Lancaster University Management School
Working Paper
2005/042

An analysis of costs in institutions of higher education in England

Geraint Johnes, Jill Johnes and Emmanuel Thanassoulis

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

AN ANALYSIS OF COSTS IN INSTITUTIONS OF HIGHER EDUCATION IN ENGLAND

Geraint Johnes[1]

Jill Johnes[1]

Emmanuel Thanassoulis[2]

May 2005

Word count: 9675

ABSTRACT

Cost functions are estimated, using both random effects and stochastic frontier methods, for institutions of higher education in England. The paper advances on the existing literature by employing finer disaggregation by subject, institution type, and location, and by introducing consideration of quality effects. The findings are that, amongst undergraduates, medical students are the most costly, and non-science students the least; amongst postgraduates, those on taught courses are costly, while research students are relatively inexpensive. Provision in London is found to be more costly than that elsewhere. Estimates of economies of scale and economies of scope vary according to the choice of estimating technique. The random effects model suggests that ray economies of scale and economies of scope are ubiquitous. The stochastic frontier model suggests some product-specific economies of scale in research, but diseconomies elsewhere, and product specific economies of scope in undergraduate science, but diseconomies elsewhere. This has implications for achieving any expansion in higher education.

Keywords: higher education, cost functions
JEL Classification: I20

1. Lancaster University Management School, Lancaster LA1 4YX, UK
2. Aston Business School, Aston University, Birmingham B4 7ET, UK

An earlier version of the paper was presented at the CRISS conference in Siena, 25-26 September 2004. The authors thank participants at that conference for useful comments, and are grateful also to Pam Lenton for her work in setting up the database used to produce the results reported herein.

1. Introduction

During the last three decades, the UK higher education sector has been under pressure to provide its services as efficiently as possible, whilst undergoing huge changes in its size and structure. In 1992, former polytechnics were granted the status of universities. Since then, in the period from 1996 to 2003, total student numbers in the UK higher education sector have increased by 33.5%, income from research grants and contracts has increased by 67.1% and expenditure has grown by 45.7%. Despite the drive for efficiency, little detailed information is available about the structure of costs in the UK higher education sector, yet, in an environment of expanding output and increasing costs, the importance of such a knowledge can surely not be overstated[1].

Any efficient expansion of output requires a knowledge of marginal cost, average cost, and economies of scale and scope. Institutions of higher education (IHEs) are multi-product firms and are generally agreed to produce two main outputs, namely teaching and research (Cohn & Cooper 2004). An additional output is known as the *third leg* output and encompasses, *inter alia*, the provision of advice and other services to business, the storage and preservation of knowledge, and the provision of a source of independent comment on public issues (Verry & Layard 1975).

Any analysis of costs in higher education must therefore acknowledge and explicitly take into account in the estimation technique the multi-product nature of production. The first attempt to do so occurred only 15 years ago (Cohn *et al* 1989), and more recently still in analyses of costs in UK higher education (Glass *et al* 1995a; 1995b). Only limited progress has been made in the intervening years. Issues which still need to be resolved include the definition of the multiple outputs of IHEs, identification of inputs to and exogenous factors influencing the higher education production process, the specification of an appropriate statistical function which allows for both economies of scale and scope, and the choice of estimating technique.

The purpose of this paper is therefore to resolve these issues in order to estimate an up-to-date cost function of the English higher education sector[2]. Such a function

[1] Concerned by this, the funding councils have devoted much energy to the development of a Transparent Approach to Costing (TRAC) and to the full Economic Costing (fEC) of research projects. Thus work remains incomplete, and, while not uninstructive, suffers an inherent flaw in that it is an accounting based exercise which cannot accommodate available economies through improved efficiency, scale or scope.

[2] We focus on England in order to avoid problems arising from differences across the constituent countries of the UK in both funding mechanisms and the structure of the education system.

can then be used to establish whether there are economies of scale or scope in English higher education institutions, in an effort to identify the best way of achieving expansion in the sector. In addition, the effect on costs of institution type, location and quality, areas of research which have been largely ignored in earlier analyses of costs, will also be examined in the subsequent analysis.

The paper is in 5 sections of which this is the first. Section 2 considers the methodological issues, and a short review of the literature on costs in UK higher education is provided in section 3. An empirical analysis of costs in more than 120 IHEs in England over a 3 year period is presented in section 4 and conclusions are drawn in section 5.

2. Methodology

Before proceeding to conduct an empirical analysis, there are three methodological questions to address: (i) what is an appropriate functional form for the cost equation? (ii) how can economies of scale and of scope be quantified in a multi-product context? and (iii) what is the appropriate estimating technique to use?

2.1 Functional Form

A cost function is an equation that allows costs (C) to be evaluated, given information about the level of output being produced by an organisation and information about the price (or quality) of the organisation's inputs. This is written for IHE k as:

$$C_k = f\left(y_{ik}, w_{lk}\right)$$

where y_{ik} = output i of IHE k ($i = 1,\ldots n$; $k = 1, \ldots ,K$) and w_{lk} is the price of input l ($l = 1,\ldots,m$) used in IHE k.

The first cost functions estimated for UK higher education were simple linear functions (Verry & Layard 1975), but these were restrictive and could not accommodate the possibility of varying levels of economies of scale and varying levels of output. A linear cost function is also unable to model economies of scope (or synergy) that are due to joint production. Thus more sophisticated cost functions must be hypothesised, and they should be capable of:

(i) explaining how economies of scale can occur for some output profiles, yet diseconomies of scale can occur for other output profiles.

(ii) explaining how economies of scope can occur for some output profiles, yet diseconomies of scope can occur for other output profiles.

(iii) ensuring that estimates of costs are sensible under conditions where a firm produces positive quantities of some output types, but zero amounts of other outputs; for example, the cost function should predict sensible levels of costs for an institution that does not engage in research. This rules out a lot of simple (and commonly used) cost functions that involve logarithmic transformations.

Baumol *et al*. (1982) have suggested three forms for the cost function that allow the three requirements above to be satisfied. These are:

(a) the constant elasticity of substitution (or CES) cost function

$$C_k = \alpha_0 + \left[\sum_i \beta_i y_{ik}^{\delta_i} \right]^{\rho} + v_k$$

where y_{ik} is as above, α_0, δ_i and ρ are parameters to be estimated, and v_k is an error term.

(b) the flexible quadratic cost function

$$C_k = a_0 + \sum_i a_i F_{ik} + \sum_i b_i y_{ik} + (1/2)\sum_i \sum_j c_{ij} y_{ik} y_{jk} + v_k$$

where y_{ik} and v_k are as above, a_0, a_i, b_i and c_{ij} are coefficients to be estimated, and F_{ik} is a dummy variable such that $F_{ik} = 1$ if output i in IHE k is positive, and zero otherwise.

