Lancaster University Management School
Working Paper
2005/064

Higher education institutions' costs and efficiency: taking the decomposition a further step

Geraint Johnes and Jill Johnes

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

HIGHER EDUCATION INSTITUTIONS' COSTS AND EFFICIENCY: TAKING THE DECOMPOSITION A FURTHER STEP

Geraint Johnes
Jill Johnes

Lancaster University Management School
Lancaster LA1 4YX
United Kingdom

T: +44 1524 594215
F: +44 1524 594244
E: G.Johnes@lancs.ac.uk

October 2005

ABSTRACT

A multiproduct cost function is estimated for English higher education institutions using a panel of data from recent years. The panel approach allows estimation by means of a random parameter stochastic frontier model which provides considerable new insights in that it allows the impact on costs of inter-institutional differences in the cost function itself to be distinguished from inter-institutional differences in efficiency. The approach used here therefore resembles in some respects the non-parametric methods of efficiency evaluation. We report also on measures of average incremental cost of provision and on returns to scale and scope.

JEL Classifications: C14, C23, C51, D20, I20
Keywords: stochastic frontier, random parameter models, costs, higher education

1. Introduction

Research on efficiency measurement has, since the seminal work of Farrell (1958) bifurcated, with economists typically following the route of statistical analysis (Aigner et al., 1977) and management scientists characteristically opting for a non-parametric route grounded in linear programming (Charnes et al., 1978). The former approach has come to be known as stochastic frontier analysis, the latter as data envelopment analysis (DEA). The relative merits and demerits of the two approaches are by now well known: the parametric statistical approach benefits from the availability of the toolkit of statistical inference, but imposes a common functional form and common parameters on all decision-making units; the alternative non-parametric approach is attractive in that it does not impose a common loss function on all units, but it lacks a statistical apparatus and its results may be sensitive to the presence of outliers.

Recent developments in the analysis of panel data have made available a new approach which combines the merits of both the statistical and non-parametric methodologies while suffering from none of the drawbacks. Tsionas (2002) and Greene (2005) have developed random parameter formulations of the stochastic frontier model which (in common with data envelopment analysis) allow a separate loss function to be estimated for each decision-making unit while (in common with traditional frontier models) retaining the apparatus of statistical inference. In essence these models are simply a generalisation of the random effects frontier model introduced by Battese and Coelli (1995); while the random effects model allows only the constant to vary across decision-making units, however, the random parameters model allows any number of the other coefficients to vary as well. A distinction between these models and data envelopment analysis is that the cross-unit variation is constrained to follow a specified statistical distribution; this constraint allows us to retain the toolkit of statistical inference.

In the context of higher education institutions, the development of this new methodology is particularly significant. It is well understood that HEIs do not represent an homogenous group. Some are old, some are new, some are big some are small, some focus on certain subject groups, others focus on others, some are comprehensive in their provision, others are more specialised, some are research intensive, others not, and so on. Early studies of cost functions for UK institutions (such as Glass *et al*., 1995a, 1995b) focused purely on traditional universities. Later studies (for example, Johnes, 1997) looked at all universities, but excluded other providers of higher education such as colleges. The most recent work (Johnes et al., 2005) includes higher education colleges as well as universities, but devotes much space to the separate estimation of cost functions specific to certain pre-specified groups of institutions. This approach is far from ideal, however, because the distinctions between traditional universities, former polytechnics, and colleges of higher education have become increasingly blurred over time. An alternative approach, and the one on which the present paper is founded, is to develop an integrated framework for the estimation of costs, but to let the data decide the parameters of the cost function that apply uniquely to each institution.

