Lancaster University Management School
Working Paper
2003/082

Measuring The Efficiency Of Universities: A Comparison Of Multilevel Modelling And Data Envelopment Analysis

Jill Johnes

The Department of Economics
Lancaster University Management School
Lancaster LA1 4YX
UK

MEASURING THE EFFICIENCY OF UNIVERSITIES: A COMPARISON OF MULTILEVEL MODELLING AND DATA ENVELOPMENT ANALYSIS

Jill Johnes

Department of Economics
Lancaster University
Lancaster LA1 4YX
United Kingdom
e-mail: **j.johnes@lancaster.ac.uk**

November 2003

ABSTRACT

Data envelopment analysis (DEA) and multilevel modelling (MLM) are applied to a data set of 54578 graduates from UK universities in 1993 in order to assess the teaching performance of universities. A methodology developed by Thanassoulis & Portela (2002) allows each individual's DEA efficiency score to be decomposed into two components: one attributable to the university at which the student studied, and the other attributable to the individual student. From the former component a measure of each institution's teaching efficiency is derived and compared to the university effects from various multilevel models. The comparisons are made within four broad subjects: pure science; applied science; social science and arts. The results show that the rankings of universities derived from the DEA efficiencies which measure the universities' own performance (i.e. having excluded the efforts of the individuals) are not strongly correlated with the university rankings derived from the university effects of the multilevel models. The data were also used to perform various university-level DEAs. The university efficiency scores derived from these DEAs are largely unrelated to the scores from the individual-level DEAs, confirming a result from a smaller data set (Johnes 2003). However, the university-level DEAs provide efficiency scores which are generally strongly related to the university effects of the multilevel models.

JEL Classification: I21, C14, C16
Keywords: data envelopment analysis, multilevel modelling, efficiency measurement, higher education

Note: I am grateful to Geraint Johnes, Steve Bradley and Pam Lenton for comments whilst preparing this paper. All errors are my own.

MEASURING THE EFFICIENCY OF UNIVERSITIES USING MICRO DATA: A COMPARISON OF MULTILEVEL MODELLING AND DATA ENVELOPMENT ANALYSIS

1. INTRODUCTION

Over the last twenty years, there have been various attempts to construct performance indicators for institutions of higher education (IHEs) using a variety of approaches. The methodology used has been largely determined by the availability of data. Thus, early attempts to construct performance indicators relied on university-level data and so used regression based methods or non-parametric methods such as data envelopment analysis (DEA) to construct the efficiency measures. More recently, data on individuals attending university have become available, and so it is possible to use techniques such as ordered probit and multilevel modelling (MLM) to derive measures of the performance of the IHEs attended by those individuals.

The multitude of methods available for constructing performance indicators, however, necessitates some comparison of conclusions derived from each possible approach. If, for example, the conclusions regarding the efficiency of IHEs vary radically with choice of technique, this has serious implications for the use of performance indicators to policy-makers. There has been some limited research into comparing performance indicators derived using various methods applied to university level data. There has been very little research into comparing efficiency measures derived using alternative techniques in the context of micro-level data. The aim of this paper is therefore to remedy this by comparing two techniques for constructing performance indicators from individual level data: MLM and DEA. The comparison is of particular interest because DEA is a non-parametric method which measures efficiency relative to a frontier, whereas MLM is a parametric method which allows observations to lie above or below the line of best fit (i.e. it is non-frontier). There is evidence that these alternative approaches (i.e. parametric and non-parametric) can provide differing efficiency measures when applied to aggregate-level data, and so it is of interest to establish whether the same is true at the micro level.

The paper is in 5 sections of which this is the first. Section 2 describes and compares possible approaches, including MLM and DEA, to measuring efficiency. A data set compiled by the Universities Statistical Record (USR) of more than 117000 students leaving university in 1993 provides the sample data on which the comparisons of MLM and DEA are made. These data are described in detail in section 3, while the results of the analyses are presented in Section 4. Conclusions are drawn in Section 5.

2. APPROACHES TO EFFICIENCY MEASUREMENT

Early parametric studies of efficiency measurement in higher education applied ordinary least squares (OLS) methods in order to estimate a production function for IHEs. Specifically, suppose that producer *i* converts *m* inputs (*x*) into output (*y*), and the process is represented by equation (1):

$$y_i = f(x_{i1},...,x_{im}) + u_i \qquad (1)$$

Performance could then be assessed by measuring the distance of an IHE from the production function. In effect, therefore, the OLS residual (u_i) provided a measure of each IHE's performance. This approach has been applied using output measures such as graduates' degree results, undergraduate non-completion rates, graduates' first destinations and research output (Johnes & Taylor 1987; 1989a; 1989b; 1990; 1992; Johnes 1996).

One problem with this approach is that the production function estimated is an average of the observed production points of all IHEs rather than a frontier around their observed production points. Thus an IHE can lie above the production function as well as below it, whereas in reality points should lie only on or inside the production frontier (and u_i should therefore be zero or strictly negative). This problem can be solved by applying instead corrected OLS (COLS) i.e. the parameters of equation (1) are estimated using OLS, and the intercept is then shifted up until all residuals (denoted by u_i) are non-positive and at least one is zero. An example of measuring efficiency in the context of education using this approach can be found in Barrow (1991).

Although COLS provides an estimated production frontier (and is therefore superior to OLS in this respect) it is worth noting that the *ranking* of IHEs is actually identical to that obtained using OLS. Moreover, both OLS and COLS assume deterministic errors i.e. the whole of the distance of an IHE from the production frontier is attributed to inefficiency, and no allowance is made for random fluctuations or measurement errors. In the context of higher education, where difficulties in measuring the inputs and outputs abound, this may be a considerable disadvantage.

More recently, a technique which allows the error to be split into two components, one a consequence of inefficiency and the other the result of stochastic error, has been applied in the context of efficiency measurement in higher education. In stochastic frontier analysis (SFA) the production frontier is written as:

$$\ln(y_i) = \ln[f(x_{i1},...x_{im})] + \varepsilon_i \qquad (2)$$

where $\varepsilon_i = v_i - u_i$, $v_i \sim N\left(0,\sigma_v^2\right)$, u_i and v_i are statistically independent and $u_i \geq 0$ (Aigner *et al* 1977). The component v_i is normally distributed and reflects stochastic errors in the data, while u_i is one-sided (typically exponential or half-normal) and reflects technical efficiency. The parameters of the function can be estimated using modified OLS (Førsund *et al* 1980; Lovell 1993) or maximum likelihood estimation (MLE) methods. Examples of education production functions estimated using SFA are Deller & Rudnicki (1993) and Kang & Greene (2002). SFA has mainly been used to estimate cost functions[1] in higher education (see, for example, Johnes 1998; 1999; Stevens 2001; Izadi *et al* 2002).

While SFA allows for stochastic errors in the data, it remains a controversial tool of analysis because of the need to assume a specific distribution for the efficiencies. Moreover, it is difficult to apply SFA in a situation where there are both multiple inputs and multiple outputs. IHEs produce at least two distinct outputs (teaching and research) and each of these broad categories can be further divided (for example,

[1] In this case, $\varepsilon_i = v_i + u_i$ where u_i and v_i are as defined in the text, since observations can lie on or above the cost function.

teaching can be divided into undergraduate, taught postgraduate and research postgraduate; research can be divided by type of research output). For this reason, DEA becomes an attractive tool for estimating the multi-product production frontiers of higher education, and indeed DEA has frequently been applied in this context for a variety of countries including the UK (Tomkins & Green 1988; Beasley 1990; 1995; Johnes & Johnes 1992; 1993; Athanassopoulos & Shale 1997; Sarrico *et al* 1997; Sarrico & Dyson 2000), the USA (Ahn *et al* 1989; Ahn & Seiford 1993), Australia (Madden *et al* 1997; Coelli *et al* 1998; Avkiran 2001; Abbott & Doucouliagos 2003) and Finland (El-Mahgary & Lahdelma 1995; Korhonen *et al* 2001).

All these approaches, however, can be criticised for their use of aggregate data. First of all, it is argued that the use of data measured at the level of the production unit does not allow for variation of within unit relationships (Woodhouse & Goldstein 1988). Second, the residuals of regression based models can vary substantially depending on the predictors included in the model (Woodhouse & Goldstein 1988), and this therefore unit rankings to vary across various possible statistical models.

Measuring the efficiency of production units using aggregate level DEA is also open to serious objections (Woodhouse and Goldstein 1988). The efficiency score of a DMU is computed as the ratio of the weighted outputs to weighted inputs. In a simple one output (y), one input (x) case (for example, y = achievement level of graduates and x = average entry score) DMU k's efficiency is measured as a simple ratio of output to input i.e. $\theta_k = y_k / x_k$. However, there is a relationship between achievement level of graduates and entry scores, and supposing the relationship takes the form $y_k = a + bx_k$, then DMU k's efficiency is represented by $\theta_k = a/x_k + b$, and is therefore inversely proportional to the input measure. This argument can be extended to the situation with multiple inputs and multiple outputs (Woodhouse and Goldstein 1988).

The availability in recent years of large data sets of individuals who attended university has resulted in a number of studies aiming to identify the significant determinants of a given output measure (for example, the earnings of graduates, the labour market destination of graduates, whether or not a student completes his degree course, and the degree results of university leavers) with a view to measuring the efficiency of the institution attended (Naylor *et al* 2000; Smith *et al* 2000; Smith & Naylor 2001a; 2001b; Bratti 2002; Bratti *et al* 2003). MLM offers an alternative method of analysis to the OLS regression, logit and probit models of these studies, and explicitly incorporates institutional effects into the relationship between individuals' outcomes and the inputs.

