in Australian universities, 1998-2003. *Economics of Education Review* 27: 285-298.