To motivate the analysis a little further, consider a comparison between four institutions. One is an ancient university, where learning is delivered primarily through small group tutorials. This university has high costs because the student:staff ratio is necessarily low. But it delivers learning in a form that might be deemed desirable, albeit not one that would be cost-effective if applied to the mass of higher education institutions.[1] The second institution might also have high costs, but in this case they are due to locational factors; perhaps the institution is located in the nation's capital, where space and other costs are relatively high. The third institution has relatively high costs because (within the subject mix categories used in the analysis) it teaches expensive subjects; for instance, medicine may be more costly to deliver than other science subjects, but our analysis fails to disaggregate subjects sufficiently to identify medicine as a separate output. The fourth institution has moderate costs, as it does not have an adverse location or a need to employ unusually expensive teaching technologies. Now in a simple cross-section frontier analysis, the first three institutions may appear to be inefficient because of their high costs. In fact, however, there are reasonable explanations for these high costs, and these should not necessarily be put down to inefficiency. It is clear, therefore, that it is desirable that we should establish a method whereby unobserved heterogeneity in the cost function across institutions, on the one hand, and inefficiency, on the other, can be disentangled. That is the aim of this paper.

We employ recent developments in order to analyse the cost function for each higher education institution in England. Both random effects and more general random parameters models are estimated using panel data for three years, 2000-01 through 2002-03. Hence differences in intercept and slope coefficients across institutions can be estimated alongside differences in institutions' efficiency. The next section discusses the data. Results and analysis are provided in the following section. The paper ends with a conclusion and suggestions for further research.

2. Data

Our data are drawn from English institutions of higher education over the three year period from 2000-01 through 2002-03. Some 121 institutions are included in the analysis; this includes ancient universities (such as Oxford and Cambridge), traditional universities (comprising all those institutions with university status prior to 1992), new universities (granted university status in or since 1992), and colleges of higher education.[2] The sample therefore includes a heterogeneity of institutional types, and it is likely that it would be inappropriate to impose on any model of costs based on this sample a parametric form that does not allow coefficients to vary at least somewhat across observations.

All data are obtained from the Higher Education Statistics Agency (HESA): aggregate student numbers are published in Students in Higher Education Institutions, and financial statistics are available from Resources in Higher Education Institutions; institution-specific information about student numbers, disaggregated by subject area, was obtained from unpublished HESA sources. All financial data used in the study have been adjusted to 2002-03 values[3]. Student numbers are expressed as full-time equivalents. The costs measure includes both current and capital (in the form of depreciation) expenditures, but excludes 'hotel' type costs. These last costs, which measure costs due to the provision of student residences and catering, vary considerably from institution to institution, but they

[1] We realise, of course, that this is contentious. The assumption here is that the user of the analysis has a will to see teaching technologies of this kind preserved in some institutions but not others.
[2] A small number of instutions which changed significantly in character over the three year period, and for which therefore consistent data series are not available, is excluded from the sample.
[3] RPI inflators of 1.0366 and 1.0294 were applied to 2000/01 and 2001/02 figures respectively.

are costs that are generally recovered directly by imposing user charges, and their level in any one institution does not necessarily reflect the level of educational provision (the core business of the institution) to any great degree. In common with many other studies (dating back as far as Cohn *et al*., 1989), we use research income (both from research grants and contracts and from the funding council) as a proxy for research output. The limitations of this approach have been well rehearsed in the literature. We note that this measure is very highly correlated with more output-oriented measures (such as those derived from Research Assessment Exercise scores – see, for example, http://www.gla.ac.uk/rae/ukweight2001.xls), and we can therefore be confident that the use of our financially based measure does not bias the key results of the present paper. In common with the majority of previous empirical studies (Cohn *et al*. 1989; Glass *et al*. 1995a; 1995b; Johnes 1997; Stevens 2005) we do not include a measure of the third mission output of higher education institutions (namely knowledge and skills transfer). This is a deviation from the approach in the most recent study (Johnes *et al*. 2005) but is a necessary omission because the complexity of the statistical technique means that the estimation of the parameters is particularly demanding (see section 3). A parsimonious specification of the model is also desirable in order to avoid problems of multicollinearity. It is for these reasons that students are divided into only two broad subject groups, namely science (including medicine) and non-science.