MLM assumes that the data to be analysed are hierarchical, for example students (level 1) are nested within universities (level 2), and universities are nested within type of IHE (level 3- for example pre- or post-1992 university). Consider a sample of IHEs producing graduates with specific degree results, and let y_{ij} denote the degree result of the *i*th student in the *j*th university, then a basic multilevel model is given by

$$\begin{aligned} y_{ij} &= \beta_{0ij} \\ \beta_{0ij} &= \beta_0 + u_{0j} + e_{0ij} \end{aligned} \qquad (3)$$

where $\beta_0 + u_{0j}$ is the university-specific contribution to degree results and e_{0ij} is an individual's deviation from the university's contribution. Furthermore, the university-specific contribution is divided into β_0, which is the mean value across all universities and u_{0j}, which is the deviation from the mean.

The term of interest in the context of measuring the efficiency of the IHEs which form the sample is u_{0j} - the amount by which each university deviates from the mean value – known, in this context, as the university effects. Since the universities are assumed to be a random sample from the population of universities, the u_{0j} are therefore distributed among universities, and are assumed to be normally distributed with mean zero and variance σ_{u0}^2. The student residuals (e_{0ij}) are also normally distributed with mean zero and variance σ_{e0}^2. The u_{0j} can be estimated as follows:

$$\hat{u}_{0j} = \frac{n_j \sigma_{u0}^2}{n_j \sigma_{u0}^2 + \sigma_{e0}^2} \cdot \frac{\sum_i \left(y_{ij} - \hat{y}_{ij} \right)}{n_j} \qquad (4)$$

where n_j = the number of students at university j. Each university's estimated effect $\hat{u}_{0j}$ has a sampling error hence confidence intervals can also be computed.

The simple model can be adapted to incorporate predictors of the dependent variable. Suppose the variables expected to have an effect on y_{ij} are $x_{1ij},...,x_{kij}$, then a multilevel model where y_{ij} depends on $x_{1ij},...,x_{kij}$, the intercepts vary across IHEs but the slope coefficients are constant, is given by:

$$\begin{aligned} y_{ij} &= \beta_{0ij} + \beta_1 x_{1ij} + ... + \beta_k x_{kij} \\ \beta_{0ij} &= \beta_0 + u_{0j} + e_{0ij} \end{aligned} \qquad (5)$$

Thus the mean intercept is β_0, but the intercepts for the individual universities lines vary around this by u_{0j} (the level 2 residuals) with variance σ^2_{u0}. In addition, each individual student's degree result varies around the universities' lines by e_{0ij} (the level 1 residuals) with variance σ^2_{e0}. The addition of the explanatory or input variables leads to the interpretation of the estimated residuals $\hat{u}_{0j}$ from this model as indicators of a university's effectiveness in terms of 'value added' i.e. having taken into account inter-university variations in the input variables.

Finally, it is worth considering the situation where both the intercept and the slope coefficients vary. For simplicity, the equations below illustrate the model where the coefficient of just one explanatory variable (x_{1ij}) varies, but this can be extended to any or all of the explanatory variables[2]:

$$\begin{aligned} y_{ij} &= \beta_{0ij} + \beta_{1j} x_{1ij} + \beta_2 x_{2ij} + ... + \beta_k x_{kij} \\ \beta_{0ij} &= \beta_0 + u_{0j} + e_{0ij} \\ \beta_{1j} &= \beta_1 + u_{1j} \end{aligned} \qquad (6)$$

[2] In equation (6) note that the slope coefficient varies by institution only. The model, however, can be extended so that the slope coefficient varies by both institution and individual.

In model (6), as in model (5), the intercepts of the individual universities vary around the mean (of β_0) by the amount u_{0j} with variance σ_{u0}^2. In addition, the student's individual degree results vary around the universities' lines by e_{0ij} with variance σ_{e0}^2. In contrast to model (5) however, the coefficient on x_1 is not a constant. In fact the mean slope coefficient (across all universities) on x_1 is β_1, and the slopes for the individual universities vary around this by the amount u_{0j} with a variance of σ_{u1}^2. The relationship between the level 2 slope and intercept residuals is measured by an additional parameter of the model, namely the level 2 intercept/slope covariance denoted by σ_{u01}. If this parameter is positive, for example, this indicates that universities with higher intercepts also tend to have steeper slopes i.e. the estimated lines for each university fan out to the right.

The disadvantage of MLM in the context of performance measurement, however, is that observations can lie both above and below the line of best fit, contrary to the theory of production[3]. An alternative approach to MLM which constructs a frontier rather than an average line of best fit exists in the form of DEA, where the DMU, rather than being a school, university, department or district is, in fact, the individual pupil or student (Thanassoulis 1999; Portela & Thanassoulis 2001; Thanassoulis & Portela 2002). For example, consider a student at an IHE whose output is his degree result and his input his entry qualification. Each student's efficiency score is then obtained from applying DEA to all students in the higher education sector, but this efficiency score would incorporate a component which was a consequence of the student's own efforts and a component which was a consequence of the efficiency of teaching at the university attended by the student. In order to assess the efficiency of the IHEs, it would therefore be necessary, as a first step, to decompose the students' efficiency scores into these two components using a method pioneered by Portela & Thanassoulis (2001) and Thanassoulis & Portela (2002). Consider a hypothetical data set of students from two universities, each producing graduates with degrees, the quality of which is measured by degree results, using initial student quality, measured

[3] It should be noted that Thanassoulis *et al* (2003) adapt MLM to derive various frontier measures of efficiency in order to compare them to DEA measures of efficiency which are described in the remainder of the section.

by entry qualification. The output and input data can be plotted for all students (see Figure 1).

The boundary EFCD envelops all students and can be termed the student-within-all-universities efficiency boundary, students lying on segments EF and CD being on the boundary but not efficient (because of slacks). Thus, using the traditional (output-oriented) DEA definition of efficiency, student F, who lies on the student-within-all-universities efficiency frontier, has an efficiency score of 1, whereas student Y, who lies inside the student-within-all-universities efficiency frontier, has an overall efficiency level of OY/OY" which is less than 1. In other words, OY/OY" represents the proportion of degree achievement obtained by student Y relative to the best achievement obtained by students from all universities, *and* given student Y's initial qualifications.

This student-within-all-universities efficiency score, however, conceals the effect that the university has on the student's level of achievement. Students from university T, for example, have their own efficiency boundary (termed the student-within-own-university efficiency boundary), defined by ABCD. Similarly, the student-within-own-university efficiency boundary for university S is EFGH. Thus student Y (from university T) has a student-within-own-university efficiency score of OY/OY', which represents the proportion of degree achievement obtained by student Y relative to the best achievement obtained by students from university T only *and* given student Y's initial qualifications. The distance Y'Y" gives a measure of the impact of student Y's university on his degree result. The university-within-universities efficiency score, specific to student Y, is defined as the ratio OY'/OY", and varies with the level of input.

The efficiency of each IHE can then be examined by comparing the array of university-within-universities efficiency scores of each IHE's own students. This component is a measure of the efficiency of the IHE itself, and is not contaminated by the effects of students' own efforts. However, a comparison of all three components can offer useful insights, particularly to decision-makers within each institution, into how greater efficiency can be achieved (Portela & Thanassoulis 2001; Thanassoulis & Portela 2002).

3. DATA AND METHODOLOGY

Both MLM and DEA will be applied to the data relating undergraduate teaching output to undergraduate teaching inputs in order to rank the universities in terms of efficiency. The advantage of performing the statistical analysis first is that the variables found to be significant determinants of undergraduate teaching output can be used as inputs in the DEA, thereby overcoming the problem of specification usually encountered in a DEA[4].

The analysis requires a full data set of the performance and personal characteristics of individuals leaving their IHE in a given year. Such a data set, compiled by the Universities Statistical Record (USR)[5] of more than 117000 students (from pre-1992 universities) leaving university in 1993 fulfils the criteria required and therefore forms the basis of the analysis[6].

The output of the undergraduate teaching process can be measured in a variety of ways. Measures based on the degree results achieved by graduates, the salary obtained by graduates in the labour market or the propensity of students to leave university without a degree are commonly used to reflect undergraduate teaching output (Johnes & Taylor, 1990; Johnes, 1996; Smith and Naylor 2001b). The approach taken in the present paper attempts to capture the success or failure of students in one composite measure by employing weights to reflect degree classification or lack of degree. The relative weights of degree classifications are represented by an index derived by Mallier and Rodgers (1995) based on income differentials by degree classification. The weight for students who do not achieve a degree is derived from the salary differential between those who complete and those who do not (Johnes & Taylor

[4] Increasing the number of inputs and/or outputs in a DEA can increase the proportion of efficient DMUs in the data set and lead to a higher overall average efficiency (Chalos 1997).

[5] The data set was made available by the USR and the UK Data Archive.

[6] Students who were classed as aegrotat or enhanced first degree, or left university for non-academic reasons have been deleted. In addition, students from Scottish universities or whose main entry qualification was Scottish Certificate of Education have been deleted in order to avoid problems which may arise from the inclusion of individuals who are from a system of education which differs from that in the rest of the UK. Students whose A level score is unknown are also deleted, as are students of medicine, dentistry or veterinary studies where the length of degree differs from the degree length of the students remaining in the sample.

1991). The weighting system is described in table 1[7], which also shows the distribution of degree results for the data set.

