Johnes, Jill

Efficiency and mergers in English higher education 1996/97 to 2008/09: Parametric and non-parametric estimation of the multi-input multi-output distance function

Original Citation

Johnes, Jill (2014) Efficiency and mergers in English higher education 1996/97 to 2008/09: Parametric and non-parametric estimation of the multi-input multi-output distance function. Manchester School, 82 (4). pp. 465-487. ISSN 1467-9957

This version is available at http://eprints.hud.ac.uk/id/eprint/23104/

EFFICIENCY AND MERGERS IN ENGLISH HIGHER EDUCATION 1996/97 TO 2008/09:

PARAMETRIC AND NON-PARAMETRIC ESTIMATION OF THE MULTI-INPUT MUTLI-OUTPUT DISTANCE FUNCTION

Jill Johnes[*]

Department of Economics

Lancaster University

LA1 4YX

E: j.johnes@lancaster.ac.uk

T: +44 (0)1524 592102

F: +44 (0)1524 594244

November 2011
Revised first December 2012
Revised second April 2013

Abstract

This paper explores the issue of efficiency in English higher education using data envelopment analysis and stochastic frontier analysis to estimate an output distance function (which incorporates measures of both quantity and quality of teaching and research inputs and outputs) over a thirteen-year period. The study compares the efficiency estimates derived from various estimation methods, and uses the results to provide guidance to researchers, managers and policy-makers on undertaking efficiency studies. The length of the study under consideration allows a preliminary statistical investigation of the effects on efficiency of merger activity in higher education.

INTRODUCTION

Publicly funded sectors are under pressure to deliver more for less, and none more so than the English higher education sector. This implies that there are efficiency savings to be made (Mandelson 2009) yet estimates of technical efficiency for the sector vary from 85% to 95% (Athanassopulos and Shale 1997; Flegg *et al.* 2004; Flegg and Allen 2007a; 2007b; Johnes 2008). Whether these estimates offer an accurate reflection of current efficiency is doubtful. First most are based on data covering periods prior to 2005, and so do not reflect the expanded sector which we observe in English higher education today. Second none of the efficiency studies has investigated the effect of merging on efficiency. As one suggestion for increasing efficiency is to merge institutions (Griffiths 2010) this is a serious omission. This study aims to provide an up-to-date picture of efficiency in English higher education, and to present a preliminary analysis of the effects on efficiency of mergers. As such, it will be useful to researchers, managers and policy-makers alike.

A merger is defined as the union of two or more institutions to form an entirely new entity. Examples of recent mergers in English higher education include: London Metropolitan University, which was created by merging the University of North London and London Guildhall University in 2002; the University of Manchester, which was formed by merging the Victoria University of Manchester with the University of Manchester Institute of Science and Technology in 2004; the University of Cumbria which evolved in 2007 from a merger of St. Martin's College, the Cumbria Institute of the Arts and the Cumbrian campuses of the University of Central Lancashire; and the University College Falmouth which merged with Dartington College of Arts in 2008. From a theoretical perspective, a merger may have efficiency benefits which accrue from returns to scale, as a consequence of increased administrative, economic and academic efficiency , or returns to scope if the merging institutions have complementary activities (Skodvin 1999; Harman 2000). While there has been qualitative work on higher education mergers based on case studies (Skodvin 1999), there are very few quantitative studies of the empirical effect of merging on efficiency (Hu and Liang 2008; Mao *et al.* 2009), and none in the context of English higher education.

The efficiency of an organisation can be assessed by examining its observed production relative to best practice in the industry. This requires estimation of the distance function. The distance function approach has a number of advantages: it handles a production situation with multiple inputs and multiple outputs; it does not assume any particular optimizing behaviour on the part of the firms; it does not require a knowledge of prices of either inputs or outputs; and it does not require prices to be exogenous (Coelli and Perelman 1999; Coelli 2000; Uri 2003). But the actual estimation of the distance function in the multiple input multiple output context poses something of a problem. Parametric estimation has obvious advantages in that it allows for stochastic errors, statistical inferences can be drawn, and the estimated parameters can provide potentially useful information on, for example, returns to scale and scope, and elasticities. The disadvantages are that results may be sensitive to the choices of functional form and error distribution. In addition, the parametric estimation of a multi-output multi-input distance function places

huge demands on the data. Non-parametric estimation of the distance function, in contrast, easily handles the multi-dimensional nature of production, and makes no assumptions regarding functional form or error distributions. The downside is that it does not allow for stochastic errors (such as measurement errors and random shocks), it is difficult to derive measures such as shadow prices and elasticities, and efficiency scores are biased upwards in the context of small sample sizes.

The purpose of the paper is to provide an up-to-date picture of efficiency in English higher education by estimating a multi-output multi-input distance function for English higher education institutions (HEIs) over the period 1996/97 to 2008/09 using both parametric and non-parametric methods. The sample's diversity and length allow a crude examination of performance in merging (both pre- and post-merger) and non-merging HEIs. An empirical analysis of the estimated efficiency effects of merger in the context of English higher education is new to the literature and is particularly relevant if mergers are likely in the wake of financial cuts. In addition, by using both parametric and non-parametric estimation methods, we are able to assess whether policy conclusions are likely to vary by method of estimation and, if so, provide guidelines to policy-makers regarding the most appropriate method for their needs.

The remainder of the paper is in five sections. Section I introduces the distance function methodology including estimation issues. The model specification is discussed in section II, and the data which underpin the analysis are described in section III. The results of estimating an output distance function using both stochastic frontier analysis (SFA) and data envelopment analysis (DEA) are reported and interpreted in section IV. Finally, conclusions are drawn in section V.

I DISTANCE FUNCTIONS

Methodology

HEIs produce multiple outputs from a variety of inputs. An assumption of separate production (commonly made) leads to the estimation of an individual production function for each output; but this does not capture the obvious jointness of production observed (Chizmar and Zak 1983). We therefore assume that HEIs use a vector of inputs $x \in \mathbb{R}_+^K$ to produce a vector of outputs $y \in \mathbb{R}_+^M$. Inputs to higher education, such as student intake, are often pre-determined by government policy and so an output-oriented perspective (whereby inputs are fixed and outputs are expanded proportionally) is used here.

The production technology of the HEI is defined as

$$P(x) = \{y \in \mathbb{R}_+^M : x \text{ can produce } y\} \tag{1}$$

The output distance function (Shephard 1970) is defined on the output set $P(x)$ as:

$$D(x, y) = min_\theta\{\theta : (y/\theta) \in P(x)\} \tag{2}$$

The distance function $D(x, y)$ is non-decreasing, positively linearly homogeneous of degree +1 in y, convex in y, and decreasing in x. It follows that

$$D(x, y) \leq 1 \Longleftrightarrow y \in P(x) \tag{3a}$$

$$D(x, y) = 1 \Longleftrightarrow y \in \text{Bound}P(x) \tag{3b}$$

where $\text{Bound}P(x)$ is the frontier of the output set (see Coelli *et al.* 2005). If *y* is located on the boundary of the production possibility set, $D(x, y) = 1$ and this represents technical efficiency; if $D(x, y) < 1$, *y* lies inside the frontier and technical inefficiency exists.

Parametric estimation

The parametric approach assumes a functional form for the distance function, estimates of the parameters are provided, and the significance of these can be tested. The wrong choice of functional form, however, introduces problems of misspecification the consequences of which can be serious: parameter estimates will be biased and efficiency estimates may be incorrect. Multicollinearity and omission of relevant variable(s) may also cause problems.