Descriptive statistics appear in Table 1. One thing is very clear from these: the standard deviations for all variables are high in relation to the mean. While the means reported in the table refer, in a statistical sense, to a typical institution, the notion of such a typical institution can be very misleading. The higher education sector in England is one characterised by great heterogeneity. Nonetheless, the representative model of an institution that is suggested by the means in Table 1 is one that will strike many as familiar: the university has several thousand students, roughly evenly split between the 'arts' and 'sciences', and with about one in five students studying at postgraduate level. Mean costs are a little above £85 million. These vary considerably from institution to institution, depending upon the level of production of the various outputs. The precise nature of the mapping from outputs to costs is the subject matter of the next section of this paper.

3. Methodology and Results

Cost functions in economic theory represent an envelope or boundary which describes the lowest cost at which it is possible to produce a given vector of outputs. As it is an envelope that we wish to model, it is necessary to employ frontier methods of estimation rather than the more conventional best fit technology.

The conventional approach to stochastic frontier estimation, based upon cross-section data, is due to Aigner *et al*. (1977). In this model, the equation

$$y_i = \alpha + \boldsymbol{\beta}'\mathbf{x}_i + v_i \pm u_i \qquad (1)$$

is estimated using maximum likelihood, where v_i denotes normally distributed white noise error and u_i is a second residual term that is intended to capture efficiency differences across observations. This could in principle follow any non-normal distribution, though the half-normal is a common assumption.

A particularly appealing feature of this approach is that, following the insight of Jondrow *et al*. (1982) it is possible to recover observation-specific estimates of the efficiency residual. This estimator is given by

$$\mathrm{E}[u_i|\varepsilon_\mathrm{i}] = \sigma\lambda\{\phi(a_i)/[1-\Phi(a_i)] - a_i\}/(1+\lambda^2) \qquad (2)$$

where $\sigma = (\sigma_v^2 + \sigma_u^2)^{1/2}$, $\lambda = \sigma_u / \sigma_v$, $a_i = \pm\varepsilon_i\lambda/\sigma$, and ϕ(.)and Φ(.) are, respectively, the density and distribution of the standard normal.

When using panel data, it is appropriate to modify (1) to

$$y_{it} = \alpha_i + \boldsymbol{\beta}'_i\,\mathbf{x}_{it} + v_{it} \pm u_{it} \qquad (3)$$

where $v_{it} \sim N[0,\sigma_v^2]$, $u_{it} = |U_{it}|$, $U_{it} \sim N[0,\sigma_{ui}^2]$, and v_{it} is independent of u_{it}. Equation (2) is similarly modified, for the panel data case, to

$$\mathrm{E}[u_{it}|\varepsilon_\mathrm{it}] = \sigma\lambda\{\phi(a_{it})/[1-\Phi(a_{it})] - a_{it}\}/(1+\lambda^2) \qquad (4)$$

There are various ways in which one could implement this specification; for instance it would be possible to identify subgroups of the sample and estimate each parameter separately for each subgroup (Johnes *et al*. 2005). This is, in effect, the latent class estimator (Caudill, 2003). An alternative which we shall pursue in the present paper, is to model the β_i as random parameters. Greene (2005) summarises the problem by defining the stochastic frontier as (3) above, the inefficiency distribution as a half-normal with mean $\mu_i = \boldsymbol{\mu}'_i\mathbf{z}_i$ and standard deviation $\sigma_{ui} = \sigma_u\exp(\theta'_i\mathbf{h}_i)$, and the parameter heterogeneity is modelled as follows:

$$\left.\begin{aligned} (\alpha_i,\boldsymbol{\beta}_i) &= (\bar{\alpha},\bar{\boldsymbol{\beta}}) + \Delta_{\alpha,\beta}\mathbf{q}_i + \boldsymbol{\Gamma}_{\alpha,\boldsymbol{\beta}}\mathbf{w}_{\alpha,\boldsymbol{\beta}_i} \\ \boldsymbol{\mu}_i &= \bar{\boldsymbol{\mu}} + \boldsymbol{\Delta}_{\boldsymbol{\mu}}\mathbf{q}_i + \boldsymbol{\Gamma}_{\boldsymbol{\mu}}\mathbf{w}_{\mu i} \\ \boldsymbol{\theta}_i &= \bar{\boldsymbol{\theta}} + \boldsymbol{\Delta}_{\boldsymbol{\theta}}\mathbf{q}_i + \boldsymbol{\Gamma}_{\boldsymbol{\theta}}\mathbf{w}_{\boldsymbol{\theta}i} \end{aligned}\right\} \qquad (5)$$

Here the random variation appears in the random parameters vector $\mathbf{w}_{ji}$ (where *i* is the index of producers and *j* refers to either the constant, the slope parameter, or – in more general specifications of the model - the moments of the inefficiency distribution represented by μ and θ); this vector is assumed to have mean vector zero and, in the case where parameters are assumed to be normally distributed, the covariance matrix equals the identity matrix.

The parameters of this model cannot be estimated by traditional maximum likelihood methods because the unconditional log likelihood includes within it a term containing an unclosed integral. The obvious approach to adopt in this situation is to simulate the likelihood using Monte Carlo methods. Convergence to the solution of the problem therefore entails selection of numerous random draws of parameters, and so this is inevitably a computationally intensive exercise. Speed of solution can be reduced by employing Halton (1960) sequences of quasi-random draws. Such sequences have properties that resemble random series of numbers (and so can be used for simulation) but are in fact non-random and designed to facilitate rapid convergence in numerical integration problems. In the present case we have employed 100 Halton sequences; this is equivalent to the use of almost 1000 random simulations and is therefore in line with normal practice in Monte Carlo simulations. The simulated log likelihood function that must be maximised is

$$\log L_S = \sum_{i=1}^{N} \frac{1}{R} \sum_{r=1}^{R} \{ \sum_{t=1}^{T} \ln \Phi\{[\mu_{ir} /(\sigma_{uir} / \sigma_v) \pm (y_{it} - \alpha_{ir} - \boldsymbol{\beta}'_{ir}\, \mathbf{x}_{it})(\sigma_{uir} / \sigma_v)] / \sqrt{\sigma_{uir}^2 + \sigma_v^2}\} -$$

$$\frac{1}{2}\{[\mu_i \pm (y_{it} - \alpha_{ir} - \boldsymbol{\beta}'_{ir}\, \mathbf{x}_{it})] / \sqrt{\sigma_{uir}^2 + \sigma_v^2}\}^2 + \ln \frac{1}{\sqrt{2\pi}} - \ln \Phi(\mu_i / \sigma_{uir}) - \ln \sqrt{\sigma_{uir}^2 + \sigma_v^2}\} \quad (6)$$

The model is estimated using Limdep.

It is straightforward to observe that the traditional random effects model is a special case of the random parameters model; to be specific, the former is the case of the latter where only one parameter, namely the constant term, is allowed to vary across observations. In the results reported below, we report the random effects case as a point of comparison.

The recent literature on costs in higher education institutions is firmly built on the foundations provided in the literature on multiproduct cost function. This literature, which developed from the investigation of contestable markets, has highlighted the difficulty of choosing a cost function that makes sense in a multiproduct context. Baumol *et al*. (1982) propose three possible functional forms: the CES, the quadratic, and the hybrid translog. Problems attach to the first of these (Johnes, 2004), and the last is demanding both in terms of data and its highly nonlinear specification. We therefore restrict our analysis in the present paper to the quadratic cost function.