There are numerous possible inputs which could affect the quality and quantity of undergraduate teaching output. The most obvious is the quality of the student on arrival at university, and there is strong evidence of a positive relationship between previous academic achievement and degree results (Freeman 1970; Kapur 1972; Tarsh 1982; Crum & Parikh 1983; Sear 1983; Rudd 1984; Montague & Odds 1990; Johnes 1992; Chapman 1994; Rodgers & Ghosh 2001; Smith & Naylor 2001a; Bratti 2002). The possibility that the effect of entry qualification varies by subject of degree (Entwistle & Wilson 1977; Sear 1983) suggests that the analysis should be performed separately for different subject groups (see for example Smith 1990; Jenkins & Smith 1993; Bratti 2002). Table 2 shows the distribution of degree results across four broad subject categories (pure science, applied science, social science and arts) along with the mean entry score (ASCORE) for each degree class in each subject[8]. As a consequence of the observed differences in distribution of degrees across subjects and the difference between subjects in the relationship between mean entry score and degree classification the analysis in the next section is performed separately for each subject group.

It is generally accepted that personal characteristics of the students themselves also affect their outcome in the undergraduate teaching process. Such a rich dataset allows the construction of numerous variables to reflect the personal characteristics of the students. The gender, age, marital status, country of origin, and socio-economic status of a student may all affect the level of their achievement at university. Indeed, there is clear evidence that females achieve better degree results than males (Rudd 1984; Rodgers & Ghosh 2001; Smith & Naylor 2001a); and that married students and students who are not from abroad perform better in their degree than, respectively,

---

7 Alternative weighting systems are also used but the results reported in the next section are not sensitive to the weighting system employed.

8 Subject groups are: arts comprising subject codes 71-96 and 104-106; social studies comprising subject codes 53-70 and 102-103; pure science comprising subject codes 5-18, 24-33, 99 and 101; applied science comprising subject codes 20-23 and 34-52, where the subject codes are as defined in Appendix II of *University Statistics volume 2 1992-1993* (Universities Statistical Record. Note that students of medicine, dentistry and veterinary studies are excluded, as are students studying across disciplines.

unmarried students and those from abroad (Smith & Naylor 2001a). The evidence regarding the effect of age and socio-economic status, however, is mixed.

Additional factors which may also affect performance at university include the type of degree (i.e. part-time or full-time) and the type of living accommodation whilst at university. Variables to represent these aspects are also included in the analysis.

A full list and description of all the variables included in the analysis and the evidence of their effect on performance from previous studies is provided in table 3.

4. RESULTS

A ML model[9] of the form of equation (3) with no explanatory variables was estimated to form the base against which subsequent models could be compared. Models of the form of equation (5), where the explanatory variables are defined in table 3, were estimated for each of the four subject groups defined. The results presented in table 4a provide three sets of results for each subject: the model with no explanatory variables; the model with only pre-university entry qualification (ASCORE) as the explanatory variable; and a final model in which all explanatory variables included are significant[10].

When no explanatory variables are included in the model, the amount by which the variation in degree results is due to differences between universities varies across the subjects as indicated by the intra-university correlation. Thus, 2.61% of the variation is due to differences between universities in arts, 2.81% in pure science, and this rises to 3.51% and 4.50% in social science and pure science (respectively).

Although entry qualification is a significant explanatory variable in the model, there are, however, also big differences across subjects in the effect of initial entry qualification (ASCORE) on academic achievement (DEGVALUE). While the

[9] All ML models were estimated using the package MLwiN.

[10] The change in the value of –2*loglikelihood determined whether or not to retain a variable: a variable which caused this statistic to fall by more than $\chi^2_{0.05,1}$ = 3.99 was retained in the model; otherwise it was excluded.

variable ASCORE is a significant explanatory variable in all subjects, the percentage of the unexplained variation which is explained by including ASCORE in the model is 10.62% and 11.79% for pure science and applied science (respectively), but is much lower at around 5 to 6% for social science and arts. This result that the association between entry score and degree results is strongest amongst science graduates and weakest amongst the arts and social science graduates confirms earlier findings (Entwistle & Wilson 1977; Sear 1983).

The inclusion of additional background characteristics of the students in the model reveal further differences between subjects. First, while student gender (measured by FEMALE) is a consistently significant explanatory variable across all subjects, the effect is positive in pure, applied and social science subjects, but negative in the arts.

Second, there are some variables which appear to be significant in one subject but not the others. For example, being married (MARITAL) is significantly positive in the arts; students' age (AGE) is significantly positive in pure science, in contrast to earlier findings (Walker 1975); and not being on a part-time course has a significantly positive effect in social science. Living in halls of residence (HALLS) has a significantly positive effect on degree results in both pure science and social science.

There are, however, some consistencies in the determinants of academic achievement across subjects. The nationality of students (UK) has a significantly positive effect, while having attended an independent school has a significantly negative effect in all subjects, and these confirm earlier findings (Smith & Naylor 2001a). Living at home whilst at university has a significant positive effect on achievement in three subjects, as does attendance at a comprehensive school prior to coming to university.

It is clear that the effect of background characteristics in explaining the initial unexplained variation in achievement is much lower than for entry qualification, and this is true across all subjects. Thus, entry qualification is the single most important explanatory variable in explaining academic achievement for all 4 subject categories considered.

Allowing the coefficient on ASCORE to vary at university level (see results in table 4b) causes a reduction in the value of –2*loglikelihood of more than 3.99 (i.e. $\chi^2_{0.05,1}$) in each of the subject areas. The greatest effect, however, is observed in the arts. The magnitude and significance of the coefficient on the explanatory variables are largely unchanged by allowing the slope on ASCORE to vary by university.

An additional result of the models in table 4b which is worthy of note is that the university level intercept/slope covariance is negative for all subjects apart from social science. Thus, in pure science, applied science and arts, a steeper slope on ASCORE is related to a higher intercept, and the opposite is the case for social science. The effect, however, is only significant in pure science.

Various DEA[11] models designed to correspond as closely as possible to the MLM results of tables 4a and 4b were performed for each subject group, and these are described in table 5. A presentation of the full sets of efficiencies for each subject group is unnecessary since it is the correspondence between the universities' efficiencies derived from DEA and MLM which is the focus of the analysis. Thus the DEA results are summarised in table 6. Finally, the correlation between the efficiency rankings derived for each university[12] are displayed in table 7. The main findings are discussed below.

First, the rankings of the universities based on the university effects of a ML model with no explanatory variables (defined in equation (3)) are most highly correlated with rankings derived from the student-within-all-universities efficiencies. This is true for all subjects.

Second, it is generally the case that as more variables are added to the ML model, the university rankings derived from the university effects are also highly and

[11] The DEA models were estimated using a Fortran programme provided by Geraint Johnes, Lancaster University.

[12] The rankings for each university based on student-within-own-university efficiencies are derived as follows. For each university, an unweighted mean of the individuals' student-within-own-university efficiencies is calculated to provide a mean efficiency score. Universities are then ranked according to this mean efficiency score. A similar process is applied to student-within-all-universities efficiencies and university-within-universities efficiencies to obtain a mean efficiency score and hence ranking for each university based on each of these variables respectively.

significantly correlated with university rankings derived from the student-within-all-universities efficiencies. Pure science is an exception to this generalisation.

The rankings of universities derived from the university-within-universities efficiencies are significantly correlated with the rankings derived from the university effects *only* from the simplest MLM (i.e. the one with no explanatory variables), with the exception of arts, where no such significant correlations are observed. Moreover, even in the subjects where these correlations are significant, the magnitude of the correlation is not particularly high (the highest rank correlation coefficient is 0.537 in pure science). Thus the individual level DEA measure of efficiency based on the university-within-universities efficiencies (i.e. the measure from which the effects of the efforts of individuals have been excluded) provides rankings of universities which are *unrelated* to the rankings derived from the university effects of ML models which include explanatory variables, and are most closely related to the university effects derived from *ML models with no explanatory variables*. The rankings of universities derived from the university effects from the ML models (with various specifications) are, however, highly and significantly correlated with the rankings derived from the student-within-all-universities efficiencies (i.e. the measure which includes the effects of students' own efforts on their academic achievement).

In order to make further comparisons, university-level DEAs were performed to compare the results with those derived from the individual-level DEAs and the MLM analysis (see table 8 for definitions of the university-level DEA runs). The correlations of university rankings derived from university-level DEAs are provided in columns/rows 10 and 11 of table 7. It should be noted that a university-level DEA which includes a large number of inputs produces a large number of universities with an efficiency score of 1, and hence there is little discrimination between universities in terms of efficiency. Thus the correlations will be discussed in the context only of the simple university-level DEA (column/row 10).

First, the rankings of universities based on the university-level DEA are significantly related to the rankings derived from the student-within-all-universities efficiencies, and are *not* significantly correlated with the rankings derived from the university-within-universities efficiencies. This is true across all subjects. This confirms the result also found in Johnes (2003) based on a small sample of economics graduates.

Second, it is particularly interesting that there is a strong and highly significant correlation between the rankings derived from the university-level DEA and the rankings from all specifications of the ML models. This is observed across all subjects.

Finally, we can turn attention to just how reliable are the rankings of universities derived from various methods. It is possible with MLM to calculate for each university not only a university effect but also an associated 95% confidence interval. From this a caterpillar plot (Goldstein & Spiegelhalter 1996) of universities' effects and associated confidence intervals can be produced in order to assess whether there are significant differences between universities in terms of their performance. These plots are shown in figures 2a-2d for each of the subject groups using the models in table 4a with only one explanatory variable (ASCORE) and with constant slopes across individuals and universities.

There appears to be some difference between subjects in the degree to which the confidence intervals overlap. In pure science, for example, the 11 bottom-ranked universities perform significantly worse than the median (represented by the horizontal line) and the 10 top-ranked universities perform significantly better than the median. No distinction can be made, however, between the middle 27 universities. In the remaining three subjects, there is considerably more overlap, and 34 universities in arts, 35 universities in applied science, and 36 universities in social science cannot be separated.