(c) the hybrid translog cost function

$$\ln C_k = \alpha_0 + \sum_l \alpha_l \ln(w_{lk}) + \sum_i \beta_i [(y_{ik}^{\vartheta} - 1)/\vartheta] + \frac{1}{2}\sum_l \sum_m \gamma_{lm} \ln w_l \ln w_m +$$
$$\frac{1}{2}\sum_i \sum_j \delta_{ij} [(y_{ik}^{\vartheta} - 1)/\vartheta][(y_{jk}^{\vartheta} - 1)/\vartheta] + \sum_l \sum_i \rho_{li} \ln w_l [(y_i^{\vartheta} - 1)/\vartheta] + v_k$$

where y_{ik}, w_{lk} and v_k are as above, and the Greek letters are all parameters to be estimated. The hybrid translog does not contain terms in the logarithm of y variables, and, so long as the estimated value of θ is not precisely zero, it is therefore possible to use this functional form to evaluate costs when some y are zero. Clearly the hybrid translog function is highly non-linear and (in common with the CES, but not the quadratic) must be estimated by maximum likelihood rather than by conventional least squares methods.

In the subsequent analysis, we shall use the quadratic cost function to examine the structure of costs in English IHEs; the choice of the quadratic function is governed by the desirability of a simple specification: a simple specification is particularly desirable when using stochastic frontier methods to estimate the parameters of the model, since standard software packages do not allow such methods to be applied to highly non-linear models. [3]

2.2 Economies of Scale and Scope

2.2.1 Ray economies of scale

Ray economies (or diseconomies) of scale are defined in the multi-product case as the cost savings (or dissavings) arising when the size of the aggregate output expands but the composition of output (i.e. output mix) remains constant. The size of the ray economies of scale (S_R) is calculated in the general case as:

$$S_R = \frac{C(y)}{\sum_i y_i C_i(y)}$$

Where $C(y)$ is the cost of producing the output vector y and $C_i(y)$ is the marginal cost of producing the *i*th output so that $C_i(y) = \partial C(y)/\partial y_i = MC_i$. If $S_R > 1$ ($S_R < 1$) then there are ray economies of scale (diseconomies of scale).

2.2.2 Product-specific economies of scale

Product-specific economies (or diseconomies) of scale are the cost savings (or dissavings) which occur when the level of one product increases while the levels of the rest of the outputs remain fixed. The incremental cost of producing output *i* ($IC(y_i)$) is defined as:

$$IC(y_i) = C(y_n) - C(y_{n-i})$$

where $C(y_n)$ is the total cost of producing all the outputs at the levels in y_n, while $C(y_{n-i})$ is the total cost of producing all the outputs at the levels in y_n *except output i* which is zero. The average incremental cost of product *i* is then defined in the general case as:

$$AIC(y_i) = \left[C(y_n) - C(y_{n-i})\right]/y_i = IC(y_i)/y_i$$

If the average incremental cost of product i exceeds its marginal cost then we have *product-specific returns to scale* for product i. Thus, product-specific returns to scale for product *i* ($S_i(y)$) are:

$$S_i(y) = AIC(y_i)/C_i(y)$$

[3] An example of the application of these methods to the CES cost function is, however, given by Izadi *et al*. (2002). We eschew the use of the CES function here in light of the adverse comments made on this form by Johnes (2004).

If $S_i(y) > 1$ ($S_i(y) < 1$) then there are product-specific economies (diseconomies) of scale for product *i*. The size of product specific returns to scale is of particular relevance if there is a desire to increase the level of one of the products (for example undergraduate teaching) whilst holding the level of the other products constant. Clearly both ray and product-specific returns to scale would vary in general with the mix and levels of the outputs of each IHE.

2.2.3 Global economies of scope

Economies of scope measure the cost savings (or otherwise) arising from producing two or more products jointly in a multi-product firm rather than in a firm specializing in the production of one output. In higher education, for example, two types of economies of scope can arise: the economies from the production of all the outputs (eg teaching, research and third mission) using shared inputs, and the economies from the production of different disciplines using shared inputs.

Global economies of scope arise if the cost of producing all outputs together in one firm is less than the cost of producing each output in a separate firm. Thus if $C(y)$ is the cost of producing all *n* outputs jointly at the levels in *y*, and $C(y_i)$ is the cost of producing the *i*th output in a specialised firm at the same level as in *y* then if

$$C(y) < \sum_i C(y_i)$$

we have global economies of scope. The degree of global economies of scope is measured by S_G where

$$S_G = \left[\sum_i C(y_i) - C(y) \right] \Big/ C(y)$$

If $S_G > 0$ (< 0) then global economies (diseconomies) of scope exist for producing the outputs jointly rather than in separate firms.

2.2.4 Product-specific economies of scope

A measure of product-specific economies of scope (SC_i) is given by:

$$SC_i = [C(y_i) + C(y_{n-i}) - C(y)] / C(y)$$

where notation is as above and $C(y_{n-i})$ is the cost of producing all outputs jointly at their levels in y, except the *i*th one. If $SC_i > 0$ then there are complementarities from producing output *i* with the other outputs. If $SC_i < 0$ then the converse is the case. This measure is of particular relevance in the discussion of how specialist IHEs should be. Should undergraduate degrees, for example, be produced jointly with all other university outputs, or are there advantages of specialisation by type of output? Alternatively, would there be advantages from having IHEs specialise in a particular subject area? Clearly both global and product-specific economies of scope would vary with the levels of outputs in *y*.

2.3 Estimation Methodology

The final methodological issue which must be addressed concerns the means of estimation. In the present case we have observations on *k* IHEs over *t* time periods, and so an appropriate technique for panel data estimation should be employed. The main focus of modelling panel data is how to model the heterogeneity across observations (here, IHEs).

The simplest solution is to adopt a fixed effects approach which allows the unobserved individual effects to be correlated with the included explanatory variables. Modelled simply, the fixed effects model shifts the regression function for each unit in the set. One problem with this is that the model can only apply to observations in the sample for which the intercept has been estimated; for units outside the sample (e.g. proposed new units) it is impossible to predict the intercept and hence predict the outcome for such units. Another difficulty, particularly relevant in the case of the present dataset, is that where the time dimension of the panel is short, most of the variation in the dependent and independent variables is across observations; introduction of fixed effects can then introduce severe multicollinearity and diminish the precision of the coefficient estimates.

Alternatively, if it is assumed that the individual effects are strictly uncorrelated with the explanatory variables, the individual-specific intercept terms can be modelled as randomly distributed over all units, i.e. a random effects (RE) approach[4]. The advantage of this is that the number of parameters to be estimated is less than for a fixed effects model. However, if the initial assumption is inappropriate, then estimates may be inconsistent. For IHE *k* in time period *t* the model is represented by

$$C_{kt} = y_{kt}^{'}\beta + (\alpha + u_k) + \varepsilon_{kt}$$

where:

C_{kt} is the observation of the dependent variable for the *k*th IHE in the *t*th time period;

y_{kt} is the matrix of *n* explanatory variables (not including a constant);

[4] See Scheffe (1956), Wallace and Hussain (1969) and Nerlove (1971).

α is the intercept term and denotes the mean of the unobserved heterogeneity;

$u_k \sim IID(0,\sigma_\alpha^2)$ and is the random heterogeneity specific to the *k*th observation and is constant over time; (The assumption of constancy over time can be relaxed at the expense of setting u_k a pre-specified function of time, a refinement we have not pursued in this paper.)

$\varepsilon_{kt} \sim IID(0,\sigma_\varepsilon^2)$ and is uncorrelated over time.