The desirable properties of the functional form are that it should be i) flexible; ii) easy to estimate; and iii) permit the imposition of homogeneity (Coelli and Perelman 2000). The translog fulfils all three criteria and has been used to estimate distance functions in various contexts ranging from agriculture to telecommunications (Paul *et al.* 2000; Uri 2003; Karagiannis *et al.* 2004; Balcombe *et al.* 2007). The translog distance function is defined below for *N* HEIs using inputs x_k (*k* = 1,...,*K*) to produce outputs y_m (*m* = 1,...,*M*):

$$\ln D_{it}(x, y) =$$

$$\alpha_0 + \sum_{m=1}^{M} \alpha_m \ln y_{mit} + \frac{1}{2} \sum_{m=1}^{M} \sum_{n=1}^{M} \alpha_{mm} \ln y_{mit} \ln y_{nit} + \sum_{k=1}^{K} \beta_k \ln x_{kit} +$$

$$\frac{1}{2} \sum_{k=1}^{K} \sum_{l=1}^{K} \beta_{kl} \ln x_{kit} \ln x_{lit} + \sum_{k=1}^{K} \sum_{m=1}^{M} \delta_{km} \ln x_{kit} \ln y_{mit} \qquad i = 1, 2, \dots, N \tag{4}$$

where subscript *it* refers to the *i*th HEI in the *t*th time period. Distance function restrictions require the following conditions to hold:

a) Homogeneity of degree +1 in outputs

$$\sum_{m=1}^{M} \alpha_m = 1 \quad \text{and} \tag{5a}$$

$$\sum_{n=1}^{M} \alpha_{mn} = 0 \quad m = 1, 2, \dots, M \quad \text{and} \tag{5b}$$

$$\sum_{m=1}^{M} \delta_{km} = 0 \quad k = 1, 2, \dots, K \tag{5c}$$

b) Symmetry:

$$\alpha_{mn} = \alpha_{nm} \quad m, n = 1, 2, \dots, M \quad \text{and} \tag{6a}$$

$$\beta_{kl} = \beta_{lk} \quad k, l = 1, 2, \dots, K \tag{6b}$$

By the homogeneity in outputs restriction, $D(x, \omega y) = \omega D(x, y)$ and so the *M*th output can be chosen arbitrarily such that $\omega = 1/y_M$. Thus equation (4) can be written as:

$$-\ln y_{Mit} = \alpha_0 + \sum_{m=1}^{M-1} \alpha_m \ln \left(\frac{y_{mit}}{y_{Mit}} \right) + \frac{1}{2} \sum_{m=1}^{M-1} \sum_{n=1}^{M-1} \alpha_{mn} \ln \left(\frac{y_{mit}}{y_{Mit}} \right) \ln \left(\frac{y_{nit}}{y_{Mit}} \right) + \sum_{k=1}^{K} \beta_k \ln x_{kit} +$$

$$\frac{1}{2} \sum_{k=1}^{K} \sum_{l=1}^{K} \beta_{kl} \ln x_{kit} \ln x_{lit} + \sum_{k=1}^{K} \sum_{m=1}^{M-1} \delta_{km} \ln x_{kit} \ln \left(\frac{y_{mit}}{y_{Mit}} \right) + \varepsilon_{it} \qquad i = 1, 2, \dots, N \tag{7}$$

where $\varepsilon_{it} = -\ln D_{it}(x, y)$

It has been suggested that there is endogeneity in this model caused by the explanatory variables being related to the error term (ε_{it}) (Coelli *et al.* 2005; O'Donnell 2011). This would inevitably lead to simultaneous equations bias and require estimation using instrumental variables (Atkinson *et al.* 2003) or Bayesian methods (Fernández *et al.* 2000; O'Donnell 2011). It has been argued, however, that such bias is not a problem in an output distance function which (as here) uses a translog functional form (Coelli and Perelman 2000).

The quantity which is of interest in equation (7) is the distance (or efficiency) $\ln D_{it}(x, y)$ which is measured by the error term. The choice of parametric estimation method depends on the assumptions made about this error term. Three alternative estimation methods will be used in the following analysis, all of which assume that the error can be split into two components i.e. $\varepsilon_{it} = v_{it} - u_{it}$.

a) Random effects (RE) estimation assumes v_{it} to represent statistical noise and $u_{it} = u_i$ (i.e. technical inefficiency) to be time-invariant; it makes no distributional assumptions about u_i.

b) A time-invariant (TI) stochastic frontier model (Aigner *et al.* 1977) assumes v_{it} and u_{it} are IID such that $v_{it} \sim N(0, \sigma_v^2)$, $u_{it} = u_i \sim N^+(\mu, \sigma^2)$ where N^+ represents a truncated-normal distribution truncated at 0. In the following, u_i is treated as a random variable and the TI stochastic frontier model of equation (7) is estimated by maximum likelihood methods.

c) A time-varying decay (TVD) model assumes that v_{it} and u_{it} are IID such that $v_{it} \sim N(0, \sigma_v^2)$, $u_{it} = \{exp[-\eta(t - T_i)]\}u_i$ where T_i is the last period in the *i*th panel, η is a decay parameter to be estimated, and u_i is the base level of inefficiency which in this case is the inefficiency for the last period observed for unit *i* (Battese and Coelli 1992). In the following, the TVD stochastic frontier model of equation (7) is estimated in a random effects framework by maximum likelihood methods.

Non-parametric estimation

The DEA approach does not require a functional form and allows each firm to have different objectives. However, there are no parameter estimates (hence no significance tests), information such as elasticities or economies of scope can be difficult to derive, and results can be severely affected by the presence of outliers. Given the diversity of the English higher education sector this could be a severe drawback. DEA makes no allowance for stochastic errors; while recent developments have introduced bootstrapping methods to the DEA methodology, such methods address issues of sampling variability rather than stochastic error (Coelli *et al.* 2005).

Taking a DEA approach, the technical efficiency of the firm or decision making unit (DMU) *i* at time *t* ($TE_{it} = D_{it}(x,y)$) is defined as (Charnes *et al* 1978; 1979):

$$TE_{it} = \frac{\sum_{m=1}^{M} a_{mt} y_{mit}}{\sum_{k=1}^{K} b_{kt} x_{kit}} \qquad (8)$$

where y_{mit} and x_{kit} are as already defined; a_{mt} is the weight applied to output m in time t and b_{kt} is the weight applied to input k in time t. For each DMU, weights are found by maximizing efficiency subject to the constraints that weights must be non-zero and universal. DEA can be applied in the context of constant returns to scale (CRS) or variable returns to scale (VRS) . DMU i is efficient in time t if $TE_{it} = D_{it}(x,y) = 1$.

Few studies have compared efficiency values of HEIs derived using both parametric and non-parametric output distance functions. Efficiencies derived from parametric and non-parametric estimations of cost functions for the Canadian higher education sector are significantly though not particularly highly correlated (McMillan and Chan 2006). A higher correlation is found in the context of German universities, but the DEA and SFA models are not entirely comparable (Kempkes and Pohl 2010).

II MODEL SPECIFICATION

Inputs and outputs

The input and output variables used in this study are constructed from annual statistics for all HEIs in England published by the Higher Education Statistics Agency (HESA). HEIs can be seen as using primary inputs, capital and labour to produce teaching and research outputs. Five measures of inputs are specified (see table 1 for detailed definitions). Primary inputs are undergraduates and postgraduates. These vary by institution in terms of quality on entry. Undergraduate students (UGINPUT) are weighted by average entry score (QUAL), as reflected by the average A level score of those students entering on the basis of A levels, to produce the quality-adjusted input variable UGINQUAL. There is no satisfactory recorded measure of quality for postgraduates and so postgraduate student numbers (PGINPUT) are used to reflect this input. The number of academic staff (STAFF) and the expenditure on administration and central services (ADMIN) reflect, respectively, academic and non-academic labour inputs. Finally, capital inputs are measured by expenditure on library, computing and other learning resources (ACSERV). Three measures of outputs are included in the model. Undergraduate teaching output is measured by number of undergraduate first degree qualifications (UGOUTPUT) weighted by degree classification (see table 1) to produce the output variable UGOUTQUAL; postgraduate teaching output is measured by number of postgraduate degree qualifications (PGOUTPUT); income received for research purposes (RESEARCH) is included to reflect research output.

Table 1 here

There are some obvious shortcomings of the model specification. First, owing to data constraints, postgraduate inputs and outputs are not adjusted for quality (see Johnes 2008 for further discussion). Moreover, while postgraduate teaching input and output are correlated (as indeed are the undergraduate teaching input and output), HEIs experience non-completion to vastly differing extents (HEFCE 2010; HESA 2012). The inclusion of student numbers as input and graduation numbers as output is intended to capture this effect.