It is also possible to calculate 95% confidence intervals for the university-level DEA efficiency scores using bootstrapping procedures (Simar and Wilson 1998, 1999)[11]. Caterpillar plots illustrating these results are also displayed (in figures 3a-3d) and are more indicative of overlaps between universities in terms of their efficiency. Indeed, no significant differences can be found, in terms of performance, between universities in any of the subject categories.

5. CONCLUSION

The purpose of this paper has been to compare the results from DEA applied to a data set of individuals with those from MLM applied to the same data set, in the context of student achievement at university. The MLM results suggest that there are some differences between subjects in the determinants of achievement at university, although entry qualification is the single most important explanatory variable in all subjects. The results of the MLM are used to specify the inputs in the subsequent individual-level DEA.

An individual-level DEA provides three sets of efficiencies: student-within-all-universities efficiencies; student-within-own-university efficiencies; and university-within-universities efficiencies. The last quantity purportedly measures the efficiency of the university having taken out the effects of the efforts of the individuals, while the first quantity includes both the university's and individuals' efforts. It is of interest, therefore, that it is the student-within-all-universities efficiencies which provide university rankings most closely correlated with the rankings from MLM university effects. This is true across all subjects. The university-within-universities efficiencies provide university rankings which are generally not significantly correlated with the rankings from multilevel models which include explanatory variables.

When the individual data are adapted to perform university-level DEAs, the results indicate that the university-level DEA efficiency scores are more strongly correlated with the efficiency scores calculated from the student-within-all-universities efficiencies than those from the university-within-universities measures, and this confirms an earlier result from a more limited data set (Johnes 2003).

Remarkably, however, the university-level DEA efficiency scores provide university rankings which are highly correlated with the rankings from the university effects of multilevel models.

Finally, a closer inspection of the efficiency scores and their associated confidence intervals from the multilevel models suggest that it is impossible to separate many of the middle-performing universities in terms of their efficiency. This is confirmed by an examination of the university-level DEA scores and confidence intervals (derived using bootstrapping methods).

This paper has investigated whether the choice of analytical technique is important when measuring performance. The results suggest that the DEA and MLM performed on individual data provide different measures of efficiency. It appears that the level of analysis (individual or university) is also important. Closer inspection of the efficiency scores and confidence intervals from MLM indicate that these results are particularly pertinent in the measurement of performance at the extremes since middle-ranking universities cannot be separated in terms of their performance.

Table 1: Distribution of degree classification for sample data

Degree classification	Weight	number	%
no degree	1.90	1529	2.6
pass	2.00	1004	1.7
third class honours	2.20	3377	5.7
lower second class honours	2.30	19200	32.2
upper second class honours	2.45	28766	48.3
first class honours	2.85	5666	9.5

Table 2: Distribution of degrees and mean entry score by broad subject of degree

	%	mean ASCORE	%	mean ASCORE	%	mean ASCORE	%	mean ASCORE	%	mean ASCORE
Degree classification	ALL USED OBS		PURE SCIENCE		APPLIED SCIENCE		SOCIAL SCIENCE		ARTS	
no degree	2.6	17.60	3.2	17.01	4.7	15.95	1.9	19.76	1.4	19.29
pass	1.7	18.19	2.4	17.87	4.0	17.22	0.8	19.87	0.6	20.49
third class honours	5.8	18.41	9.2	18.76	10.2	17.18	2.7	19.54	2.5	18.19
lower second class honours	32.3	20.02	31.5	19.85	34.2	18.52	34.1	21.12	30.9	19.93
upper second class honours	47.8	22.49	40.7	21.99	34.9	21.19	54.7	23.21	56.5	22.67
first class honours	9.7	25.20	13.1	25.39	12.1	24.77	5.8	25.57	8.1	24.90
Total	54578	21.52	18395	21.21	7317	19.90	13779	22.44	15087	21.83

Table 3: Definition of the variables constructed to represent the inputs of the undergraduate teaching process

Input variable	Definition	Evidence from previous studies of the effect on performance		
		Positive	Negative	No significant effect
ASCORE	Score based on best 3 A levels or equivalent (i.e. 2 AS levels = 1 A level) For A levels: A = 10; B = 8; C = 6; D = 4; E = 2. For AS levels: A = 5; B = 4; C = 3; D = 2; E = 1. Note that duplicate subjects are not counted.	Freeman 1970; Kapur 1972; Tarsh 1982; Crum & Parikh 1983; Sear 1983; Rudd 1984; Montague & Odds 1990; Johnes 1992; Chapman 1994; Rodgers & Ghosh 2001; Smith & Naylor 2001a; Bratti 2002		Bee & Dolton 1985 Connolly & Smith 1986
FEMALE	1 = female, 0 = male	Rudd 1984 Rodgers & Ghosh 2001 Smith & Naylor 2001a		Bee & Dolton 1985
MARITAL	1 = married, 0 = single	Smith & Naylor 2001a		
UK	1 = from UK, 0 = otherwise	Smith & Naylor 2001a		Johnes & Taylor 1987
NOTPT	1 = on a part-time course; 0 = not on a part-time course	Smith & Naylor 2001a		
HOME	1 = live at home; 0 = does not live at home	Smith & Naylor 2001a	Johnes & Taylor 1987	
HALLS	1=live in halls of residence; 0 = does not live in halls of residence	Johnes & Taylor 1989b[1]		
AGE	age in 1993	Walker 1975 Eaton & West 1980 Smith & Naylor 2001a	Barnett & Lewis 1963 Barnett *et al* 1968 Kapur 1972 Entwistle & Wilson 1977 Bee & Dolton 1985	Nisbett & Welsh 1972 Smith 1990
IND	1 = attended an independent school prior to entering university; 0 = otherwise	Barnett & Lewis 1963[2]	Smith & Naylor 2001a	Rodgers & Ghosh 2001
COMP	1=attended a comprehensive school prior to entering university; 0 otherwise			Johnes 1997[1]

Notes:

1. Student performance is reflected by completion rather than degree attainment in the case of these two studies.
2. 2. In fact the variable used in this study reflected attendance at an independent or grammar school.

Table 4a: MLM results – intercept variable at level 1 and level 2; all slope coefficients constant across levels

	PURE SCIENCE (n = 18395)						APPLIED SCIENCE (n = 7303)					
Variable	Est	SE	Est	SE	Est	SE	Est	SE	Est	SE	Est	SE
constant	-0.035	0.026	0.013	0.018	-0.238	0.043	-0.036	0.036	0.002	0.023	-2.24	0.162
ASCORE[1]			0.355	0.008	0.383	0.008			0.344	0.012	0.376	0.012
FEMALE					0.192	0.014					0.183	0.028
MARITAL												
UK					0.114	0.039					0.089	0.007
NOTPT												
HOME					0.092	0.029					0.129	0.042
HALLS					0.090	0.018						
AGE					0.087	0.007						
IND					-0.043	0.020					-0.094	0.030
COMP					0.060	0.016					0.074	0.027
-2loglikelihood	51744.95		49937.17		49526.13		20476.91		19752.77		19499.88	
university σ^2 intercept (σ^2_{u0})	0.028	0.006	0.012	0.003	0.011	0.003	0.045	0.012	0.014	0.005	0.012	0.004
student σ^2 intercept (σ^2_{e0})	0.970	0.010	0.880	0.009	0.861	0.009	0.956	0.016	0.869	0.014	0.840	0.014
intra-university correlation	0.0281		0.0135		0.0126		0.0450		0.0159		0.0141	
% unexplained variation accounted for			10.62%		12.63%				11.79%		14.89%	

Table 4a continued

	SOCIAL SCIENCE (n = 13779)						ARTS (n = 15087)					
Variable	Est	SE	Est	SE	Est	SE	Est	SE	Est	SE	Est	SE
constant	-0.027	0.029	-0.004	0.022	-1.114	0.080	-0.030	0.026	-0.005	0.017	-0.784	0.062
ASCORE[1]			0.210	0.009	0.261	0.010			0.236	0.009	0.298	0.009
FEMALE					0.162	0.016					-0.071	0.016
MARITAL											0.165	0.061
UK					0.035	0.003					0.035	0.003
NOTPT					0.141	0.036						
HOME					0.289	0.032						
HALLS					0.060	0.023						
AGE												
IND					-0.142	0.019					-0.053	0.021
COMP											0.062	0.020
-2loglikelihood	38696.68		38194.480		37757.89		42412.77		41749.54		41329.4	
university σ^2 intercept (σ^2_{u0})	0.035	0.008	0.018	0.004	0.018	0.005	0.026	0.006	0.009	0.003	0.008	0.002
student σ^2 intercept (σ^2_{e0})	0.963	0.012	0.930	0.011	0.901	0.011	0.968	0.011	0.928	0.011	0.903	0.010
intra-university correlation	0.0351		0.0190		0.0196		0.0261		0.0096		0.0088	
% unexplained variation accounted for			5.01%		7.92%				5.73%		8.35%	

Table 4b: MLM results – intercept variable at level 1 and level 2; slope coefficient on ASCORE variable at level 2