By assumption, u_k and ε_{kt} are mutually independent and are independent of x_{kt} $\forall k,t$, and so the OLS estimators for α and β are unbiased and consistent. However, since the composite error term $u_k + \varepsilon_{kt}$ exhibits a pattern of autocorrelation (unless $\sigma_\alpha^2 = 0$), generalized least squares (GLS) is used to estimated the RE model.

It is open to argument, however, whether this method, which plots a 'best fit' function to the data, is appropriate in the cost function context involved here. A more appropriate method may be that of stochastic frontier analysis (SFA) (Aigner and Chu, 1968), which estimates a frontier around the data. SFA works by disaggregating the residual term into a Gaussian component (usually attributed to measurement error) and a non-Gaussian component (often distributed half-normally or exponentially and attributed to efficiency variations across the units of observation). Moreover, SFA can be adapted to the context of panel data following the work of Batesse and Coelli (1995) and Greene (2002). An interesting by-product of the SFA estimation method, following the contribution of Jondrow *et al*. (1982), is that estimates of the technical efficiency may be obtained for each unit of observation.

3. Literature review

Numerous studies of the costs of higher education have been conducted in a variety of countries. In this section, we report largely on those that have concerned the UK. The first reported study is an exception, and is included because it is the earliest study of university costs that is based on modern understanding of multi-product organisations (Cohn *et al* 1989). These authors use a flexible fixed quadratic function and a cross section of 1887 IHEs to relate total education transfers and expenditures in 1981-82 to three outputs: FTE undergraduate enrolment; FTE postgraduate enrolment; and grants received by the IHE for research. The problem of the possibility of different objectives across IHEs is addressed by splitting the sample into public and private institutions, and performing the analysis separately for each subsample.

Variations in labour input prices are controlled for by the inclusion of average faculty salary and its square (which are found to be insignificant).

The results indicate that there are ray economies of scale up to the mean output level in the public sector and up to 6 times the mean output level in the private sector. Product-specific economies of scale, however, are observed only in the public sector, and only for postgraduate teaching and research output (although economies are also observed in undergraduate teaching in IHEs producing low levels of output). Economies of scope are found in both the public and the private sectors of higher education.

Several studies of costs in the UK higher education sector have also employed a quadratic functional form for the cost function (Johnes 1996; 1998). These studies differ from the Cohn *et al* (1989) study, however, because each of the outputs (undergraduate student load, postgraduate student load, and value of research grants) is split by broad subject category, namely arts and science. In addition, the latter of the studies, which are based on data for 50 UK universities in the years 1989-90 and 1990-91, uses SFA as well as least squares to estimate the cost function. Ray economies of scale and product-specific economies of scale for science postgraduate teaching and for science research output are found using both estimating methods. Global economies of scope are observed when least squares is used but not when SFA is used to estimate the quadratic cost function reinforcing the importance of eliminating inefficiency before addressing the issue of economies of scope or scale.

The broad subject category split for the teaching outputs permits a closer examination and comparison of average incremental costs (AICs). AICs in arts undergraduate teaching are less than those for science undergraduate teaching with the exception of IHEs which are former Colleges of Advanced Technology (CATs), where the opposite is the case. This is probably because the former CATs exploit economies of scale in science but not in arts. Turning to postgraduate teaching, the AIC of science postgraduate teaching exceeds that for science undergraduate teaching in small universities, but the two are virtually identical for a typical university with average levels of output. In the case of arts, postgraduate teaching has a lower AIC than undergraduate teaching in the typical university.

A CES functional form has been used in a number of studies of UK higher education (Johnes 1997; 1999; Izadi *et al* 2002). All three of these studies use data for 1994/95 (thereby including both pre- and post-1992 universities in the cost function for the first time), and disaggregate only undergraduate students into arts and science categories. The later two studies (Johnes 1999; Izadi *et al* 2002) differ from the first as they estimate a frontier cost function using SFA. AICs of postgraduate teaching are found to exceed those for undergraduate teaching, and AICs of science undergraduate teaching exceed those of arts undergraduate teaching. Ray economies of scale are close to unity but product-specific economies of scale are observed for arts undergraduate teaching, postgraduate teaching and research. There are no economies of scale for science undergraduate teaching. Economies of scope are observed nowhere.

The translog functional form has been used in several studies of the UK higher education sector (Glass *et al* 1995a; 1995b; Stevens 2005). The early studies use data for periods when the binary divide was still in existence, and do not disaggregate the three main outputs by subject. These studies confirm the existence of ray economies of scale, and product-specific economies are consistently observed for undergraduate teaching (Glass *et al* 1995a; 1995b). There is no evidence of global economies of scope, and product-specific economies of scope are observed only for postgraduate teaching, while diseconomies of scope are identified in undergraduate teaching (Glass *et al* 1995a).

More recent data (academic years 1995/96, 1996/97, 1997/98 and 1998/99) form the basis of analysis in the Stevens (2005) study, which differs further from the Glass *et al* studies (1995a; 1995b) because SFA is used to estimate the translog cost function. The only output to be disaggregated by subject is undergraduate teaching. This study is particularly remarkable for its attempt to control for quality by the inclusion of two variables namely A level/Scottish Highers score of the entry cohort, and the percentage of firsts and upper seconds achieved. Interpretation of the results regarding these variables is difficult because of the multicollinearity between the variables[5] but there seems to be some evidence that variables reflecting quality of output are important in determining cost efficiency, suggesting that this issue would be worth pursuing in future research.

In general, therefore, the literature on costs in UK higher education is limited in that institutions are treated as a homogeneous group. Potential differences between location of and types of IHEs have not been investigated (compare this with Cohn *et*

[5] The A level variable is positive until the degree variable is included, at which point the coefficient on A level becomes negative and the coefficient on the degree variable is positive.

al 1989 where differences between different types of universities are found). In addition, the disaggregation of outputs by subject is limited to just two types, science and arts, yet within the former group, for example, there are likely to be cost differences between clinical and laboratory-based teaching. Furthermore, only three outputs (ignoring any possible subject disaggregation) are considered: undergraduate and postgraduate teaching and research. The third mission activities of IHEs are not included as a determinant of costs in any study. Also, teaching outputs vary not just by subject, but also by type of qualification, and potential cost differences arising from these aspects have been ignored. Finally, most studies have not exploited the data to the full by using panel methods and frontier analysis. Our aim in the present paper is therefore to improve on the received literature along these dimensions.

4. Analysis

The sample of institutions included in the analysis comprises all IHEs in England. This sample therefore includes ancient universities, such as Oxford and Cambridge, traditional universities (in the pre-1992 sector), new universities (mainly former polytechnics that were granted university status in 1992), and colleges of higher education. Universities are represented nationally by Universities UK, while the colleges of higher education are represented by the Standing Conference of Principals (SCOP).

A panel of data has been collected across three years, 2000-01 through 2002-03. The data include information about total operating costs (net of residence and catering costs) measured in December 2002 values), undergraduate and postgraduate student load by subject area, research activity, third leg activities, degree results, and the quality of the student intake for each institution in each year (precise definitions of variables are provided in Table 1). The data have all been provided by the Higher Education Statistics Agency (HESA). Descriptive statistics for the sample data can be found in Table 2.