Second, in contrast to most previous studies (an exception is Athanassopulos and Shale 1997), an attempt is made here to adjust undergraduate input for quality. The method is inexact because the average A level score is observed only for students entering each university with A level qualifications rather than for all students. We therefore weight UGINPUT (number of all undergraduate students) by QUAL (average entry score of those entering on the basis of A levels), and so QUAL is assumed to provide an adequate reflection of inter-university differences in the entry quality of all undergraduates. An additional problem is that the units of measurement of the entry score data vary between two periods: thus a transformation is applied to entry scores for 1996/97 to 2001/02 to convert them to the same scale as that used in 2002/03 to 2008/09. Sensitivity of the results to this adjustment is tested by producing additional results for the restricted period of 2002/03 to 2008/09.

The third problem concerns the quality of undergraduate teaching output. Using the number of 'good' degrees (Flegg *et al.* 2004; Flegg and Allen 2007a; 2007b) would be inappropriate here given the diversity of the HEIs (and hence the undergraduate qualifications) in the sample. But an unweighted count of all degree qualifications (Johnes 2008) would disadvantage those HEIs producing higher quality. This study seeks to address these problems by weighting graduates from undergraduate programmes by class of degree.

The weighting of both undergraduate inputs and outputs using entry and exit qualifications presupposes that the steady increases in the entry and graduation grades of students over time represent real increases in quality rather than grade inflation. In order to assess the sensitivity of the results to quality weightings, the analysis has been performed without taking quality into account in either undergraduate teaching inputs or outputs and results are similar to those found here (Johnes 2010).

Fourth, subject differences are not reflected in either teaching or research outputs. Disaggregation of outputs even by broad subject areas would cause an unacceptable loss of degrees of freedom in the parametric model. Inter-university variations in subject mix, however, may be accounted for to some extent, first by the inclusion of HEI type dummies, and second by the estimation methods: RE allows for unobserved heterogeneity, and DEA calculates weights which are unique to each DMU in the data set.

Fifth, the measure of research output is open to debate. Publications data are not easily available; research ratings produced by the Research Assessment Exercises (RAEs), although attractive in terms of incorporating both quantity and quality, (Glass *et al.* 2006), may be biased (Clerides *et al.* 2011) and are available only at intervals over the study period (1996, 2001 and 2008). We therefore resort to using the more easily available measure of research income. Such a measure of research output is now generally accepted in these types of studies (Flegg *et al.* 2004; Flegg and Allen 2007a; 2007b; Worthington and Lee 2008); indeed HESA, in its published data on research outputs includes research grants and contracts and funding council grants for research amongst its measures (HESA 2012). Moreover the RAE results for the year 2008 correlate highly with the grant income measure for the same year (Spearman's rho is 0.942).

Grant income is also an attractive measure of research in that it provides an up-to-date picture of research activity and output in the current academic year; so problems of time lag between the input to the process and the output from it, which might be encountered when using, for example, citation counts or number of patents (Hashimoto and Haneda 2008), are not of such importance when using the research income measure. The possibility of an inter-temporal relationship between the inputs and outputs of the higher education process is worthy of further discussion. One study, which develops a dynamic DEA model to capture the inter-temporal aspect, compares the results of the dynamic model with those derived from a static (or conventional) DEA model in the context of higher education, and finds considerable overall agreement between the efficiencies produced from the two approaches (Emrouznejad and Thanassoulis 2005). Given this conclusion and the assertion that there are no clear criteria for deciding on the appropriate length of lag and subdivision into periods required by this approach (Emrouznejad and Thanassoulis 2005), we opt for the static model here.

Finally, the outputs included here encompass only the teaching and research functions of HEIs. Universities also produce 'third mission' or social output. This includes such services as the storage of knowledge, the provision of advice to business and comment on issues of public interest. In line with most previous studies, no attempt is made here to find and include a measure of this output, and results may be biased as a consequence of the omission.

Efficiency over time

HEIs are likely to experience productivity change over a 13-year period. This might be due to changes in efficiency (movements in relation to the frontier) or to changes in technology (shifts in the frontier). Previous studies examining English higher education over earlier periods suggest positive but small annual average increases in productivity which are more a consequence of advances in technology than improvements in technical efficiency (Flegg *et al.* 2004; Johnes 2008). Changes over time will be investigated here as follows: in the TI parametric models, technical efficiency changes are assumed to be zero while technology changes are investigated using a set of individual year dummies. These year dummies are also included in the TVD parametric model to capture technology change, while the inclusion of the parameter η allows for technical efficiency change. The DEA models will be estimated first using the pooled data (thereby assuming a common frontier and hence zero technology change over the whole period) and second by taking a within-year approach to allow the frontier (and hence technology) to change over time. A caveat should be noted regarding this approach: the inevitable reduction in sample size resulting from the within-year estimation may affect the efficiency estimates (Zhang and Bartels 1998).

HEI type

Three types of HEIs will be distinguished in this study. First, pre-1992 HEIs are traditional universities which had university status prior to the Further and Higher Education Act of 1992. These institutions undertake

teaching (undergraduate and postgraduate) and research in a whole range of subjects including (unlike other types of institution) medical and veterinary sciences. The second group of post-1992 HEIs are former polytechnics which, by the provision of the Further and Higher Education Act, have, since 1992, been allowed to award their own degrees and use the title university. The third group of HEIs are institutions which are, or have recently been, colleges of higher education. Some of these HEIs are specialist institutions concentrating on a particular discipline. Since 2003, these colleges have been allowed to apply for university (and degree-awarding) status. Previous research suggests that technical efficiency is highest amongst the former colleges of higher education, followed by the post- and pre-1992 HEIs (Johnes 2008).

Model

The precise specification of the parametric distance function to be estimated is:

$$-\ln y_{3it} = \alpha_0 + \sum_{m=1}^{2} \alpha_m \ln\left(\frac{y_{mit}}{y_{3it}}\right) + \frac{1}{2}\sum_{m=1}^{2}\sum_{n=1}^{2} \alpha_{mn} \ln\left(\frac{y_{mit}}{y_{3it}}\right)\ln\left(\frac{y_{nit}}{y_{3it}}\right) + \sum_{k=1}^{5} \beta_k \ln x_{kit} +$$

$$\frac{1}{2}\sum_{k=1}^{5}\sum_{l=1}^{5} \beta_{kl} \ln x_{kit} \ln x_{lit} + \sum_{k=1}^{5}\sum_{m=1}^{2} \delta_{km} \ln(x_{kit})\ln\left(\frac{y_{mit}}{y_{3it}}\right) + \sum_{t=2}^{13} \tau_t t_{it} + \omega_{1i}T_{1i} + \omega_{2i}T_{2i} + v_{it} -$$

$$u_{it} \qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad i = 1,2,\dots,N \qquad (9)$$

where t_{it} are time dummies (where the first year of the sample is the base year) included to capture changes in productivity over time; T_{1i} and T_{2i} are dummy variables to reflect (former) colleges of higher education and post-1992 HEIs respectively. The numeraire is y_3 = UGOUTQUAL. In order to check the sensitivity of the results to this choice of numeraire, results presented in this paper have also been generated using respectively y_1 (PGOUTPUT) and y_2 (RESEARCH) as numeraire. In line with previous research (Coelli and Perelman 2000; Paul and Nehring 2005), results are remarkably insensitive to choice of numeraire and all conclusions remain the same.

Two variants of this parametric model are estimated in the subsequent analysis: in model 1, t = 1,…,13 whereas in model 2 t = 7, …, 13, for reasons discussed previously. In both cases, the error u_{it} is estimated using, respectively, a) RE TI, b) SFA TI and c) SFA TVD, thus providing a total of 6 sets of parametric efficiency estimates. All input and output variables are mean-corrected (Cuesta and Orea 2002).

In the context of DEA, two model variants are estimated: in model 3, t = 1,…,13 whereas in model 4 t = 7, …, 13. In both cases, estimation is performed, respectively, on the assumption of a) CRS and b) VRS (output-oriented), and each of these variants is estimated on i) the pooled sample and ii) within year, thus providing 8 sets of non-parametric efficiency estimates. In all cases we present the bias corrected efficiencies calculated using the homogeneous bootstrapping algorithm (Simar and Wilson 2008). The bootstrap method provides DEA efficiency estimates which are corrected for sampling variability.

While every effort is made to make the DEA and SFA models as similar to each other as possible (in order to compare like with like), it should be borne in mind when interpreting the comparisons of efficiencies that

there are underlying differences. In order to help with the interpretation, table 2 highlights the similarities and differences between the models.