	PURE SCIENCE (n = 18395)				APPLIED SCIENCE (n = 7303)			
Variable	Est	SE	Est	SE	Est	SE	Est	SE
constant	0.003	0.017	-0.247	0.043	-0.010	0.024	-2.157	0.162
ASCORE[1]	0.352	0.011	0.382	0.011	0.351	0.017	0.385	0.017
FEMALE			0.192	0.014			0.183	0.028
MARITAL								
UK			0.115	0.039			0.089	0.007
NOTPT								
HOME			0.092	0.030			0.129	0.042
HALLS			0.087	0.018				
AGE			0.086	0.007				
IND			-0.043	0.020			-0.096	0.030
COMP			0.059	0.016			0.077	0.027
-2loglikelihood	49919.23		49512.66		19744.21		19489.82	
university σ^2								
intercept (σ^2_{u0})	0.011	0.003	0.010	0.003	0.015	0.005	0.013	0.004
ASCORE/intercept (σ_{u01})	-0.003	0.001	-0.002	0.001	-0.002	0.002	-0.001	0.002
ASCORE (σ^2_{u1})	0.003	0.001	0.002	0.001	0.005	0.002	0.005	0.002
student σ^2								
intercept (σ^2_{e0})	0.878	0.009	0.859	0.009	0.866	0.014	0.836	0.014
intra-university correlation (calculated for mean ASCORE)	0.0124		0.0115		0.0170		0.0153	
% unexplained variation accounted for	10.92%		12.93%		11.99%		15.18%	

Table 4b continued

	SOCIAL SCIENCE (n = 13779)				ARTS (n = 15087)			
Variable	Est	SE	Est	SE	Est	SE	Est	SE
constant	-0.011	0.018	-1.111	0.080	-0.020	0.016	-0.790	0.062
ASCORE	0.213	0.014	0.262	0.013	0.241	0.017	0.301	0.015
FEMALE			0.163	0.016			-0.069	0.016
MARITAL							0.147	0.062
UK			0.034	0.003			0.035	0.003
NOTPT			0.140	0.036				
HOME			0.291	0.032				
HALLS			0.053	0.023				
AGE								
IND			-0.142	0.019			-0.055	0.021
COMP							0.061	0.020
-2loglikelihood	38179.99		37747.85		41695.01		41295.36	
university σ^2 intercept (σ^2_{u0}) ASCORE/intercept (σ_{u01}) ASCORE (σ^2_{u1})	0.011 0.002 0.005	0/003 0.002 0.002	0.013 0.002 0.003	0.004 0.002 0.001	0.007 -0.001 0.008	0.002 0.002 0.003	0.007 0.000 0.005	0.002 0.001 0.002
student σ^2 intercept (σ^2_{e0})	0.929	0.011	0.900	0.011	0.922	0.011	0.899	0.010
intra-university correlation (calculated for mean ASCORE)	0.0117		0.0142		0.0075		0.007	
% unexplained variation accounted for	5.91%		8.52%		6.54%		8.85%	

Note: ASCORE and DEGVALE are standardised to have zero mean and unit variance in the MLM analysis.

Table 5: Specification of outputs and inputs in various DEA runs

Model	Outputs	Inputs
PURE SCIENCE STUDENTS ONLY		
Model 1	DEGVALUE	ASCORE
Model 2	DEGVALUE	ASCORE, FEMALE, UK, HOME, HALLS, AGE, NOTIND[1], COMP
APPLIED SCIENCE STUDENTS ONLY		
Model 3	DEGVALUE	ASCORE
Model 4	DEGVALUE	ASCORE, FEMALE, UK, HOME, NOTIND[1], COMP
SOCIAL SCIENCE STUDENTS ONLY		
Model 5	DEGVALUE	ASCORE
Model 6	DEGVALUE	ASCORE, FEMALE, UK, NOTPT, HOME, HALLS, NOTIND[1]
ARTS STUDENTS ONLY		
Model 7	DEGVALUE	ASCORE
Model 8	DEGVALUE	ASCORE, MALE[2], MARITAL, UK, NOTIND[1], COMP

Notes:

1. NOTIND takes the value 1 for students who did not attend an independent school prior to university, and the value 0 otherwise.
2. MALE takes the value 1 for males and 0 for females.

Table 6: Summary of efficiency measures across all students

Model 1[1]	Min	Max	Mean	Median
Student-within-own-university efficiency	.6667	1	.8435	.8596
Student-within-all-universities efficiency	.6667	1	.8461	.8596
University-within-universities efficiency	.6667	1	.9971	1
Model 2[1]	Min	Max	Mean	Median
Student-within-own-university efficiency	.6667	1	.8465	.8596
Student-within-all-universities efficiency	.6667	1	.8708	.8596
University-within-universities efficiency	.6667	1	.9739	1
Model 3[1]	Min	Max	Mean	Median
Student-within-own-university efficiency	.6667	1	.8342	.8070
Student-within-all-universities efficiency	.6667	1	.8412	.8550
University-within-universities efficiency	.6667	1	.9924	1
Model 4[1]	Min	Max	Mean	Median
Student-within-own-university efficiency	.6667	1	.8345	.8070
Student-within-all-universities efficiency	.6667	1	.8552	.8596
University-within-universities efficiency	.6667	1	.9775	1
Model 5[1]	Min	Max	Mean	Median
Student-within-own-university efficiency	.6667	1	.8425	.8596
Student-within-all-universities efficiency	.6667	1	.8477	.8596
University-within-universities efficiency	.7018	1	.9943	1
Model 6[1]	Min	Max	Mean	Median
Student-within-own-university efficiency	.6667	1	.8440	.8596
Student-within-all-universities efficiency	.6667	1	.8684	.8596
University-within-universities efficiency	.6667	1	.9742	1
Model 7[1]	Min	Max	Mean	Median
Student-within-own-university efficiency	.6667	1	.8488	.8596
Student-within-all-universities efficiency	.6667	1	.8516	.8596
University-within-universities efficiency	.6667	1	.9970	1
Model 8[1]	Min	Max	Mean	Median
Student-within-own-university efficiency	.6667	1	.8492	.8596
Student-within-all-universities efficiency	.6667	1	.8638	.8596
University-within-universities efficiency	.6667	1	.9844	1

Note:

1. See table 5 for definition of models.

Table 7: Rank correlations

a) Pure science

		(1) Student-in-own-uni 1	(2) Student-in-all-unis 1	(3) University eff 1	(4) Student-in-own-uni 2	(5) Student-in-all-unis 2	(6) University eff 2	(7) MLWIN 0	(8) MLWIN 1	(9) MLWIN 2	(10) DEA 1	(11) DEA 2
(1)	Student-in-own-uni 1	1.000										
(2)	Student-in-all-unis 1	0.813**	1.000									
(3)	University efficiency 1	-0.055	0.363**	1.000								
(4)	Student-in-own-uni 2	0.280	-0.035	-0.536**	1.000							
(5)	Student-in-all-unis 2	0.754**	0.964	0.397**	-0.007	1.000						
(6)	University efficiency 2	0.212	0.528**	0.628**	-0.762**	0.521**	1.000					
(7)	MLWIN 0	0.823**	0.998**	0.359*	-0.034	0.960**	0.537**	1.000				
(8)	MLWIN 1	0.238	0.101	-0.153	0.232	0.142	-0.107	0.093	1.000			
(9)	MLWIN 2	0.223	0.134	-0.032	0.104	0.166	0.012	0.130	0.956**	1.000		
(10)	DEA 1	0.687**	0.649**	0.186	0.121	0.656**	0.263	0.641**	0.645**	0.671**	1.000	
(11)	DEA 2	0.081	0.019	0.043	-0.045	0.025	0.081	0.032	0.272	0.397**	0.265	1.000

Notes:
Student-in-own-uni 1 (2) refers to the student-within-own-university efficiency derived from model 1 (2) of table 5.
Student-in-all-unis 1 (2) refers to the student-within-all-universities efficiency derived from model 1 (2) of table 5.
University eff 1 (2) refers to the university-within-universities efficiency derived from model 1 (2) of table 5.
MLWIN 0 refers to the university effects from the ML model with no explanatory variables (see table 4a).
MLWIN 1 refers to the university effects from the ML model with one explanatory variable namely ASCORE (see table 4a).
MLWIN 2 refers to the university effects from the ML model with 8 explanatory variables (see table 4a).
DEA 1 (2) refers to the university efficiency scores derived from model 1 (2) in table 8.
An explanation of how the rankings are calculated is given in footnote 12 in the text.
* significant at the 5% significance level (2-tailed test); ** significant at the 1% significance level (2-tailed test)

b) Applied science

		(1) Student-in-own-uni 1	(2) Student-in-all-unis 1	(3) University eff 1	(4) Student-in-own-uni 2	(5) Student-in-all-unis 2	(6) University eff 2	(7) MLWIN 0	(8) MLWIN 1	(9) MLWIN 2	(10) DEA 1	(11) DEA 2
(1)	Student-in-own-uni 1	1.000										
(2)	Student-in-all-unis 1	0.505**	1.000									
(3)	University efficiency 1	-0.377*	0.371*	1.000								
(4)	Student-in-own-uni 2	0.663**	0.134	-0.531**	1.000							
(5)	Student-in-all-unis 2	0.500**	1.000**	0.379*	0.125	1.000						
(6)	University efficiency 2	-0.154	0.487**	0.695**	-0.688**	0.496**	1.000					
(7)	MLWIN 0	0.497**	0.990**	0.391*	0.144	0.989**	0.489**	1.000				
(8)	MLWIN 1	0.265	0.466**	0.264	0.268	0.465**	0.153	0.464**	1.000			
(9)	MLWIN 2	0.288	0.517**	0.293	0.255	0.513**	0.178	0.522**	0.976**	1.000		
(10)	DEA 1	0.372*	0.539**	0.159	0.355*	0.537**	0.108	0.528**	0.910**	0.893**	1.000	
(11)	DEA 2	0.595**	0.253	-0.295	0.629**	0.249	-0.350*	0.242	0.425**	0.466**	0.480**	1.000