<Table 1 here>

<Table 2 here>

Average operating costs in the institutions in our sample amount to about £86m per year. The typical institution has just over 6000 undergraduates and around 1700 postgraduates. The medical subject group accounts for only around 200 of the undergraduates (on average - though most institutions do not provide these subjects at all, and the mean for those institutions that do provide them is therefore much higher at 1395), while other science accounts for 2600 of the undergraduates, and non-science is the largest group at 3400 students.

The descriptive statistics for the entire group of IHEs conceals some considerable variations between IHE type, and these are revealed to some extent in the lower part of Table 2. Thus, post-1992 institutions have the largest number of undergraduates on average at over 10000, compared with over 6000 in traditional IHEs and only 2000 in SCOP colleges. Postgraduate numbers are fairly evenly distributed between traditional and post-1992 IHEs (at a mean level of just under 2500), but are much lower in SCOP colleges (at 450). Research activity is heavily concentrated in traditional institutions which have, on average, more than 10 times and more than 100 times the research income of post-1992 institutions and of SCOP colleges, respectively.

There is likely to be diversity within these specified groups because of the historical development of the institutions. Some institutions within the traditional university sector, for example, have developed from colleges of advanced technology, and, as such, the subject mix that is provided by these institutions is heavily skewed towards the sciences. Others of the traditional universities, often but not always the so-called 'civics', have - in view of their presence in large cities - developed substantial medical schools. In addition, while Table 2 reveals that the post-1992 sector of higher education has a lower level of research activity, on average, compared to the traditional institutions, some of the IHEs in this sector are competing in this domain with some of the traditional universities. Finally, aggregate student numbers conceal the variety of qualifications provided in many former polytechnics, and in the colleges of higher education. Indeed, many of these institutions provide not only degree level education, but also a considerable amount of education leading to qualifications below that of a bachelor degree, and at postgraduate level, these institutions provide specialist training in vocational areas such as teaching.

It is therefore clear that there is considerable diversity across IHEs in the English higher education sector. This suggests that there may also be differences between IHEs in the way in which costs are determined, for example, higher education colleges may well have different cost functions to those that attach to Oxford and Cambridge. We therefore propose to analyse separately three groups of institutions: SCOP colleges; new universities; and traditional universities, the definition of the groups arising from the obvious distinctions between these groups highlighted in Table 2. While there are admittedly also differences within each group, it is impossible to disaggregate the data further for the analysis without running into degrees of freedom problems.

4.1 Estimates of Costs

The RE, and SFA estimates of the quadratic multi-product cost function, applied to data from the full sample of institutions, are reported in Table 3. The quadratic cost function estimated here includes interaction and quadratic terms involving student numbers of all types and research but not for third mission activities in order to preserve degrees of freedom. A test of the null hypothesis that the coefficients of all the interaction and quadratic terms are zero indicates that the interaction and quadratic terms are jointly significant[6].

<Table 3 here>

In order to compare and interpret the results, values of AICs for a typical IHE i.e. one which has mean output levels are estimated for both the RE and SFA models (see table 4A)[7]. Since the size of AIC depends on the value assumed for the output vector, AICs are also estimated in Table 4B across small (80% of mean output) and large (120% of mean output) IHEs. This is done for the RE model only.

<Tables 4A and 4B here>

The results obtained from both models are broadly similar (see panel A of Table 4). The AIC of undergraduate medicine is varies from £17600 per annum from the SFA model to £21000 per annum from the RE model); that associated with undergraduate science is around £6,000 a year, while undergraduate non-science has an AIC of about £3,500 per year. Meanwhile, postgraduate education has an AIC which varies from £7,500 (when estimated using SFA) to about £10,500 per year (when estimated using RE). It should be noted that the SFA method predicts cost levels that that are in general below those estimated using RE, and this is as we might expect given that SFA, unlike RE, is a frontier method.

The estimated AICs are indicative rather than definitive and should be used with some caution. The AICs as noted above are computed by setting all outputs at prespecified values (e.g mean or 120% of mean levels). The further the actual outputs

---

[6] $\chi^2_{(15)} = 95.61$ for the RE model.

[7] It is recalled that the AIC of output *i* evaluated at mean output levels is the cost increase per unit of output *i* when all outputs bar *i* are set to mean levels and the level of output *i* is increased from zero to mean. AICs at other output levels are defined in an analogous manner.

at an IHE are from the levels used in the calculation the bigger the likely discrepancy in its actual AICs from those estimated in Table 4.

The costs reported in panel A of Table 4 may be compared with the Higher Education Funding Council for England (HEFCE) resource rates for the four subject price groups for undergraduates. For the year 2002/03 (which is the appropriate comparison with the figures estimated here) these were, respectively for groups A, B, C and D, £12939, £5750, £4313 and £2875. Group A refers to clinical medicine, dentistry and veterinary science; group B refers to laboratory based subjects; group C refers to subjects with a studio, laboratory or fieldwork element; and group D refers to all other subjects. Groups A and B correspond exactly to our definitions of medicine and other science respectively, while groups C and D are combined into undergraduate non-science. The pattern of AICs in these figures corresponds with the pattern observed in the statistical estimates and this raises the question of whether the statistical cost functions are simply describing a funding formula. This is unlikely to be the case for several reasons. First, while the pattern is the same, the level in the case of undergraduate medicine is considerably different. Second, the specification of the model includes several variables that are not formula funded – including postgraduate numbers, research, and third mission work. Indeed, only about 40% of universities' income comes from the funding council. The specification of the model is, moreover, non-linear, whereas formula funding is linear in nature. Third, the analysis is in line with other work of this kind – including work such as that of Cohn *et al.* (1989) which was conducted in the USA where resources are not allocated by formula. Finally, and in our view most tellingly, the use of SFA provides a safeguard against the misrepresentation of expenditures as costs since the functions estimated by SFA tell us what the parameters would be for a technically efficient institution. Nevertheless the possibility cannot be entirely dismissed that, as Bowen (1980) has argued, 'each institution raises all the money it can' and 'each institution spends all it raises'.

4.2 Estimates using sub-samples

Table 5 reports the AIC estimates that are obtained for 3 distinct groups of IHEs: SCOP colleges; post-1992 universities; and pre-1992 universities.[8] The results

[8] The stochastic frontier models for subgroups of institutions proved to be unrobust – possibly as a consequence of the small numbers in the subgroups of IHEs - and the results from this model are not therefore reported here.

highlight some differences between type of institution. Undergraduate science has an AIC of around £8000 in SCOPs and traditional IHEs compared with around £3000 in post-1992 institutions, while undergraduate non-science has an AIC of around £3000 at all types of institutions. Indeed, the AIC of undergraduate science is actually lower than that for undergraduate non-science in post-1992 IHEs. This is a remarkable result which is not observed in the other IHE types. It is likely a consequence of the fact that AICs are computed by setting one output at a time to zero, and in the case of the post-1992 sector, the data used to estimate the cost functions did not include a zero level for undergraduate science or non-science outputs (see minimum values reported in Table 2). This therefore represents an extrapolation beyond the valid domain and makes the estimated AICs for these outputs in the post-1992 sector particularly questionable.