Table 2 here

III DATA

The analysis is based on a panel of data covering 13 years from 1996/97 to 2008/09. The sample is unbalanced for a number of reasons. First, some HEIs merged during the study period. Following merger the new institution was treated as a different entity from the HEIs which merged to form it (this approach is similar to Cuesta and Orea 2002). Second, some HEIs entered the higher education sector during the period. Third, HEIs which produced a zero amount of any output or used a zero amount of any input during a given year were removed from the sample in that year. Finally, four HEIs were removed entirely from the sample: Open University was removed because of its large size and unique nature of teaching provision; the University of London (Institutes and activities) was also excluded on the grounds that the composition of the component HEIs recorded under this umbrella changed over time; University of Buckingham was deleted because it is not publicly funded, and Heythrop College because it only became publicly funded during the time period under consideration. The number of HEIs included in each year therefore varies from 108 to 113, and the panel totals 1444 observations.

Descriptive statistics are provided in table 3 for all variables across the pooled sample. The average HEI produces just over 1000 graduates from postgraduate degrees and over 2500 graduates from undergraduate degrees. In addition, it receives over £72 million in research grants. The inputs of the average HEI are nearly 2000 postgraduate students, 7000 undergraduate students, and just over 900 academic staff. In addition it spends over £9 million on academic services and nearly £17 million on administration. An examination of the inputs and outputs over time suggests that there has been an upward trend in all inputs over the entire period; outputs have risen until the final year when there has been a slight fall.

Table 3 here

The standard deviations of the input and output variables reported in table 3 demonstrate the wide diversity between HEIs in England. Institutions vary tremendously in size, but the standard deviation of research grants (nearly £80 million) implies that research activity varies particularly widely across the sector. An examination of inputs and outputs by HEI type reveals that former colleges of higher education are considerably smaller in terms of both inputs used and outputs produced than either of the other two groups of HEIs (details are available at http://www.lancs.ac.uk/people/ecajj/). The emphasis on the different production activities also varies by type. Research activity is dominant in pre-1992 universities where the value of research income is typically twice that in post-1992 HEIs and 6 times that in former colleges of higher education. Undergraduate teaching, on the other hand is predominantly in the post-1992 HEIs (both before and after adjusting for quality), and postgraduate outputs are produced almost equally

by both pre- and post-1992 HEIs at around 3 times the volume of former colleges of higher education. Postgraduate inputs are higher in pre- than post-1992 HEIs, as are staff, administrative and capital inputs. The relative size of undergraduate input varies according to whether quality is taken into account: the volume of undergraduate inputs is smaller in pre- than post-1992 HEIs, but this is reversed when quality is incorporated into the measure.

IV RESULTS

Efficiency scores

The mean efficiencies derived from estimating equation (9) using the parametric approaches described in section I are presented in table 4 (a complete table of estimated parameters is available at http://www.lancs.ac.uk/people/ecajj/). The presence of technical inefficiency in the SFA TI (SFA TVD) model is assessed using the test of H_0: $\sigma_u^2 = 0$ (H_0: $\mu = 0$) (Coelli *et al.* 2005). In both SFA models, H_0 is rejected and so inefficiency is significant. SFA efficiencies are remarkably similar across models and samples (the mean score is around 80%) except in the case of the SFA TVD model for the sub-sample; in this case, the null hypothesis $\eta = 0$ cannot be rejected; this model cannot be preferred to the SFA TI model, and it is therefore not included in the ensuing discussion.

The first main result pertains to the level of efficiency estimated for the English higher education sector. Parametric estimates of mean efficiency are at the lower end (at around 72% to 80%); within-year DEA estimates (see table 4) are at the higher end (at around 86% to 93%); and pooled DEA estimates (see table 4) overlap with the parametric estimates (at around 75% to 84%). While there is some variation in estimates of mean efficiency (the within-year DEA estimates in particular differ from the other results), all methods agree that there is considerable scope at the lower end for HEIs to improve: minimum efficiency is at or below 50% across the models. This poor performance is explored further in figure 1 which presents the efficiency score and associated 95% confidence interval (Horrace and Schmidt 1996; Simar and Wilson 2008) for each HEI for the final year of the study, calculated using, respectively, SFA TVD and pooled DEA VRS. The parametric approach, unlike DEA, applies the same parameters of the distance function across the diverse array of HEIs in the sample, and this might be why the SFA TVD model identifies around 6 HEIs as particularly inefficient (i.e. with an efficiency score below 0.6 – see figure 1a). Given the diversity of the HEIs in this data set, perhaps a random parameter distance function might be a more appropriate approach, but this makes considerable demands on the data and is beyond the scope of the present paper.

Table 4 here

Figures 1a and 1b here

Policy-makers should be aware that estimates of the *level of* efficiency (and hence of the scope for efficiency savings) vary by estimation method. This is not such a problem if both methods provide the same ranking. Further scrutiny finds that whilst rankings for all pairs of efficiency scores are significantly

positively correlated, the level of correlation is fairly low (details of Spearman's rank correlations are available from http://www.lancs.ac.uk/people/ecajj/). The most closely correlated rankings are those derived from the DEA VRS (applied to pooled data) and the SFA TVD models (Spearman's rho is 0.49). The rank correlations between the within-year DEA efficiencies (pooled DEA efficiencies) and efficiencies from parametric models vary from 0.268 to 0.332 (0.414 to 0.494). The pooled DEA approach therefore provides estimates of efficiencies which are generally more closely correlated with the estimates from the parametric models. Thus the within-year DEA results (which also differ in terms of level of mean efficiency) should be treated with particular caution.

The plots in figure 1 suggest that the top and bottom groups of HEIs are distinct (i.e. significantly different from each other) but that we cannot distinguish between the HEIs in the middle in terms of their efficiency. But do the different methods identify the same group of high- and low-performing institutions? A scatter plot of rankings from the pooled DEA VRS and the SFA DVD models (available from http://www.lancs.ac.uk/people/ecajj/) indicates that the two methods identify common but very small sets of HEIs at the top and bottom ends of performance. This offers further confirmation that rankings of middle performing HEIs (regardless of estimation approach) are not particularly informative.

Efficiency over time

This section reports the patterns of change in efficiency over time thus providing a context for the subsequent examination of efficiency in merging and non-merging HEIs. The estimated parameters of the SFA TVD model, the only parametric model to allow for both technical efficiency and technology change, suggest that there has been no significant technology change while technical efficiency has been decreasing by 1.4% per annum over the period 1996/97 to 2008/09 and by almost 1% per annum over the period 2002/03 to 2008/09 (full details are available at http://www.lancs.ac.uk/people/ecajj/). Given that there is no significant change in the frontier over the two periods, this implies that the inefficient institutions are becoming relatively more inefficient compared to the frontier HEIs over time.

The SFA TVD and pooled DEA models are very similar in patterns of mean efficiency over time. This is perhaps not surprising since the former found no significant technology change and the latter models do not allow for technology change. The within-year DEA models, however, show a more random pattern of mean efficiency over time, and there is no obvious trend. It is likely that the problem of DEA estimation within small samples (Zhang and Bartels 1998) is affecting the results of the within-year DEA models estimated here, and these results are likely to be unreliable as a consequence.

Efficiency by HEI type

The estimated coefficients from the parametric models (see http://www.lancs.ac.uk/people/ecajj/) provide evidence that the former colleges of higher education are significantly more efficient than pre-1992 universities (full details are available at http://www.lancs.ac.uk/people/ecajj/). The DEA efficiencies (see

table 5) generally suggest that both former colleges of higher education and post-1992 HEIs are, on average, more efficient than their pre-1992 counterparts. It is therefore clear that the differences in the input and output characteristics of these different types of institutions, both measured and unmeasured (for example, differences in subject mix) feed through into differences in efficiency. The poorer performance of the pre-1992 HEIs may be a consequence of many of these HEIs being involved in teaching medical and veterinary sciences which are resource intensive. Thus attempts to measure efficiency in this diverse sector need explicitly to take into account inter-institutional differences in provision.

Table 5 here

What, then, should the policy-maker take away from the efficiency results presented so far? With regard to method of estimation, the DEA within-year approach provides little discrimination between institutions, and the efficiencies fluctuate from one period to the next. These methods are unlikely to be reliable in providing efficiency trends in the higher education sector over time, and hence should be avoided.