Notes:
Student-in-own-uni 1 (2) refers to the student-within-own-university efficiency derived from model 3 (4) of table 5.
Student-in-all-unis 1 (2) refers to the student-within-all-universities efficiency derived from model 3 (4) of table 5.
University eff 1 (2) refers to the university-within-universities efficiency derived from model 3 (4) of table 5.
MLWIN 0 refers to the university effects from the ML model with no explanatory variables (see table 4a).
MLWIN 1 refers to the university effects from the ML model with one explanatory variable namely ASCORE (see table 4a).
MLWIN 2 refers to the university effects from the ML model with 6 explanatory variables (see table 4a).
DEA 1 (2) refers to the university efficiency scores derived from model 3 (4) in table 8.
An explanation of how the rankings are calculated is given in footnote 12 in the text.
* significant at the 5% significance level (2-tailed test); ** significant at the 1% significance level (2-tailed test)

c) Social science

		(1) Student-in-own-uni 1	(2) Student-in-all-unis 1	(3) University eff 1	(4) Student-in-own-uni 2	(5) Student-in-all-unis 2	(6) University eff 2	(7) MLWIN 0	(8) MLWIN 1	(9) MLWIN 2	(10) DEA 1	(11) DEA 2
(1)	Student-in-own-uni 1	1.000										
(2)	Student-in-all-unis 1	0.592**	1.000									
(3)	University efficiency 1	-0.260	0.457**	1.000								
(4)	Student-in-own-uni 2	0.347	-0.149	-0.338*	1.000							
(5)	Student-in-all-unis 2	0.658**	0.956**	0.357*	-0.017	1.000						
(6)	University efficiency 2	0.018	0.542**	0.542**	-0.822**	0.436**	1.000					
(7)	MLWIN 0	0.640**	0.993**	0.429**	-0.113	0.945**	0.515**	1.000				
(8)	MLWIN 1	0.684**	0.682**	0.137	0.210	0.735**	0.165	0.702**	1.000			
(9)	MLWIN 2	0.709**	0.686**	0.150	0.228	0.760**	0.160	0.710**	0.937**	1.000		
(10)	DEA 1	0.542	0.530**	0.097	0.225	0.604**	0.090	0.538**	0.955**	0.882**	1.000	
(11)	DEA 2	0.470	0.336*	-0.008	0.107	0.434**	0.104	0.353*	0.521**	0.652**	0.505**	1.000

Notes:
Student-in-own-uni 1 (2) refers to the student-within-own-university efficiency derived from model 5 (6) of table 5.
Student-in-all-unis 1 (2) refers to the student-within-all-universities efficiency derived from model 5 (6) of table 5.
University eff 1 (2) refers to the university-within-universities efficiency derived from model 5 (6) of table 5.
MLWIN 0 refers to the university effects from the ML model with no explanatory variables (see table 4a).
MLWIN 1 refers to the university effects from the ML model with one explanatory variable namely ASCORE (see table 4a).
MLWIN 2 refers to the university effects from the ML model with 7 explanatory variables (see table 4a).
DEA 1 (2) refers to the university efficiency scores derived from model 5 (6) in table 8.
An explanation of how the rankings are calculated is given in footnote 12 in the text.
* significant at the 5% significance level (2-tailed test); ** significant at the 1% significance level (2-tailed test)

d) Arts

		(1) Student-in-own-uni 1	(2) Student-in-all-unis 1	(3) University eff 1	(4) Student-in-own-uni 2	(5) Student-in-all-unis 2	(6) University eff 2	(7) MLWIN 0	(8) MLWIN 1	(9) MLWIN 2	(10) DEA 1	(11) DEA 2
(1)	Student-in-own-uni 1	1.000										
(2)	Student-in-all-unis 1	0.762**	1.000									
(3)	University efficiency 1	-0.364*	0.146	1.000								
(4)	Student-in-own-uni 2	0.622**	0.253	-0.582**	1.000							
(5)	Student-in-all-unis 2	0.753**	0.995**	0.145	0.242	1.000						
(6)	University efficiency 2	-0.170	0.246	0.662**	-0.748**	0.259	1.000					
(7)	MLWIN 0	0.766**	0.991**	0.177	0.243	0.986**	0.276	1.000				
(8)	MLWIN 1	0.613**	0.704**	0.018	0.297*	0.728**	0.112	0.724**	1.000			
(9)	MLWIN 2	0.588**	0.617**	-0.055	0.315*	0.647**	0.063	0.639**	0.967**	1.000		
(10)	DEA 1	0.403**	0.387**	-0.143	0.375**	0.414**	-0.101	0.406**	0.841**	0.885**	1.000	
(11)	DEA 2	0.275	0.198	-0.193	0.333*	0.219	-0.211	0.193	0.421**	0.496**	0.570**	1.000

Notes:
Student-in-own-uni 1 (2) refers to the student-within-own-university efficiency derived from model 7 (8) of table 5.
Student-in-all-unis 1 (2) refers to the student-within-all-universities efficiency derived from model 7 (8) of table 5.
University eff 1 (2) refers to the university-within-universities efficiency derived from model 7 (8) of table 5.
MLWIN 0 refers to the university effects from the ML model with no explanatory variables (see table 4a).
MLWIN 1 refers to the university effects from the ML model with one explanatory variable namely ASCORE (see table 4a).
MLWIN 2 refers to the university effects from the ML model with 6 explanatory variables (see table 4a).
DEA 1 (2) refers to the university efficiency scores derived from model 7 (8) in table 8.
An explanation of how the rankings are calculated is given in footnote 12 in the text.
* significant at the 5% significance level (2-tailed test); ** significant at the 1% significance level (2-tailed test)

Table 8: Specification of outputs and inputs in the university-level DEA runs

Model	Outputs	Inputs
PURE SCIENCE		
Model 1	MEANDEG	MEANASCORE
Model 2	MEANDEG	MEANASCORE, %FEMALE, %UK, %HOME, %HALLS, MEANAGE, %NOTIND, %COMP
APPLIED SCIENCE		
Model 3	MEANDEG	MEANASCORE
Model 4	MEANDEG	MEANASCORE, %FEMALE, %UK, %HOME, %NOTIND, %COMP
SOCIAL SCIENCE		
Model 5	MEANDEG	MEANASCORE
Model 6	MEANDEG	MEANASCORE, %FEMALE, %UK, %NOTPT, %HOME, %HALLS, %NOTIND
ARTS		
Model 7	MEANDEG	MEANASCORE
Model 8	MEANDEG	MEANASCORE, %MALE, %MARITAL, %UK, %NOTIND, %COMP

Definitions of variables:

Variable	Definition
MEANDEG	For given subject, unweighted mean of DEGVALUE.
MEANASCORE	For given subject, unweighted mean of ASCORE.
%FEMALE	Percentage of students in given subject who are female.
%MALE	Percentage of students in given subject who are male.
%MARITAL	Percentage of students in given subject who are married.
%UK	Percentage of students in given subject who are from the UK.
%NOTPT	Percentage of students in given subject who are not pert-time.
%HOME	Percentage of students in given subject who live at home.
%HALLS	Percentage of students in given subject who live in halls of residence.
MEANAGE	For given subject, unweighted mean age of students.
%NOTIND	Percentage of students in given subject who did not attend an independent school.
%COMP	Percentage of students in given subject who did not attend a comprehensive school.

Figure 1: Decomposing efficiencies using DEA

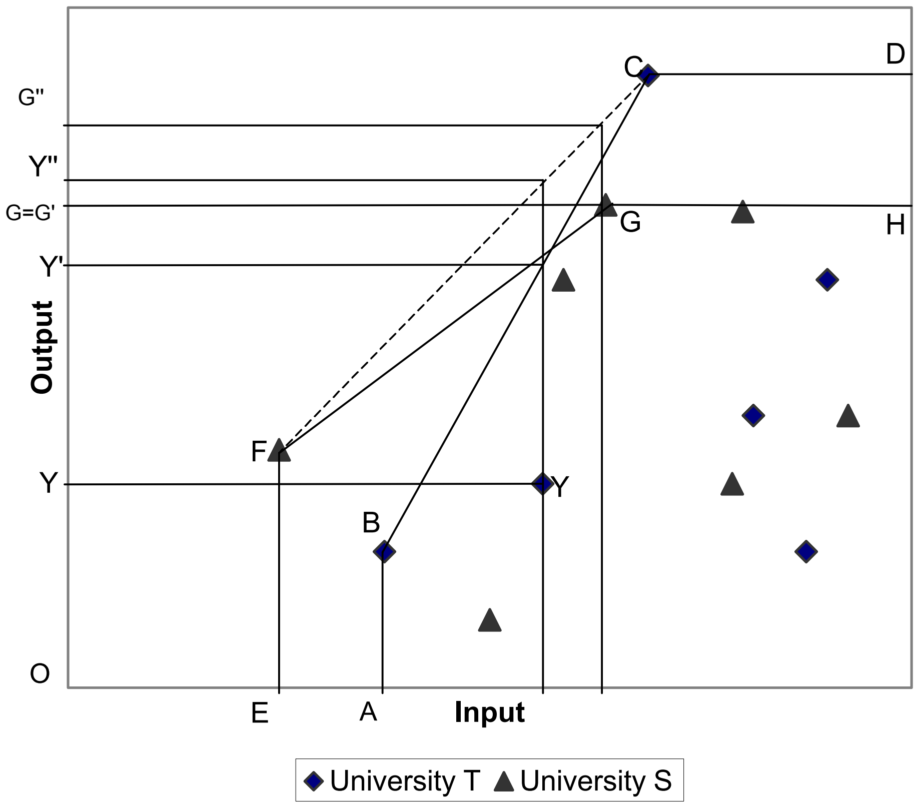


Figure 2: University effects and associated 95% confidence intervals derived using MLM