<Table 5 here>

A further notable result is the difference between institutions in the cost of postgraduates: the AIC is around £3000 in SCOP colleges, £8000 in post-1992 IHEs and £14000 in traditional universities[9]. There are various possible explanations for these observed differences between groups of institutions in their estimated AICs. First, it may be a consequence of a technical problem with the model. If, owing perhaps to a paucity of observations, variation in one or more of the variables is limited, the quadratic specification can be particularly prone to problems of multicollinearity because variables appear not only in linear form but also in a multiplicity of interaction terms. Secondly, the inter-institutional difference in AICs may be a consequence of different mixes of outputs in the different groups of IHEs. Thirdly, the inter-institutional variation in the AIC of postgraduates may arise from the fact that the postgraduate output encompasses considerable variety in the type of qualifications obtained: for example, 1 year teacher training; 1 year taught masters; or 3 year doctorate. Clearly, the resources required for each type of qualification will vary and, if (as appears to be the case from the descriptive statistics in Table 2) different types of IHE specialise in different types of postgraduate qualification, this will give rise to the variation across institutions in AIC for postgraduate teaching. This is investigated further in section 4.6.

[9] Note that there are no observations with zero output for postgraduates in pre-1992 universities, and so the AIC for this group of IHEs in particular should be interpreted with caution since its calculation involves extrapolation outside the valid domain.

4.3 Estimates of economies of scale and scope

This section reports on measures of scale and scope economies that have been derived from the full (quadratic) specification of the model, estimated across all IHEs in the sample using both the RE and SFA estimation methods. Like AICs, these measures vary depending on the levels in the output vector used for their calculation, and so are reported for a hypothetical institution which produces average levels of all outputs identified in the model, one which produces 80% of the average level of each output, and one which produces 120% of the average level of each output. The results appear in Tables 6 and 7.

<Table 6 here>

<Table 7 here>

Three things are apparent from the estimated economies of scale. First, while there appear to be ray economies of scale at all sizes considered when the RE model is used, the opposite is the case when the SFA method of estimation is applied. Second, estimates of product-specific returns to scale are generally similar across estimating method, with the exception of postgraduate teaching (where there are constant returns using the RE model, but decreasing returns with the SFA model), and undergraduate non-science (where there are slightly decreasing returns with the RE model, but slightly increasing returns using SFA). Third, these broad conclusions remain the same across all sizes of institutions considered. On theoretical grounds the SFA estimates of returns to scale are to be preferred over those based on RE. Economies of scale can only be defined for efficient cost levels and the RE model estimates an average cost function. Thus it is possible the cost level used in the numerator of S_R (see section 2.2.1) could be over-estimating true costs by more than any overestimate of total costs through the use of marginal costs in the denominator of S_R. RE and SFA agree better on product-specific economies of scale and this in part could be because the slopes of the two functions estimated can be similar in a particular direction even if the functions are located at different levels of total cost estimated.

So far as economies of scope are concerned, there is a greater disagreement between the estimating methods than is the case for economies of scale. The RE model predicts both global and product-specific economies of scope (the former are substantial), while the SFA model predicts diseconomies of scope except in the case of undergraduate science where scope economies are observed across all sizes of institution. While, on theoretical grounds, the SFA estimates are to be preferred since we wish to disaggregate costs attributable to inefficient behaviour from those due to diseconomies of scope, the measure of scope economies could in either estimating method lead to misleading results, as we are extrapolating to zero output levels. Therefore our findings here need to be seen as indicative rather than definitive. Despite this caveat, the predictions of economies of scope (as indeed those of scale) are extremely interesting, if verified, as they have implications regarding the appropriate choice of expansion policy. These are discussed further in section 4.7.

<Table 8 here>

4.4 Estimates of efficiency

A by-product of the SFA quadratic cost function displayed in Table 3 is a set of efficiencies for each IHE for each time period. These efficiencies are calculated as the ratio of the estimated to the observed cost. It should be noted that estimates of efficiencies at the extremes can be unreliable because the estimation of the frontier can be distorted by the presence of outliers, and so the descriptive statistics displayed in Table 8 for the final year of the study (2002/03) should be interpreted with a large degree of caution. A close inspection of the efficiency values reveals two findings. First, there is a positive relationship between efficiency and size. The average size (measured by the sum of all students, undergraduate and postgraduate) of the 29 IHEs with efficiency scores below 0.60 is 1137 compared with 10450 for the remaining institutions. Second, of all these IHEs with an efficiency score below 0.60, all but two of them are specialist institutions (specialising in the teaching of a particular subject area, or a specific type of student, or specialising in teaching activities only or research activities only). Clearly further investigation into the possible determinants of cost efficiency is vital.

4.5 An augmented model

The model estimated thus far is extremely simple in that the vector of explanatory variables is made up only of the various outputs produced by IHEs. A variety of additional variables might be expected to influence costs. First, the existing vector of explanatory variables takes into account inter-institutional differences in the quantity but not the quality of output produced, nor, for that matter, the quality of input. It would be unwise to include in the cost function separate measures of students' achievements at entry and exit, since it would inevitably result in a severe problem of multicollinearity. Instead, a crude measure of value added is constructed as the ratio of average weighted degree results to average A level score for each institution[10]. A second possible influence on costs is the location of an IHE, particularly location within the London conurbation where land prices and labour costs are higher than might be expected elsewhere. The vector of explanatory variables is therefore augmented with a London dummy.

The impact of adding to the model of costs a value-added measure and a London dummy can be examined in Table 9 which reports the coefficients on the variables of interest which have been added to the quadratic model estimated using RE and SFA respectively. It should be noted that the A level score data are not available for the final year of the study, nor for some IHEs, and so models are based on a 2 year panel, and those which include the quality variable are based on a reduced sample. The impact of the quality variable is insignificant. The London dummy, however, is significant (when using RE) indicating that presence in the capital adds around £5 million a year, on average, to the costs of an institution. We would expect these additional costs to be represented in the main by overheads, and not therefore impact on AICs.

<Table 9 here>

4.6 Postgraduate education

The inter-institutional differences in the types of qualifications undertaken by postgraduates have been noted in Table 2. In fact, pre-1992 universities have a considerably higher percentage of postgraduates undertaking research (37%) than either post-1992 or SCOP colleges (which have 11% and 7% respectively), while post- and pre-1992 institutions have much higher percentages of taught postgraduates (at 55% and 47% respectively) than SCOP colleges (at 26%). Of the relatively small number of postgraduates in SCOP colleges, 67% are in the 'other' category which is made up of, *inter alia*, teacher training qualifications. In the preceding analysis lack of degrees of freedom determined that only a single measure of postgraduate activity was included on the right hand side, despite this variation.

---

[10] In constructing such a variable, the weights used for the degree classes are first = 30, upper second = 25, lower second = 20, third = 15 and unclassified = 10. An alternative weighting (first=80, upper second=65, lower second=55, third=45, unclassified=37.5) was used to give a greater spread but produced almost identical results to those reported in the tables. Data on intake quality are available only for the first two years of our panel, and so Table 9 uses data only for these periods.

[12] A linear rather than quadratic function is estimated to preserve degrees of freedom and avoid multicollinearity.