This leaves two pooled DEA and three parametric approaches. The main finding is that none of these methods of efficiency estimation provides the fine discrimination in institutional rankings which might be deemed desirable. At best, they identify common (small) sets of high- and low-performing HEIs, but there is little difference in the efficiency of the remaining institutions, and rankings of these HEIs should not be used. In this respect, then, the methods are in agreement, and any one of these estimation methods can provide estimates of the very best- and worst-performing HEIs.

The choice of method therefore depends on the additional requirements (beyond identifying best- and worst-performing HEIs) of the efficiency study, and the underlying assumptions therefore become relevant (see table 2 for an overview). If the researcher is interested in efficiency over time then the parametric SFA TVD is the most flexible in terms of allowing for both efficiency and technology change. The other four methods allow for one or the other but not both. If information is required on, for example, economies of scale, economies of scope and elasticities then this is also easily calculated from the SFA TVD model; it is not so easy to derive using the DEA approaches. For these reasons, the SFA TVD model should be the preferred method.

A final caveat is needed, however: the SFA approach applies the same parameters to all units. In a diverse sector such as English higher education this is potentially a severe disadvantage of the parametric approach. The development of a random parameters variant of SFA which allows units to differ in their objectives (as is the case with DEA) is therefore a welcome advance and should be considered (data permitting) in future studies of efficiency.

Merger activity

A merger is defined as the union of two or more institutions to form an entirely new entity. Included in the sample data are 19 instances of (horizontal) merger and this enables us to examine the effects on efficiency

of merger. Of the 19 instances of merger, 4 occurred between pairs of pre-1992 HEIs, 4 between pairs of former colleges of higher education, 8 between pre-1992 HEIs and former colleges of higher education, 2 between post-1992 HEIs and former colleges of higher education, and 1 between pre- and post1992 HEIs. A comparison of the mean efficiency of three types of HEIs (pre-merging HEIs; merged HEIs; and non-merging HEIs) in table 6 shows that average efficiency is considerably higher amongst merged than pre-merger and non-merging institutions. The null hypothesis of identical means in the three groups is rejected in all cases. This finding also holds when the investigation is carried out by year (detailed results can be found at http://www.lancs.ac.uk/people/ecajj/). These results should be accompanied by a strong caveat, however: whether the difference in efficiency is a consequence of the merger or of some other underlying characteristic(s) – and merging and non-merging HEIs are very different in terms of their characteristics – is unknown.

Table 6 here

Finally, rather than look at the 'typical' pre- and post-merging HEI, such a small number of mergers permits a close examination of efficiency changes in individual merger cases. The patterns of efficiency (where efficiency is measured by the values obtained from applying DEA VRS to the pooled data) before and after merger are displayed for a sample of mergers in figures 2a to 2d. While some mergers seem to have had unambiguous efficiency gains (see, for example, graph (i) in figure 2a and graph (ii) in figure 2c), others have had less clear benefits for some of the participants (see, for example, graphs (ii) and (iii) in figure 2b). In some cases, efficiency declines over time following the merger (see, for example, graph (i) in figure 2c). Thus the positive picture of merger activity suggested by the mean efficiencies conceals the fact that there can be losers as well as winners in a merger.

Figures 2a to 2d here

V CONCLUSIONS

HEIs are likely to face tight fiscal constraints over the coming years. There has been some suggestion that cuts can be absorbed by increased efficiency which in turn may be effected by merging some HEIs. The purpose of this study is to estimate a multi-input multi-output distance function in order to provide a better understanding of potential for efficiency improvements in the sector. The results are derived from a panel data set of English HEIs over a period from 1996/97 to 2008/09. This is a period of rapid change both in terms of expansion of all inputs and outputs and in terms of considerable merger activity.

Of the seven estimation methods considered in this study, two (the within-year DEA models) should be discounted: they give levels, rankings and patterns of efficiency which do not conform well with other methods. The remaining five methods suggest that average efficiency in the English university sector over the period 1996/97 to 2008/09 is around 75% to 83%. Rank correlations between these five methods are significantly positive but low (between 0.4 and 0.5). While the SFA TVD and pooled DEA methods provide the most consistent results in this application, confidence interval plots suggest that any method of

efficiency estimation can at best discriminate only between the very highest- and lowest-performing universities; those in the middle are not significantly different from one another. University rankings should therefore come with a serious health warning and be handled with extreme caution.

In the context of merger activity, the typical HEI involved in a merger has efficiency which is similar to (or slightly higher than) the average non-merging HEI. Thus merger activity in English higher education seems not to be a reaction to a crisis in efficiency in the merging HEIs. Of particular interest to institutions considering merger is the result that the typical merged HEI is significantly more efficient than either pre-merger or non-merging HEIs, suggesting that, on average, merging is a positive activity. It should be remembered that the characteristics of the three types of institutions (pre-, post- and non-merging HEIs) are considerably different and so the efficiency differences may be a consequence of something other than the merger. Moreover, closer scrutiny of the individual mergers suggests that the effects can vary by the types of HEI participating in the merger, and that there are both winners and losers in the merging process. A more rigorous analysis of the effects of merger would construct a matched sample (or a control group) of non-merging HEIs with similar characteristics to those of the merging HEIs with which to make comparisons. Thus the results here should not be seen as definitive, but rather as a signal of the need for further empirical investigation of merger activity.

REFERENCES

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

Athanassopulos, A. and E. A. Shale (1997). 'Assessing the comparative efficiency of higher education institutions in the UK by means of data envelopment analysis.' *Education Economics* 5(2): 117-135.

Atkinson, S. E., C. Cornwell and O. HonerKamp (2003). 'Measuring and decomposing productivity change: stochastic distance function estimation vs DEA.' *Journal of Business and Economic Statistics* 21(2): 284-294.

Balcombe, K., H. Doucouliagos and I. Fraser (2007). 'Input usage, output mix and industry deregulation: an analysis of the Australian dairy manufacturing industry.' *The Australian Journal of Agricultural and Resource Economics* 51(2): 137-156.

Battese, G. E. and T. J. Coelli (1992). 'Frontier production functions, technical efficiency and panel data: with application to paddy farmers in India.' *Journal of Productivity Analysis* 3: 153-169.

Chizmar, J. F. and T. A. Zak (1983). 'Modeling multiple outputs in educational production functions.' *American Economic Review* 73(2): 18-22.

Clerides, S., P. Pashardes and A. Polycarpou (2011). 'Peer review vs metric-based assessment: testing for bias in the RAE ratings of UK Economics departments.' *Economica* 78(311): 565-583.

Coelli, T. J. (2000). 'On the econometric estimation of the distance function representation of a production technology.' *Working Paper* 2000/042 Université catholique de Louvain, Center for Operations Research and Econometrics (CORE), Louvain.

Coelli, T. J. and S. Perelman (1999). 'A comparison of parametric and non-parametric distance functions: with application to European railways.' *European Journal of Operational Research* 117: 326-339.

Coelli, T. J. and S. Perelman (2000). 'Technical efficiency of European railways: a distance function approach.' *Applied Economics* 32: 1967-1976.

Coelli, T. J., D. S. P. Rao, C. J. O'Donnell and G. E. Battese (2005). *An Introduction to Efficiency and Productivity Analysis*. New York, Springer.

Cuesta, R. A. and L. Orea (2002). 'Mergers and technical efficiency in Spanish savings banks: a stochastic distance function approach.' *Journal of Banking and Finance* 26(12): 2231-2247.

Emrouznejad, A. and E. Thanassoulis (2005). 'A mathematical model for dynamic efficiency using data envelopment analysis.' *Applied Mathematics and Computation* 160: 363-378.

Fernández, C., G. Koop and M. Steel (2000). 'A Bayesian analysis of multiple-output production frontiers.' *Journal of Econometrics* 98(1): 47-79.

Flegg, T. and D. Allen (2007a). 'Does expansion cause congestion? The case of the older British universities, 1994-2004.' *Education Economics* 15(1): 75-102.

Flegg, T. and D. Allen (2007b). 'Using Cooper's approach to explore the extent of congestion in the New British universities.' *Economics Issues* 12(2): 47-81.