The results of four estimates of this cost function appear in Table 2. The first two columns report ‘best fit’ estimates. Model 1 is a standard random effects model where the constant is allowed to vary across institutions following a normal distribution. Model 2 is a random parameters model where both the constant and the coefficient on the full-time equivalent number of science undergraduates are allowed to vary, each following a normal distribution. Extensive experimentation, not reported here for reasons of space, has shown that, apart from the constant, it is only the coefficient on the linear term in science undergraduates that consistently exhibits significant variation across institutions. Models 3 and 4 are frontier counterparts to models 1 and 2 respectively.

Given the presence of quadratic and interaction terms in our preferred specification, the results in Table 2 are not straightforward to interpret. So we move quickly on to discuss some more intuitive results that emerge from our analysis. In Table 3, we report some measures of interest that are specific to each institution.[4] The first column reports the

[4] As noted at the bottom of Table 3, some 15 small and specialist institutions (see Table 2 footnotes) are excluded from this table because the frontier models predict that they will have negative costs. This is a result of the limited choice of legitimate functional forms in the multiproduct context. While we could conceivably eradicate the negative cost predictions by using, say, a Cobb-Douglas or a translog cost function, these are not legitimate functional forms to use where institutions are characteristically producers of multiple outputs. This,

random effects that are produced in model 1. These indicate systematic differences in costs across institutions that are not explained by differences in the explanatory variables. Such differences are due to unobserved heterogeneity – though as we shall argue later, this can be further decomposed. It is noticeable that the ancient universities (Oxford and Cambridge) have markedly higher costs than is the norm. Institutions located in or around London also tend to incur relatively high costs.[5]

The second and third column show, respectively, the intercept shifts and the efficiencies that are estimated by the random effects frontier model. This model therefore allows us to decompose the unobserved heterogeneity into two components: the intercept shift is designed to capture differences in the cost technology facing institutions (such as the location of the institution), while the efficiency term reflects differences in institutions' success in reaching *their own* cost frontier (maybe owing to differences in the quality of leadership). The measure of efficiency used here is the ratio of predicted costs to the sum of predicted costs and the value of the one-sided residual. Average efficiency is around 0.75, but this varies widely from 0.07 at Trinity and All Saints College to 0.97 at Cambridge. Thus the high costs at Oxford and Cambridge are readily seen to be the result of their idiosyncratic cost function, since, given this cost function, both are relatively efficient institutions with efficiency scores of above 0.9. Likewise, several of the London institutions (for example Imperial, Kings, University College London) face high costs but are deemed to be relatively efficient despite their high costs. Some others, such as City University, would appear to have high costs because of a mixture of an idiosyncratic cost function (which captures, amongst other things, high land prices in the capital) and a smaller efficiency score than some of their peers. Examination of the efficiency scores alone reveals a tendency for measured efficiency to be relatively low in smaller and more specialised institutions[6]. This is a feature noted also by Johnes *et al*. (2005).

The final three columns decompose further the unobserved heterogeneity. In addition to the intercept shift and efficiency score, use of model 4 allows us to investigate also the extent to which the cost functions faced by different institutions vary in terms of how costs respond to numbers of science undergraduates. It is readily seen that the coefficient on science undergraduates is markedly higher at Imperial College, University College London, Oxford, and Warwick than elsewhere. The reasons for this are likely to be varied; unobserved differences in the precise subject mix within the broad science category (and in particular whether an institution has medical students) is likely to be particularly important. In these final three columns, the intercept shift and efficiency scores exhibit much the same behaviour as that observed in the earlier columns.

It is worth emphasising a caveat concerning the interpretation of the efficiency scores derived from these models (3 and 4), which are calculated on the basis of institution-specific parameters (the constant for model 3, and the constant and coefficient on science undergraduates for model 4). Allowing some parameters to vary by institution brings the technique closer to DEA, and a well-known drawback of DEA is that units can be seen to be efficient simply because they are different from others in the data set. Thus the apparent cost efficiency of Oxford and Cambridge, both high-cost institutions, is questionable. It

in turn, is because any institution producing zero quantities of some outputs (and there are such institutions) could not be modelled by a Cobb-Douglas or translog function because the log of zero is indeterminate.