A linear cost equation[12], which includes all of the earlier defined outputs but splits postgraduates into three categories (namely research, taught and other), estimated using RE and SFA, respectively, is displayed in Table 10. The pattern of undergraduate AICs is as reported in Table 4. The AIC for postgraduate tuition, however, is around £11,000 to £14,000, and is relatively high compared to the AICs for postgraduate research (£2500 to £3500) and other postgraduates (£3000 to £4000). The result regarding postgraduate research is surprising given the intensive supervision on a one-to-one basis often provided for research students. The result is possibly a consequence of the fact that research postgraduates often provide input into the undergraduate teaching and research functions of the institution, thereby reducing costs. The low AIC of the other postgraduates category may be a consequence of the inclusion in this category of teacher training where much of the course involves practical experience in the workplace. The large percentage of postgraduates in SCOPs falling into this category may therefore explain the low AIC for postgraduates reported in Table 5, where costs were estimated for each subgroup of institutions separately.

<Table 10 here>

4.7 Evaluating expansion policies

It is possible to use the estimated cost function to evaluate the costs of expanding higher education outputs, and, more importantly, to identify whether expansion should take place in existing institutions, or whether it would be better to provide new IHEs (although the latter policy would incur one-off building costs which have not been taken into account in the cost equations estimated in this paper). This section therefore considers a number of expansion scenarios and evaluates the costs of expansion using both the RE and SFA models. The application of the different estimating methods actually has different implications. In the RE model, the cost function has been calculated as a line of 'best fit' through the data, whereas in the SFA model, the cost function is a frontier around the data, and as such reflects cost practices in the most efficient IHEs. Expansion in the SFA case is therefore effected by expanding existing institutions efficiently or creating new efficient institutions. This is not the case in the RE model, since the RE cost function does not assume technical efficiency greater than that which is, on average, observed in the data.

In all cases, the models (RE and SFA) from Table 3 are used in the expansion estimations, the Oxbridge dummy is set to zero, and the base situation is considered to be one in which there are 120 institutions[13] each of which produces an identical output vector which is the average amount of each output. Initially, we consider the scenario in which higher education outputs are all simultaneously increased by 25%, and the cost effects will be evaluated using the RE and SFA models respectively. The RE model estimates total global current costs (i.e. costs of 120 IHEs producing average levels of output) to be £10008m, and if the 25% increase in output is effected by increasing output in existing IHEs, then costs rise to £12302m, a rise of 22.9%. The rise in costs from increasing the number of IHEs by 25% would, of course, be 25% (excluding the set-up costs). Thus the conclusion is that such an expansion should be achieved through increasing output in existing institutions, because of the ray economies of scale observed in the RE cost model. In comparison, the SFA model estimates total global current costs to be £8678m, and the total costs once output is expanded by 25% in existing institutions to be £10927m. This represents an increase of 25.9% which is somewhat higher than the increase which would occur if new IHEs were used to effect the expansion (excluding the set-up costs). In the SFA model, therefore, which is based on the cost patterns of efficient institutions, returns to scale are decreasing at the level of outputs under examination, and expansion might be better achieved by introducing more 'typical' IHEs to the sector rather than by increasing the size of the existing IHEs. In order to make a definitive decision, however, the set-up costs need to be compared with the discounted annual savings in costs over the lifetime of the new IHEs.

It should be noted also that the above analysis raises by 25% the outputs of all institutions irrespective of whether on an individual basis the institution faces increasing, constant or decreasing economies of scale. It is possible to develop a more discerning policy so that changes are targeted in line with the nature of returns to scale faced by each individual institution. Such an approach would likely be more effective for securing savings as it would essentially move each institution to its own most productive scale size given its mix of activities and their current levels.

Consider now the scenario in which undergraduate numbers are expanded by 25% (in all subjects), but all other outputs remain at the current average level. The RE model predicts the total costs once this increase has been effected in existing

[13] In fact, there are 121 institutions in our sample; the illustrative examples provided in the remainder of this section are easier to compute using 120, since 25% of 120 is an integer.

institutions to be £11000m (or a 9.9% increase). This compares with £11430m (or a 12.4% increase) which are the costs when 30 additional IHEs specialising only in undergraduate teaching are introduced to the sector, each producing average levels of undergraduates (in each subject) and zero levels of other outputs. Alternatively, if 10 IHEs specialising in undergraduate medicine, 10 in undergraduate science and 10 in undergraduate non-science are introduced and set up to achieve the same increase in student number, total costs are estimated to be £11382m (an increase of 13.7%). The preference in the RE model is therefore to expand existing IHEs and should come as no surprise given the product-specific economies of scale and scope predicted by this cost model for undergraduates. In comparison, the SFA model predicts these figures to be £9685m (i.e. an 11.6% increase), £9889m (a 14.0% increase) and £9782m (a 12.7% increase), respectively, and therefore confirms the result of the RE model in this instance. Product-specific economies of scale for undergraduate science and non-science, and product-specific economies of scope for undergraduate science generate the observed result.

Clearly the above calculations are merely representative of analysis that is possible given knowledge of the cost structure of institutions. In particular the calculations for economies of scope above need to be used as very broad brush indications in view of the use of institutions with numerous zero output levels, which are not found in the data used to estimate the cost functions. The calculations do also indicate the potential for conflicting policy implications arising from the choice of estimation method, which itself is underpinned by the analyst's assumptions about the attainability of technical efficiency both in existing and new institutions.

5. Conclusions

This paper extends the literature on higher education costs by refining the disaggregation of subjects, introducing third mission activities into the model, examining the cost differentials between different types of postgraduate qualification, considering the role played by location, and evaluating the importance of differentials in the quality of both student intake and output; the analysis also allows some simple examination of the cost implications of different expansion policies, and of the efficiency differentials that exist between institutions. This is all done using panel data drawn from the early years of the present decade for a sample of English institutions of higher education. As with any statistical exercise, the coefficients that we report are subject to varying degrees of precision.

Amongst undergraduates, medical students are found to be the most costly, and non-science students the least. Amongst postgraduates, those on taught courses are costly, while research students (presumably because they provide a source of cheap teaching and research assistance) are relatively inexpensive. This last finding contrasts spectacularly with the results obtained by HEFCE's cost transparency exercise. Quality considerations do not appear to impact significantly on the analysis. There is some evidence to suggest that provision in London is more costly than that elsewhere.

Estimates of economies of scale and economies of scope vary according to the choice of estimating technique. The RE model suggests that ray economies of scale and economies of scope are ubiquitous (though generally not huge). The SFA model suggests some product-specific economies of scale in research, but diseconomies elsewhere, and product specific economies of scope in undergraduate science, but diseconomies elsewhere. As a consequence, the RE model predicts that uniform expansion of all outputs can most efficiently be realised by expansion of the existing institutions rather than by creation of new ones, whereas the SFA model predicts that such an expansion should be effected by creating new (efficient) IHEs so long as the one-off set-up costs are less than the discounted savings in annual costs achieved over the lifetime of the new IHEs. When an unbalanced expansion of outputs is considered (i.e. expansion only of undergraduates), both estimating techniques indicate that such an expansion is best achieved by expanding existing IHEs than by creating specialist institutions.