Flegg, T., D. Allen, K. Field and T. W. Thurlow (2004). 'Measuring the efficiency of British universities: a multi-period data envelopment analysis.' *Education Economics* 12(3): 231-249.

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

Griffiths, S. (2010). 'Be afraid, parents'. *The Sunday TImes*. London Retrieved from http://www.timesonline.co.uk/tol/life_and_style/education/article6990896.ece accessed 25th February 2010.

Harman, G. (2000). 'Institutional mergers in Australian higher education since 1960.' *Higher Education Quarterly* 54(4): 343-366.

Hashimoto, A. and S. Haneda (2008). 'Measuring the chnage in R&D efficiency of the Japanese pharmaceutical industry.' *Research Policy* 37: 1829-1836.

HEFCE (2010). 'Research degree qualification rates: Full-time students starting in 2000-01, 2001-02 or 2003-03; Part-time and mode-switch students starting in 1999-2000.' HEFCE 2010/21 Higher Education Funding Council for England, Bristol.

HESA (2012). 'Performance Indicators in Higher Education in the UK.' Retrieved 30th November 2012, from http://www.hesa.ac.uk/index.php?option=com_content&task=view&id=2072&Itemid=141.

Horrace, W. and P. Schmidt (1996). 'Confidence statements for efficiency estimates from stochastic frontier models.' *Journal of Productivity Analysis* 7(2-3): 257-282.

Hu, Y. and W. Liang (2008). 'Malmquist index analysis of the dynamic changes in scientific research productivity of some Chinese universities before and after merger.' *Frontiers of Education in China* 3(3): 429-447.

Johnes, J. (2008). 'Efficiency and productivity change in the English higher education sector from 1996/97 to 2004/05.' *The Manchester School* 76(6): 653-674.

Johnes, J. (2010). 'Efficiency and input- and output-substitutability in English higher education 1996/97 to 2007/08.' *LUMS Working Papers* 2010/025. http://eprints.lancs.ac.uk/49007/1/Document.pdf Lancaster University Management School, Lancaster.

Karagiannis, G., P. Midmore and V. Tzouvelekas (2004). 'Parametric decomposition of output growth using a stochastic input distance function.' *American Journal of Agricultural Economics* 86(4): 1044-1057.

Kempkes, G. and C. Pohl (2010). 'The efficiency of German universities - some evidence from nonparametric and parametric methods.' *Applied Economics* 42(16): 2063-2079.

Mandelson, P. (2009). 'Higher Education Funding 2010-11'. *Letter to HEFCE* Retrieved from http://www.bis.gov.uk/wp-content/uploads/publications/Mandelson-Letter-to-HEFCE-Dec09.pdf accessed 26th February 2010.

Mao, Y.-q., Y. Du and J.-j. Liu (2009). 'The effects of university mergers in China since 1990s.' *International Journal of Educational Management* 23(1): 19-33.

McMillan, M. L. and W. H. Chan (2006). 'University efficiency: A comparison and consolidation of results from stochastic and non-stochastic methods.' *Education Economics* 14(1): 1-30.

O'Donnell, C. J. (2011). 'Econometric estimation of distance functions and associated measures of productivity and efficiency change.' *Centre for Efficiency and Productivity Analysis Working Paper Series* WP01/2011. http://www.uq.edu.au/economics/cepa/docs/WP/WP012011.pdf University of Queensland.

Paul, C. J. M., W. E. Johnston and G. A. G. Frengley (2000). 'Efficiency in New Zealand sheep and beef farming: the impacts of regulatory reform.' *Review of Economics and Statistics* 82(2): 325-337.

Paul, C. J. M. and R. Nehring (2005). 'Product diversification, production systems, and economic performance in US agricultural production.' *Journal of Econometrics* 126: 525-548.

Shephard, R. W. (1970). *Theory of Cost and Production Frontiers*. Princeton, NJ, Princeton University Press.

Simar, L. and P. W. Wilson (2008). 'Statistical inference in nonparametric frontier models: Recent developments and perspectives'. *The Measurement of Productive Efficiency*. H. O. Fried, C. A. K. Lovell and S. S. Schmidt (ed). Oxford, Oxford University Press.

Skodvin, O.-J. (1999). 'Mergers in higher education - success or failure?' *Tertiary Education and Management* 5(1): 65-80.

Uri, N. D. (2003). 'The adoption of incentive regulation and its effect on technical efficiency in telecommunications in the United States.' *International Journal of Production Economics* 86: 21-34.

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

Zhang, Y. and R. Bartels (1998). 'The effect of sample size on the mean efficiency in DEA with an application to electricity distribution in Australia, Sweden and New Zealand.' *Journal of Productivity Analysis* 9(3): 187-204.

Figure 1: Efficiency score and associated 95% confidence interval by HEI

a) SFA TVD estimation

SFA TVD efficiency by HEI 2008/09

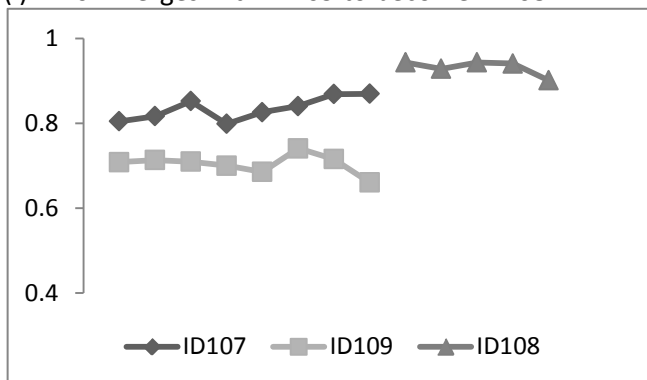**b) DEA VRS (pooled data)**

DEA VRS efficiency by HEI 2008/09
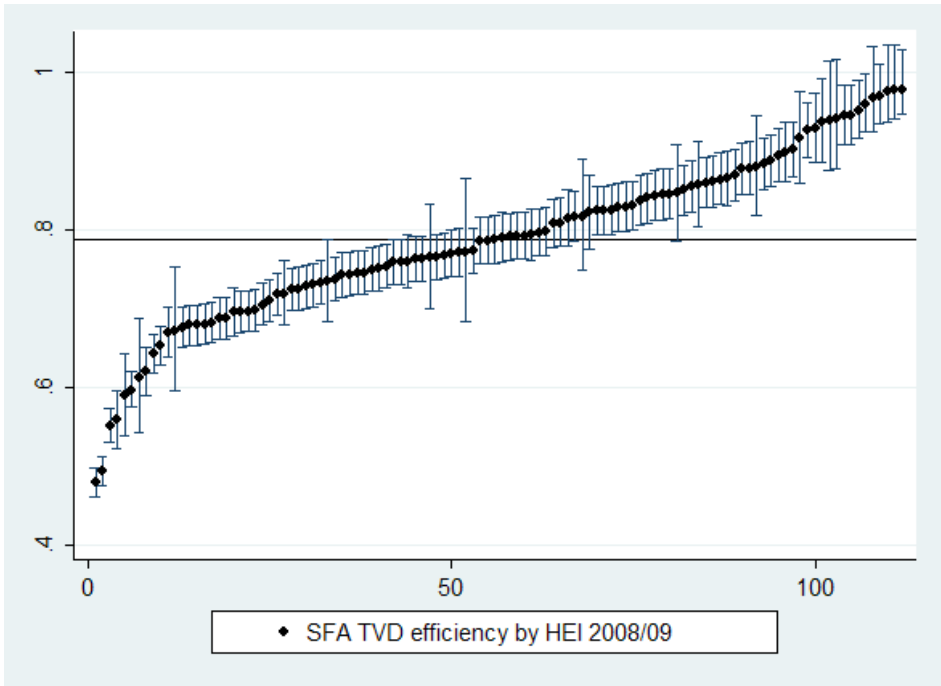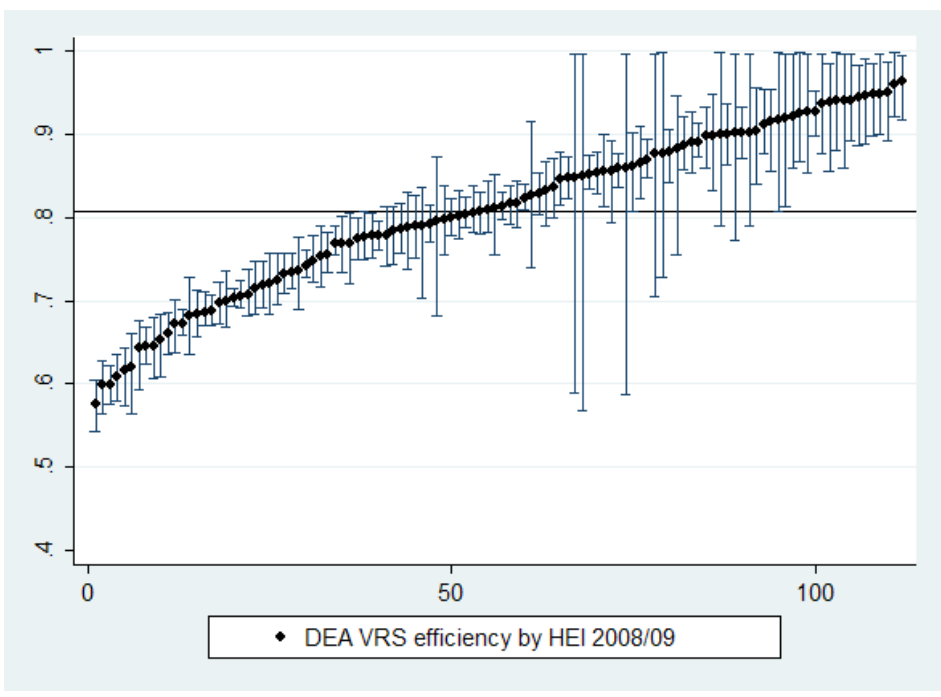
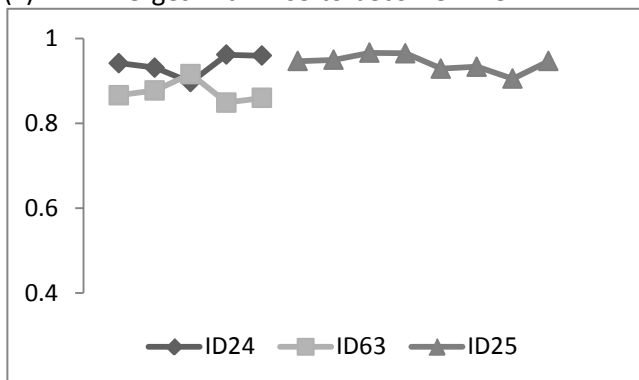**Note:** Horizontal line represents mean efficiency.