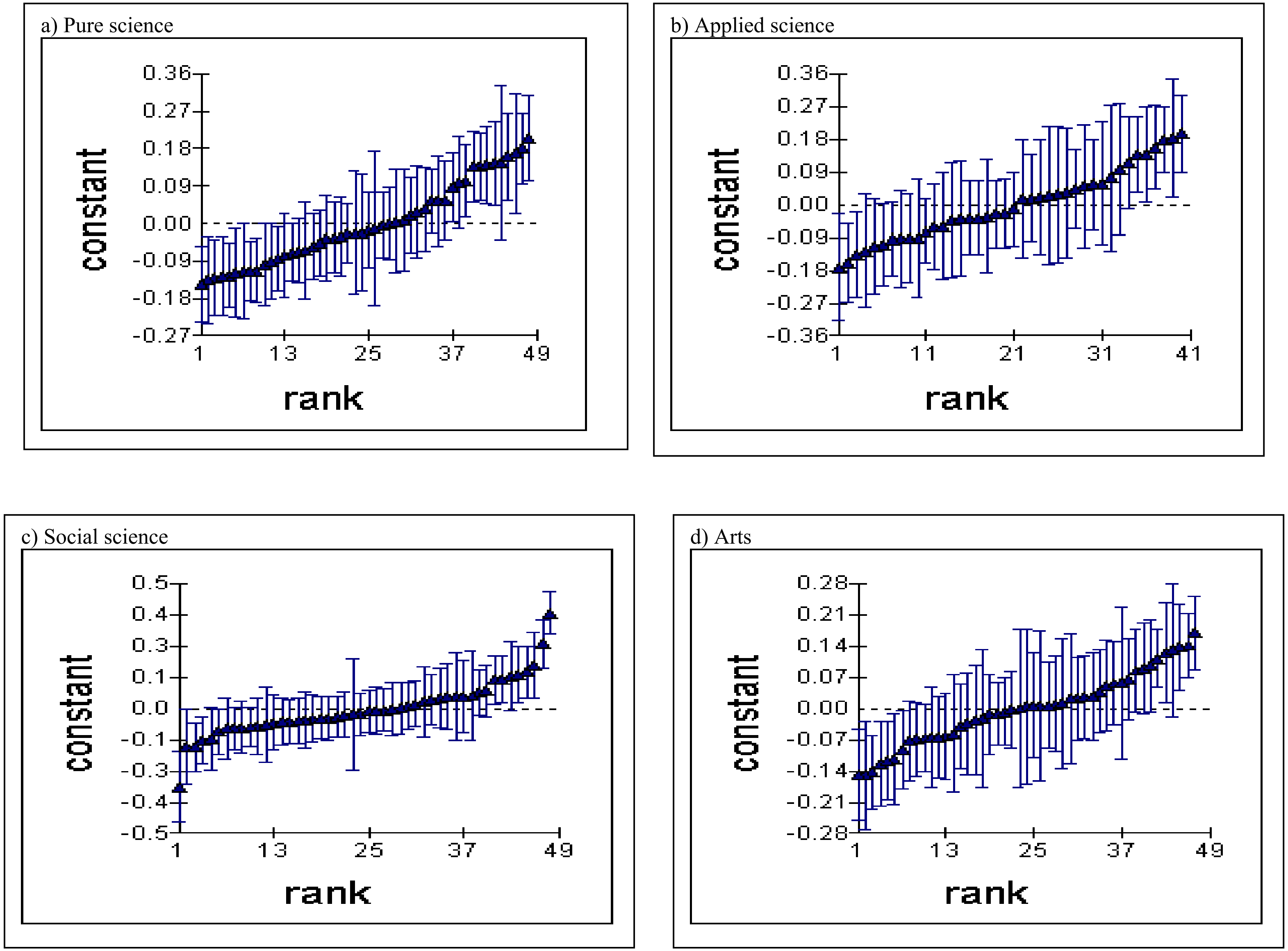
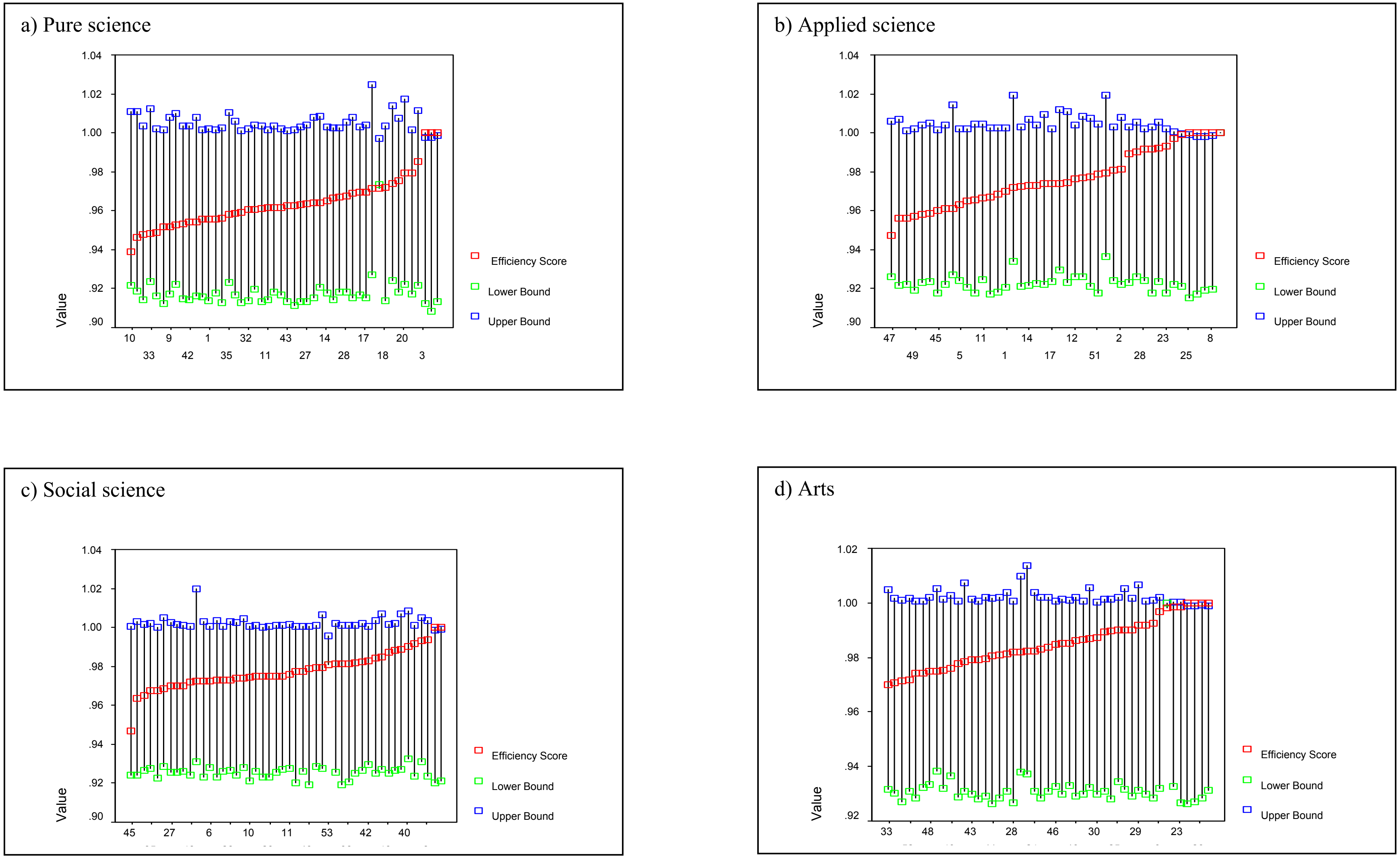


Figure 3: Efficiency scores and associated 95% confidence intervals derived using university-level DEA



References

Abbott, M & Doucouliagos, C (2002) 'The efficiency of Australian universities: a data envelopment analysis' *Economics of Education Review*, forthcoming

Ahn, T, Arnold, V, Charnes, A & Cooper, W W (1989) 'DEA and ratio efficiency analyses for public institutions of higher learning in Texas' *Research in Governmental and Nonprofit Accounting*, 5 pp165-185

Ahn, T & Seiford, L M (1993) 'Sensitivity of data envelopment analysis to models and variable sets in a hypothesis test setting: the efficiency of university operations' in Y Ijiri (ed*) Creative and Innovative Approaches to the Science of Management* Quorum Books, Westport, Connecticut, pp191-208

Aigner, D J, Lovell, C A K & Schmidt, P (1977) 'Formulation and estimation of stochastic frontier production function models' *Journal of Econometrics*, 6(1) pp21-37

Aitken, M & Longford, N (1986) 'Statistical modelling in school effectiveness studies' *Journal of the Royal Statistical Society, Series A*, 149 pp1-43

Athanassopoulos, A D & Shale, E (1997) 'Assessing the comparative efficiency of higher education institutions in the UK by means of data envelopment analysis' *Education Economics*, 5(2) pp117-134

Avkiran, N K (2001) 'Investigating technical and scale efficiencies of Australian universities through data envelopment analysis' *Socio-Economic Planning Sciences*, 35(1) pp57-80

Barnett, V D, Holder, R L & Lewis, T (1968) 'Some new results on the association between students' ages and their degree results' *Journal of the Royal Statistical Society, Series A*, 131 pp410-433

Barnett, V D & Lewis, T (1963) 'A study of the relation between GCE and degree results' *Journal of the Royal Statistical Society, Series A*, 126 pp187-216

Barrow, M (1991) 'Measuring the Local Education Authority performance: a frontier approach' *Economics of Education Review*, 10(1) pp19-27

Beasley, J E (1990) 'Comparing university departments', *Omega*, 18 pp171-183

Beasley, J E (1995) 'Determining teaching and research efficiencies' *Journal of the Operational Research Society*, 46(4) pp441-452

Bee, M & Dolton, P (1985) 'Degree class and pass rates: an inter-university comparison' *Higher Education Review*, 17 pp45-52

Bratti, M (2002) ‘Does the choice of university matter? A study of the differences across UK universities in life sciences students’ degree performance’ *Economics of Education Review*, 21 pp431-443

Bratti, M, McKnight, A, Naylor, R & Smith, J (2003) ‘Higher education outcomes, graduate employment and university performance indicators’ Department of Economics, University of Warwick

Chalos, P (1997) ‘An examination of budgetary inefficiency in education using data envelopment analysis’ *Financial Accountability and Management*, 13(1) pp55-69.