Analyses of the type conducted in this paper provide a useful complement to the accounting based TRAC studies. The identification, through statistical analyses, of economies of scale and scope, and of other regularities in the data, suggest refinements to TRAC studies that need to be investigated – presumably by way of bottom-up case-study analysis.

Inevitably a limitation of any statistical approach is that it might fail fully to capture heterogeneity across data points. In further work we are investigating two approaches to this issue. The first involves the use of non-parametric methods such as data envelopment analysis. The second involves the exploitation of the panel nature of out data in order to produce cost function estimates in a random parameter framework.

Table 1: Definition of variables used in the analysis

Variable	Description
COSTDEF (Dependent variable)	Total operating costs in £000 in constant prices. This figure is inclusive of depreciation.
UGMED	Full-time-equivalent (FTE) undergraduates in medicine or dentistry (000).
UGSCI	FTE Undergraduate in science (000). Summation of subjects allied to medicine, veterinary, biological, agriculture, physical sciences, Maths, computing, engineering and architecture. (Includes weighted average of combined category)
UGNONSCI	FTE Undergraduate in non-science subjects (000). Summation of social economics, law, business, librarianship, languages, humanities, creative arts and education. (Includes weighted average of combined)
UG	Total of UGMED, UGSCI and UGNONSCI (000)
RESEARCH	Quality related funding and research grants, in £000000, constant prices.
PG	FTE postgraduate student numbers in 000s (NB PG is the sum of PGR, PGT and PGOTHER).
PGR	FTE postgraduate student numbers on research programmes (000).
PGT	FTE postgraduate student numbers on taught courses (000).
PGOTHER	FTE postgraduate student numbers on other postgraduate courses (000).
3RD MISSION	Income from other services rendered in £000000s in constant prices.
I*XY*	Interaction of variables *X* and *Y*
*X*SQ	variable *X* squared

Table 2: Descriptive statistics for the variables in the data set

Variable[1]	Obs	Mean	Std. Dev.	Min	Max
All IHEs					
COSTDEF	363	85930.76	89934.28	1422.19	462530.00
UG	363	6.148	4.750	0	20.149
UGMED	363	0.207	0.544	0	2.724
UGSCI	363	2.552	2.243	0	7.719
UGNONSCI	363	3.388	2.615	0	12.616
PG	363	1.733	1.447	0	6.068
PGR	363	0.451	0.683	0	4.556
PGT	363	0.829	0.727	0	3.120
PGOTHER	363	0.453	0.454	0	2.346
RESEARCH	363	22.127	43.445	0	219.974
3RD MISSION	363	4.354	5.380	0	31.042
SCOP Colleges					
COSTDEF	114	17639.36	12974.01	1422.19	51046.48
UG	114	2.266	1.845	0.026	7.192
UGMED	114	0	0	0	0
UGSCI	114	0.539	0.643	0	2.310
UGNONSCI	114	1.726	1.371	0	5.621
PG	114	0.441	0.514	0	2.429
PGR	114	0.031	0.037	0	0.154
PGT	114	0.115	0.125	0	0.594
PGOTHER	114	0.295	0.412	0	1.948
RESEARCH	114	0.445	0.571	0	2.468
3RD MISSION	114	0.715	1.537	0	8.512
Post-1992 IHEs					
COSTDEF	99	88766.05	22236.79	42805.00	133524.00
UG	99	10.342	3.155	4.669	19.753
UGMED	99	0	0	0	0
UGSCI	99	4.371	1.468	1.163	7.464
NONSCI	99	5.971	2.169	2.590	12.616
PG	99	2.132	0.866	0.768	4.078
PGR	99	0.233	0.127	0	0.703
PGT	99	1.166	0.507	0.432	2.590
PGOTHER	99	0.733	0.503	0.085	2.346
RESEARCH	99	4.813	3.077	0.171	12.547
3RD MISSION	99	4.847	2.523	0.498	12.800
Traditional (pre-1992) IHEs					
COSTDEF	150	135960.93	114473.63	9616.40	462530.00
UG	150	6.332	4.735	0	20.149
UGMED	150	0.502	0.755	0	2.724
UGSCI	150	2.882	2.255	0	7.719

NONSCI	150	2.947	2.314	0	11.223
PG	150	2.452	1.577	0.110	6.067
PGR	150	0.915	0.860	0	4.556
PGT	150	1.150	0.731	0	3.121
PGOTHER	150	0.388	0.364	0	1.613
RESEARCH	150	50.033	56.889	0.319	219.974
3RD MISSION	150	6.795	6.933	0	31.042

1. See Table 1 for precise definitions of variables.

Table 3: Estimated coefficients of the quadratic cost function

Variables[1]	Random effects Coefficients (z value)	Stochastic frontier Coefficients (z value)
UGMED	15094.60	14477.65
	(1.80)	(1.28)
UGSCI	9320.50	9500.10
	(5.50)	(4.34)
UGNONSCI	3088.18	3879.17
	(2.88)	(2.36)
PG	8553.31	4822.26
	(3.03)	(1.45)
RESEARCH	921.02	986.07
	(6.67)	(4.57)
3RD MISSION	1154.03	1021.76
	(6.13)	(10.13)
IMEDSCI	3135.91	1893.80
	(1.65)	(0.66)
IMEDNSCI	6110.62	6181.28
	(3.66)	(2.06)
IMEDPG	-16723.98	-17192.70
	(6.68)	(3.10)
IMEDRES	294.60	310.15
	(3.11)	(1.29)
ISCINSCI	-341.05	-276.31
	(0.84)	(0.34)
ISCIPG	-1233.85	-1221.13
	(1.77)	(1.12)
ISCIRES	-24.97	-19.51
	(0.87)	(0.36)
INSCIPG	898.74	801.92
	(1.73)	(0.92)
INSCIRES	-85.68	-83.43
	(3.74)	(1.44)
IPGRES	257.01	253.21
	(7.13)	(2.31)
MEDSQ	-582.49	1340.00
	(0.14)	(0.26)
SCISQ	27.87	-16.03
	(0.09)	(0.02)
NONSCISQ	47.38	-22.25
	(0.26)	(0.07)
PGSQ	-0.58	644.56
	(0.00)	(.51)
RESSQ	-2.86	-3.11
	(3.45)	(1.08)
OXBRIDGE	57581.42	45107.88
	(3.49)	(1.44)
Constant	6299.86	-2679.10
	(3.07)	(0.95)

Lagrangian test for random effects[2]	chi^2=183.78	
Log likelihood function		-3704.98
n	121x3	121x3

1. See Table 1 for precise definitions of variables.
2. The Breusch and Pagan Lagrange multiplier test tests the null hypothesis that $var(u_i) = 0$ against the alternative that $var(u_i) \neq 0$. It follows a chi-squared distribution with 1 degree of freedom. Rejection of the null hypothesis (if chi-squared>3.84) suggests that the random effects model is significant (at the 5% significance level).