Figure 2: Patterns of efficiency (pooled DEA VRS pooled) in merging institutions over time (1996/97 to 2008/09)

a) Mergers of pre-1992 HEIs with pre-1992 HEIs
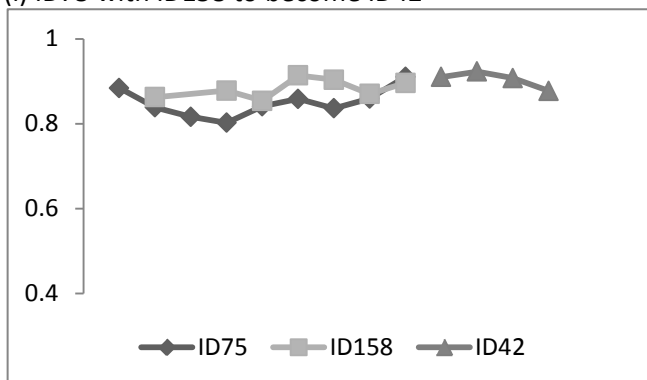
(i) ID107 merges with ID109 to become ID108

(ii) ID24 merges with ID63 to become ID25

b) Mergers of former colleges of higher education with former colleges of higher education

(i) ID75 with ID158 to become ID42

(ii) ID94 with ID173 to become ID95
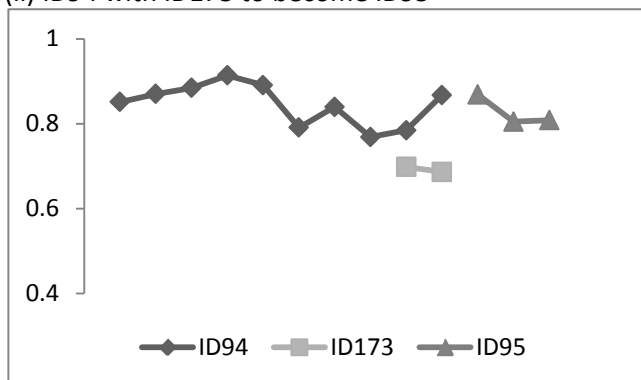
(iii) ID45 with ID55 to become ID182

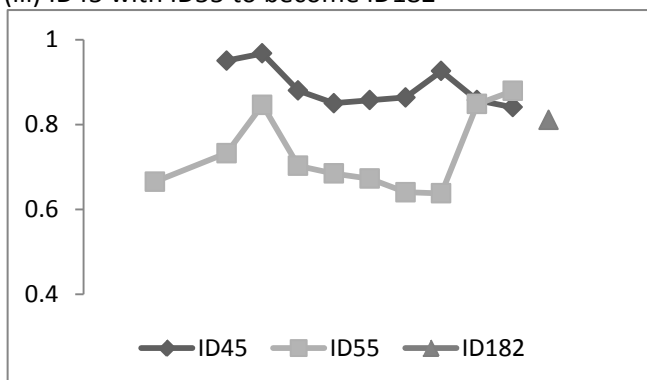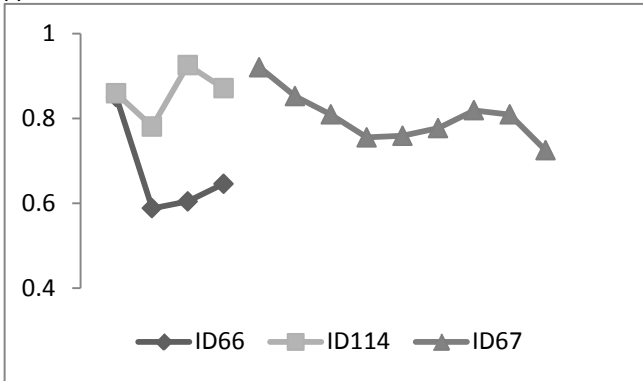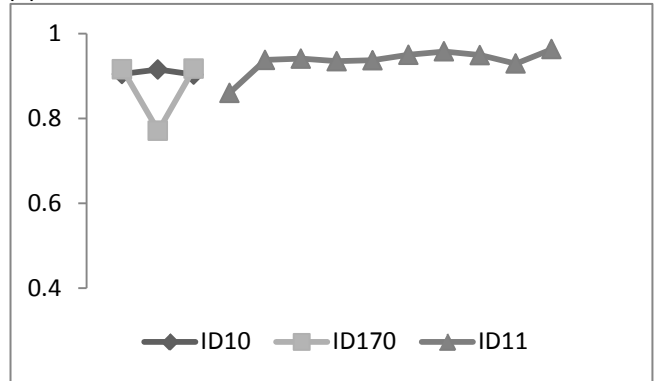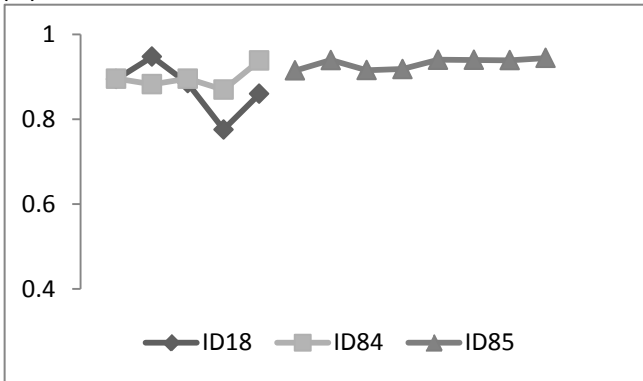**c) Mergers of former colleges of higher education with pre-1992 HEI**