[5] Birkbeck College, which specialises in part-time provision, is unusual in this respect.

[6] Pearson's correlation coefficient between the natural logarithm of efficiency derived from model 4 and the natural logarithm of the total number of all students is 0.633.

would follow that some of the seemingly less efficient institutions could potentially become more efficient by attempting to emulate Oxford and Cambridge, yet encouraging institutions to become more like Oxford and Cambridge is not a practical or desirable policy for achieving cost efficiency either in the individual institutions or in the sector as a whole.

Much of the interest in studies of the cost structures of multiproduct institutions comes from statistics on average incremental costs associated with each output, and from statistics on economies of scale and scope. Standard measures of these were defined by Baumol *et al*. (1982) and have been used in numerous studies – including Cohn et al. (1989), Johnes (1997) and Johnes *et al*. (2005) – since. These now being standard and well understood definitions, we do not define them here, but proceed to report the various statistics that emerge from analysis of the two random parameter models – model 2 which follows the 'best fit' approach, and model 4 which follows the frontier approach.

Average incremental costs are shown in Table 4. These are reported, for each of the models, for a representative institution (namely one producing the mean level of each of the outputs), and also for institutions that produce 80 per cent and 120 per cent respectively of the mean of each output type.[7] The results indicate that science undergraduates cost between twice and three times as much to produce as do non-science undergraduates, and that postgraduate education is markedly more costly than undergraduate education. It is noticeable, however, that the frontier model estimates the average incremental costs associated with postgraduate education to be markedly lower than is the case with the 'best fit' model.[8]

The results shown in Table 5 indicate that product-specific returns to scale are exhausted for undergraduates in institutions close to the representative size. Economies of scale remain unexhausted in the context of postgraduate education and research, however. These results are robust with respect to choice of estimation method. They accord with the results presented in Johnes et al. (2005). Johnes (1997), using data for an earlier period and for a smaller sample of institutions, finds that product-specific economies of scale are exhausted for science undergraduates, but not for arts undergraduates.

Findings on ray returns to scale and on returns to scope are sensitive to the choice of estimation methodology. Using a 'best fit' method, ray economies of scale appear to be unexhausted, this being in large measure due to the fact that returns to scope are positive. However, using a frontier method, these results are reversed. This finding has clear implications for the further expansion of higher education in the UK. If current efficiency levels are taken as given, any further expansion of higher education should (in order to minimise global costs) be effected within the existing institutions. If, however, efficiency could be increased, overheads would fall and hence the opening of new institutions would become a viable option.

---

[7] The results in Tables 4 and 5 are based on the average value of the coefficients for the random parameters.

[8] One interpretation of this is that the 'best fit' model refers to actual *expenditures* rather than to the *costs* that need to be spent by an efficient institution. Bowen (1980) has argued that 'each institution raises all the money it can' and 'each institution spends all it raises'. If an institution can raise funds by hiking tuition for one output type – say postgraduates – then estimation of the equation by means of a 'best fit' method will tend to indicate that more postgraduates imply more expenditure. The frontier model does not suffer from this problem, since any expenditure that is above the cost frontier is attributed to inefficiency.

4. Conclusions

Earlier studies which have estimated cost functions for institutions of higher education have failed to recognise that, owing to unobserved heterogeneity, each institution likely faces a different cost function. In this paper, we use methods that have recently become available to estimate frontier cost functions for higher education institutions within the context of a random parameter model. This brings the analysis somewhat closer to the spirit of non-parametric techniques such as data envelopment analysis (and therefore has some of its drawbacks, such as its sensitivity to the presence of outliers), and allows questions to be answered about the distinction between inefficiency and idiosyncratic cost technologies. By allowing parameters to vary across institutions, cost functions for institutions that are obviously quite different from one another can be estimated within a single, unified framework, obviating the need for separate equations to be estimated for exogenously determined groups of institutions.