Table 4A: AICs calculated the two quadratic models at the mean levels of output (full sample)

N=121 Output	Random effects AIC (£)	Stochastic frontier AIC (£)
UGMED	21220	17603
UGSCI	6196	6368
UGNONSCI	3308	3925
PG	10664	7574

Table 4B: AICs calculated using RE at 1.2 times and 0.8 times the mean levels of output (full sample)

N=121 Output	1.2 x mean AIC (£)	0.8 x mean AIC (£)
UGMED	22445	19995
UGSCI	5571	6821
UGNONSCI	3352	3264
PG	11086	10242

Table 5: Average incremental costs across subgroups calculated at mean output level using a quadratic specification

	SCOPs n = 38	Post-1992 IHEs n = 33	Traditional(pre-1992) IHEs n = 50
<u>RE</u>			
UGMED			20449
UGSCI	8241	2581	8448
UGNONSCI	3180	2890	3581
PG	2788	7725	13914

Note: The stochastic frontier models for subgroups of institutions proved to be unrobust – possibly as a consequence of the small numbers in the subgroups of IHEs - and the results from this model are not therefore reported here.

Table 6: Economies of scale (all IHEs)

a) Based on the RE model shown in column 1 of Table 3[1]

	Evaluated at:		
	Mean[2]	80% of mean	120% of mean
Ray economies	1.09	1.11	1.08
Product-specific economies			
Medicine Ug	1.01	1.00	1.01
Science Ug	0.99	0.99	0.98
Non-science Ug	0.95	0.96	0.95
Postgraduate	1.00	1.00	1.00
Research	1.07	1.05	1.08

b) Based on the SFA model shown in column 2 of Table 3[1]

	Evaluated at:		
	Mean[2]	80% of mean	120% of mean
Ray economies	0.96	0.96	0.97
Product-specific economies			
Medicine Ug	0.98	0.99	0.98
Science Ug	1.01	1.00	1.01
Non-science Ug	1.02	1.02	1.02
Postgraduate	0.87	0.89	0.86
Research	1.07	1.05	1.08

1. Oxbridge is set to zero.
2. Mean is the arithmetic mean over the 3 years (i.e. 363 observations).

Table 7: Economies of scope (all IHEs)

a) Based on the RE model shown in column 1 of Table 3[1]

	Evaluated at:		
	Mean[2]	80% of mean	120% of mean
Global economies	0.38	0.46	0.32
Product-specific economies			
Medicine Ug	0.06	0.08	0.04
Science Ug	0.17	0.17	0.18
Non-science Ug	0.07	0.09	0.06
Postgraduate	0.03	0.06	0.01
Research	0.04	0.06	0.01
Other services	0.08	0.09	0.06

b) Based on the SFA model shown in column 2 of Table 3[1]

	Evaluated at:		
	Mean[2]	80% of mean	120% of mean
Global economies	-0.18	-0.23	-0.15
Product-specific economies			
Medicine Ug	-0.05	-0.05	-0.04
Science Ug	0.07	0.04	0.10
Non-science Ug	-0.04	-0.05	-0.04
Postgraduate	-0.08	-0.08	-0.08
Research	-0.09	-0.09	-0.09
Other services	-0.04	-0.05	-0.03

1. Oxbridge is set to zero.
2. Mean is the arithmetic mean over the 3 years (i.e. 363 observations).

Table 8: Descriptive statistics for the efficiency estimates from the stochastic frontier analysis (2002/03)

	n	mean	standard deviation
All IHEss	121	0.69	0.32
post-1992 IHEs	33	0.84	0.077
pre-1992 IHEs	50	0.80	0.23
SCOP Colleges	38	0.43	0.38

Table 9: The effects of quality and location in London

Variable[1]	Quality added Coefficient (z value)	London Coefficient (z value)	Quality added and London Coefficients (z value)
RE			
quality	-113.08		692.87
	(0.04)		(0.26)
london		4488.55	5219.69
		(1.97)	(2.17)
SFA[2]			
quality	-58.74		
	(0.02)		
london		1250.02	
		(0.72)	
n	110x2	121x2	110x2

1. The variables quality and london have been added to a quadratic model (as reported in Table 3), but only the coefficients on the two variable of primary interest are reported here.
2. The SFA model failed to converge when both variables were added to the model together.

Table 10: Undergraduate and postgraduate activity disaggregated, RE estimation

Variable[1]	RE coefficient (z value)	SFA coefficient (z value)
ugmed	17102.04	17133.67
	(5.32)	(5.36)
ugsci	5232.08	4976.06
	(8.27)	(11.28)
ugnonsci	4408.90	4630.67
	(8.42)	(10.08)
pgr	3641.65	2517.67
	(1.08)	(0.93)
pgt	13887.25	10822.43
	(9.17)	(8.42)
pgoth	3045.69	4258.06
	(1.49)	(1.44)
research	1226.68	1209.07
	(20.93)	(28.98)
3rd mission	1228.21	1114.39
	(6.37)	(1.49)
oxbridge	59504.42	64322.27
	(4.92)	(3.48)
cons	6076.04	-2941.36
	(3.90)	(1.37)
log likelihood		-3750.65
n	121x3	121x3

1. See Table 1 for precise definitions of variables.

References

Aigner, D. and Chu, S-F. (1968) On estimating the industry production function, *American Economic Review*, 58, 826-839.

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

Baumol, W.J., Panzar, J.C. and Willig, R.D. (1982) *Contestable markets and the theory of industry structure*, San Diego: Harcourt Brace Jovanovich.

Bowen, H.R. (1980) *The costs of higher education*, San Francisco: Jossey Bass.

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

Cohn, E. & Cooper, S.T.(2004) Mulitproduct cost functions for universities: economies of scale and scope in Johnes, G & Johnes, J (eds) *The International Handbook on the Economics of Education*, Edward Elgar, Cheltenham.

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

Glass, J.C., McKillop, D.G. and Hyndman (1995b) The achievement of scale efficiency in UK universities: a multiple-input multiple-output analysis *Education Economics*, 3, 249-263.

Greene, W. (2002) Fixed and random effects in stochastic frontier models, mimeo available at http://pages.stern.nyu.edu/~wgreene/fixedandrandomeffects.pdf.

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

Johnes, G. (1996) Multi-product cost functions and the funding of tuition in UK universities *Applied Economics Letters*, 3, 557-561.

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

Johnes, G. (1998a) The costs of multi-product organizationa and the heuristic evaluation of industrial structure *Socio-Economic Planning Sciences*, 32, 199-209.

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. (2004) A fourth desideratum: the CES cost function and the sustainable configuration of multiproduct firms, *Bulletin of Economic Research*, 56, 329-332.

Jondrow, J., Lovell, C.A.K., Materov, I.S. and Schmidt, P. (1982) On the estimation fo technical inefficiency in the stochastic frontier production function model, *Journal of Econometrics*, 19, 233-238.

Nerlove, M. (1971) A note on error components models, *Econometrica*, 39, 383-396.

Scheffe, H. (1956) Alternative models for the analysis of variance, *Annals of Mathematical Statistics*, 27, 251-271.

Stevens, P.A. (2005) A stochastic frontier analysis of English and Welsh Universities, *Education Economics*, 13(4) forthcoming.

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

Wallace, T. Dudley and Hussain, A. (1969) The use of error components models in combining cross section with time series data, *Econometrica*, 37, 55-72.