(i) ID66 with ID114 to become ID67

(ii) ID10 with ID170 to become ID11

(iii) ID18 with ID84 to become ID85

d) Mergers of former colleges of higher education with post-1992 HEI

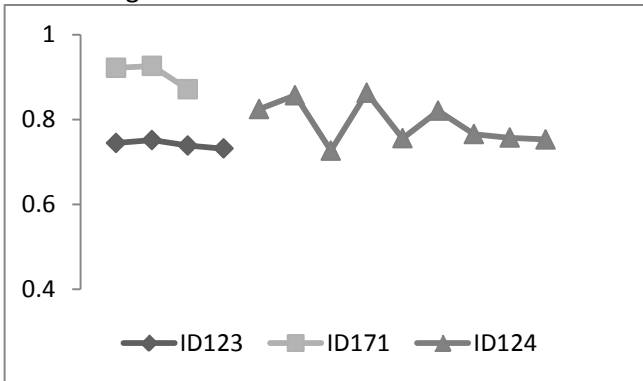ID123 merges with ID171 to become ID124

Table 1: Definitions of inputs and outputs

Variable name	Definition
Inputs:	
PGINPUT (x_1)	The total number of FTE postgraduate students (i.e. students on programmes of study leading to higher degrees, diplomas and certificates, including Postgraduate Certificate of Education (PGCE) and professional qualifications)
UGINPUT	The total number of FTE first degree and other undergraduates. The 'other undergraduates' category includes qualification aims below degree level such as Foundation Degrees and Higher National Diploma (HND)[2]
QUAL	1996/97-2001/02: Average A-level points of full-time, first year, first degree students at English HEIs converted to the tariff scale using an appropriate transformation reflecting the relationship between points and tariff; 2002/03-2008/09: Average A-level tariff scores of full-time, first year, first degree students at English HEIs
UGINQUAL (x_2)	UGINPUT*QUAL/mean(QUAL)
STAFF (x_3)	The number of full-time academic staff plus 0.5 times the number of part-time academic staff
ACSERV[1] (x_4)	Expenditure incurred on centralised academic services such as the library and learning resource centres, central computer and computer networks, centrally run museums, galleries and observatories, and any other general academic services (in £000s)
ADMIN[1] (x_5)	Expenditure on total administration and central services including expenditure on staff and student facilities (including, for example, Careers Advisory Service, all grants to student societies, emoluments to wardens of halls of residence, accommodation office, athletic and sporting facilities, excluding maintenance, and the institution's health service) and general educational expenditure (in £000s)
Outputs	
PGOUTPUT (y_1)	The number of higher degrees plus total other postgraduate qualifications awarded (including doctorate, other higher degrees, PGCEs and other postgraduate qualifications)
RESEARCH[1] (y_2)	Income received in funding council grants plus income received in research grants and contracts (in £000s)
UGOUTPUT	The number of first degree and other undergraduate degrees awarded (see definition of UG)
UGOUTQUAL (y_3)	UGOUTPUT *WEIGHT/mean(WEIGHT) where WEIGHT is: first class=30; upper second class=25; lower second class=20; third class=15; unclassified=10; other undergraduate qualification=5.

Notes:

1. These variables are deflated to July 2008 values using the higher education pay and prices index (http://www.universitiesuk.ac.uk/statistics/heppi/default.asp).

2. A full description of students included in these categories can be found in the HESA data documentation.

Degree classification weightings are: first class=30; upper second class=25; lower second class=20; third class=15; unclassified=10; other undergraduate qualification=5

3. Source of entry score data: HESA Student Record 1996/97-2008/09. Copyright Higher Education Statistics Agency Limited 2010. HESA cannot accept responsibility for any inferences or conclusions derived from the data by third parties.

4. Source of all other data: HESA , *Resources of Higher Education Institutions* Cheltenham, 1994/95 to 2008/09; HESA *Students in Higher Education Institutions* Cheltenham, 1994/95 to 2008/09.

Table 2: Comparison of parametric and non-parametric models

Model	Allows for efficiency change over time	Allows for technology change over time	Imposes CRS	Allows for stochastic error	Applies the same parameters to all observations
Parametric:					
a) RE	No	Yes	No	Yes	Yes
b) SFA TI	No	Yes	No	Yes	Yes
c) SFA TVD	Yes	Yes	No	Yes	Yes
Non-parametric					
ai) CRS pooled	Yes	No	Yes	No	No
aii) CRS within year	Yes	Yes	Yes	No	No
bi) VRS pooled	Yes	No	No	No	No
bii) VRS within year	Yes	Yes	No	No	No

Table 3: Descriptive statistics

	1996/97-2008/09 N=1444		2002/03-2008/09 N=779	
Variable	Mean	SD	Mean	SD
Outputs:				
PGOUTPUT	1096.63	913.41	1304.66	1016.85
UGOUTPUT	2673.78	1850.53	3010.73	2034.87
UGOUTQUAL	2765.97	1834.94	3049.30	1984.20
RESEARCH	72192.86	79655.38	79027.70	88492.67
Inputs:				
PGINPUT	1919.13	1517.81	2129.37	1657.73
UGINPUT	7219.74	4749.82	7759.33	5086.79
QUAL	285.92	81.15	302.11	79.22
UGINQUAL	7437.14	5852.08	8453.66	6506.89
STAFF	905.91	867.01	945.57	874.12
ACSERV	9757.03	8539.87	11233.54	9424.72
ADMIN	16827.69	12617.76	19602.08	14297.15

Notes:

1. See Table 1 for variable definitions.

Table 4: Descriptive statistics for efficiencies from parametric models for bias corrected efficiencies from DEA models

	Model 3: 1996/97 – 2008/09				Model 4: 2002/03 – 2008/09			
Parametric models	**mean**	**SD**	**Min**	**Max**	**mean**	**SD**	**Min**	**Max**
TI RE	0.747	0.073	0.520	1.000	0.715	0.077	0.528	1.000
TI SFA	0.803	0.097	0.515	0.987	0.780	0.108	0.439	0.984
TVD SFA	0.801	0.097	0.479	0.990	0.508	0.070	0.309	0.920
DEA efficiency models	**mean**	**SD**	**Min**	**Max**	**mean**	**SD**	**Min**	**Max**
a) CRS i) pooled	0.753	0.090	0.341	0.978	0.762	0.095	0.463	0.966
a) CRS ii) within year	0.869	0.083	0.422	0.984	0.862	0.087	0.553	0.980
b) VRS i) pooled	0.827	0.089	0.429	1.271	0.843	0.089	0.523	1.261
b) VRS ii) within year	0.932	0.067	0.547	1.142	0.926	0.069	0.643	1.018

Note:
See section 4 and Table 1 for definitions of the models.

Table 5: Mean DEA efficiency by HEI type

Model		Model 3				Model 4		
		Post-1992 HEIs	**Pre-1992 HEIs**	**Colleges of higher education**		**Post-1992 HEIs**	**Pre-1992 HEIs**	**Colleges of higher education**
		n = 375	n = 624	n = 445		n = 203	n = 336	n = 240
DEA	**a)CRS i) pooled**	0.780	0.726	0.767		0.801	0.727	0.777
	a) CRS ii) grouped	0.887	0.841	0.893		0.888	0.834	0.881
	b) VRS i) pooled	0.849	0.821	0.817		0.872	0.838	0.827
	b) VRS ii) grouped	0.947	0.919	0.939		0.947	0.911	0.928

Note: In the case of each set of efficiencies, the null hypothesis that the means are identical across all groups is rejected at the 5% significance level.

Table 6: Mean efficiency by merger activity

Models	Models 1 and 3			Models 2 and 4		
	1996/97 to 2008/09			2002/03 to 2008/09		
	Pre-merger n = 142	Post-merger n = 133	Non-merging n = 1169	Pre-merger n = 35	Post-merger n = 99	Non-merging n = 645
Parametric						
a) TI RE	0.744	0.794	0.742	0.710	0.753	0.710
b) TI SFA	0.797	0.890	0.793	0.771	0.872	0.766
c) TVD SFA	0.806	0.882	0.791	0.509	0.563	0.500
Non-parametric						
a)CRS i) pooled	0.764	0.780	0.748	0.777	0.793	0.756
a) CRS ii) grouped	0.873	0.899	0.865	0.861	0.892	0.858
b) VRS i) pooled	0.833	0.881	0.820	0.856	0.882	0.837
b) VRS ii) grouped	0.943	0.954	0.929	0.934	0.947	0.922