Chapman, K (1994) ‘Variability of degree results in geography in UK universities, 1973-1990: preliminary results and policy implications’ *Studies in Higher Education*, 19(1) pp89-102

Coelli, T, Rao, D S P & Battese, G E (1998) *An Introduction to Efficiency and Productivity Analysis*, Kluwer Academic, Norwell, MA

Connolly, K J & Smith, P (1986) ‘What makes a “good” degree: variations between different departments’ *Bulletin of the British Psychological Society*, 39 pp48-51

Crum, R & Parikh, A (1983) ‘Headmasters’ reports, admissions and academic performance in social sciences’ *Educational Studies*, 9 pp169-184

Deller, S C & Rudnicki, E (1993) ‘Production efficiency in elementary education: the case of Maine public schools’ *Economics of Education Review*, 12(1) pp45-57

Eaton, E G & West, H T (1980) ‘Academic performance of mature age students: recent research in Australia’ in T Hore and H T West (eds) *Mature Age Students in Australian Higher Education*, Acacia Press, Victoria

El-Mahgary, S & Lahdelma, R (1995) ‘Data envelopment analysis: visualizing the results’ *European Journal of Operational Research*, 85 pp700-710

Entwistle, N J & Wilson, J D (1977) *Degrees of Excellence: the Academic Achievement Game*, Hodder and Stoughton, London

Førsund, F R, Lovell, C A K & Schmidt, P (1980) ‘A survey of frontier production functions and of their relationship to efficiency measurement’ *Journal of Econometrics*, 13 pp5-25

Freeman, P R (1970) ‘A multivariate study of students’ performance in university examinations’ *Journal of the Royal Statistical Society, Series A*, 133 pp38-55

Goldstein, H & Spiegelhalter, D J (1996) 'League tables and their limitations: statistical issues in comparisons of institutional performance' *Journal of the Royal Statistical Society, Series A*, 159(3) pp385-443

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 pp63-71

Jenkins, A & Smith, P (1993) 'Expansion, efficiency and teaching quality: the experience of British geography departments 1986-1991' *Transactions, Institute of British Geographers New Series*, 18 pp500-515

Johnes, G (1998) 'The costs of multiproduct organizations and the heuristic evaluation of industrial structure' *Socio-Economic Planning Sciences*, 32(3) pp199-209

Johnes, G (1999) 'The management of universities: Scottish Economic Society / Royal Bank of Scotland Annual Lecture', *Scottish Journal of Political Economy*, 46 pp505-522.

Johnes, G & Johnes, J (1992) 'Apples and oranges: the aggregation problem in publications analysis' *Scientometrics*, 25(2) pp353-365

Johnes, G & Johnes, J (1993) 'Measuring the research performance of UK economics departments: an application of data envelopment analysis' *Oxford Economic Papers*, 45 pp332-347

Johnes, J (1992) 'The potential effects of wider access to higher education on degree quality' *Higher Education Quarterly*, 46(1) pp88-107

Johnes, J (1996) 'Performance assessment in higher education in Britain' *European Journal of Operational Research*, 89 pp18-33

Johnes, J (1997) 'Inter-university variations in undergraduate non-completion rates: a statistical analysis by subject of study' *Journal of Applied Statistics*, 24(3) pp343-361

Johnes, J (2003) 'Measuring teaching efficiency in higher education: an application of data envelopment analysis to graduates from UK universities 1993' Discussion Paper EC7/03 Department of Economics, Lancaster University

Johnes, J & Taylor, J (1987) 'Degree quality: an investigation into differences between universities' *Higher Education*, 16 pp581-602

Johnes, J & Taylor, J (1989a) 'The first destination of new graduates: comparisons between universities' *Applied Economics*, 21(3) pp357-374

Johnes, J & Taylor, J (1989b) 'Undergraduate non-completion rates: differences between UK universities' *Higher Education*, 18 pp209-225

Johnes, J & Taylor, J (1990) *Performance Indicators in Higher Education: UK Universities*, Open University Press and the Society for Research into Higher Education, Milton Keynes

Johnes, J & Taylor, J (1991) 'Non-completion of a degree course and its effect on the subsequent experience of non-completers in the labour market' *Studies in Higher Education*, 16(1) pp73-81

Johnes, J & Taylor, J (1992) 'The 1989 research selectivity exercise: a statistical analysis of differences in research rating between universities at the cost centre level' *Higher Education Quarterly*, 46(1) pp67-87

Kang, B-G & Greene, K V (2002) 'The effects of monitoring and competition on public education outputs: a stochastic frontier approach' *Public Finance Review*, 30(1) pp3-26

Kapur, R L (1972) 'Student wastage at Edinburgh University I: factors related to failure and dropout' *Universities Quarterly*, 26 pp353-377

Korhonen, P, Tainio, R & Wallenius, J (2001) 'Value efficiency analysis of academic research' *European Journal of Operational Research*, 130 pp121-132

Lovell, C A K (1993) 'Production frontiers and productive efficiency' in H O Fried, C A K Lovell & S S Schmidt (eds) *The Measurement of Productive Efficiency*, Oxford University Press, Oxford, pp3-67

Madden, G, Savage, S & Kemp, S (1997) 'Measuring public sector efficiency: a study of economics departments at Australian Universities' *Education Economics*, pp153-167

Mallier, T and Rodgers, T (1995) 'Measuring value added in higher education: a proposal' *Education Economics*, 3(2) pp119-132

Montague, W & Odds, F C (1990) 'Academic selection criteria and subsequent performance' *Medical Education*, 24 pp151-157

Naylor, R, Smith, J & McKnight, A (2000) 'Occupational earnings of graduates: evidence for the 1993 UK university population' Department of Economics, University of Warwick

Nisbett, J & Welsh, J (1972) '*The mature student*' Educational Research, 14 pp204-207

Portela, M C S & Thanassoulis, E (2001) 'Decomposing school and school-type efficiency' *European Journal of Operational Research*', 132(2) pp357-373

Rodgers, T & Ghosh, D (2001) 'Measuring the determinants of quality in UK higher education: a multinomial logit approach' *Quality Assurance in Education*, 9(3) pp121-126

Rudd, E (1984) 'A comparison between the results achieved by women and men studying for first degrees in British universities' *Studies in Higher Education*, 9 pp47-57

Sarrico, C S & Dyson, R G (2000) 'Using data envelopment analysis for planning in UK universities – an institutional perspective' *Journal of the Operational Research Society*, 51 pp789-800

Sarrico, C S, Hogan, S M, Dyson, R G & Athanassopoulos, A D (1997) 'Data envelopment analysis and university selection' *Journal of the Operational Research Society*, 48 pp1163-1177

Sear, K (1983) 'The correlation between A level grades and degree results in England and Wales' *Higher Education*, 12 pp609-619

Simar, L & Wilson, P W (1998) 'Sensitivity analysis of efficiency scores: how to bootstrap in nonparametric frontier models' *Management Science*, 44 (1) pp49-61

Simar, L & Wilson, P W (1999) 'Performance of the bootstrap for DEA estimators and iterating the principle' Discussion Paper, Université Catholique de LouvainSmith, J, McKnight, A & Naylor, R (2000) 'Graduate employability: policy and performance in higher education in the UK' *The Economic Journal*, 110 ppF382-F411

Smith, J, McKnight, A & Naylor, R (2000) 'Graduate employability: policy and performance in higher education in the UK' *The Economic Journal*, 110 ppF382-F411

Smith, J & Naylor, R (2001a) 'Determinants of degree performance in UK universities: a statistical analysis of the 1993 student cohort' *Oxford Bulletin of Economics and Statistics*, 63 pp29-60

Smith, J & Naylor, R (2001b) 'Dropping out of university: a statistical analysis of the probability of withdrawal for UK university students' *Journal of the Royal Statistical Society, Series A*, 164(2) pp389-405

Smith, P K (1990) 'The distribution of psychology degree classes in the UK' *Bulletin of the British Psychological Society*, 4 pp147-152

Stevens, P A (2001) 'The determinants of economic efficiency in English and Welsh universities' National Institute of Economic and Social Research, London, Discussion Paper no. 185

Tarsh, J (1982) ‘The correlation between A levels and degree performance’ Unit for Manpower Studies, mimeo

Thanassoulis, E (1999) ‘Setting achievement targets for school children’ *Education Economics*, 7(2) pp101-119

Thanassoulis, E & Portela, M C S (2002) ‘School outcomes: sharing the responsibility between pupil and school’ *Education Economics*, 10(2) pp183-207

Thanassoulis, E, Simpson, G, Battisti, G & Charlesworth-May, A (2003) 'DEA and multilevel modelling as alternative methods for assessing pupil and school performance' Discussion Paper, Aston Business School

Tomkins, C & Green, R (1988) ‘An experiment in the use of data envelopment analysis for evaluating the efficiency of UK university departments of accounting’ *Financial Accountability and Management*, 4(2) pp147-164

Walker, P (1975) ‘The university performance of mature students’ *Research in Education*, 14 pp 1-13

Woodhouse, G & Goldstein, H (1988) ‘Educational performance indicators and LEA league tables’ *Oxford Review of Education*, 14 pp301-319