Our findings on returns to scale and scope, and on average incremental costs have much in common with the received literature. Findings that are new primarily concern the decomposition of cost differentials into components due to differences in cost technology, on the one hand, and efficiency, on the other. So, for example, while Izadi *et al*. (2002) comment on the London Business School (which in that study had a low measured efficiency score) as an idiosyncratic case, it is clear from the present analysis that the higher than expected costs of that institution are due in part to an unusual cost technology, and in part to efficiency issues.

Simple frontier models exist that simultaneously determine efficiency scores and explain them by reference to a vector of (environmental) variables. Such models have not yet been extended so that they can be used in a random parameter context. That would be an obvious development of the present work that must be left to the future.

References

Aigner, Dennis J., Lovell, Charles A. and Schmidt, Peter (1977) Formulation and estimation of stochastic frontier production function models, *Journal of Econometrics*, 6, 21-37.

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

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

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

Caudill, Steven (2003) Estimating a mixture of stochastic frontier models via the EM algorithm: a multiproduct cost function application, *Empricial Economics*, 28, 581-598.

Charnes, Abraham, Cooper, William W. and Rhodes, Eduardo (1978) Measuring the efficiency of decision-making units, *European Journal of Operational Research*, 2, 429-444.

Cohn, Elchanan Rhine, Sherrie L.W. and Santos, Maria C. (1989) Institutions of higher education as multi-product firms: economies of scale and scope, *Review of Economics and Statistics*, 71, 284-290.

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

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

Greene, William (2005) Reconsidering heterogeneity in panel data estimators of the stochastic frontier model, *Journal of Econometrics*, 126, 269-303.

Halton, John H. (1960) On the efficiency of certain quasi-random sequences of points in evaluating multidimensional integrals, *Numerische Mathematik*, 2, 84-90.

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

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

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

Johnes, Geraint, Johnes, Jill, Thanassoulis, Emmanuel, Lenton, Pam and Emrouznejad, Ali (2005) *An exploratory analysis of the cost structure of higher education in England*, London: Department for Education and Skills Research Report 641.

Stevens, Philip A. (2005) The determinants of economic efficiency in English and Welsh universities, *Education Economics*, 13(, forthcoming.

Tsionas, Efthymios G. (2002) Stochastic frontier models with random coefficients, *Journal of Applied Econometrics*, 17, 127-147.

Table 1 Descriptive Statistics

Variable	Mean	Standard deviation
costs (£x10^{-7}, 2003 prices)	8.593	8.990
science undergraduates (‘000)	2.760	2.519
non-science undergraduates (‘000)	3.388	2.615
postgraduates (‘000)	1.733	1.447
research income (£m, 2003 prices)	22.125	43.431

Table 2 Regression Results

Variable	Model 1: RE Gaussian	Model 2: RPM Gaussian	Model 3: RE Frontier	Model 4: RPM Frontier
constant	2.106 (53.84)[1]	1.083 (27.72)	-0.328 (4.41)	-0.328 (4.44)
ug sci	0.372 (9.25)	0.925 (22.80)	0.898 (12.67)	0.898 (12.90)
ug non-sci	0.280 (9.34)	0.014 (0.45)	0.211 (4.06)	0.211 (4.05)
pg	1.125 (15.22)	1.538 (20.29)	1.186 (8.83)	1.186 (8.85)
research	0.088 (28.11)	0.083 (25.96)	0.089 (16.40)	0.089 (16.46)
$(\text{ug sci})^2$	0.075 (8.58)	0.030 (3.44)	0.005 (0.31)	0.005 (0.31)
$(\text{ug non-sci})^2$	-0.014 (2.33)	0.004 (0.60)	0.023 (2.18)	0.023 (2.31)
pg^2	-0.107 (4.14)	-0.161 (6.03)	-0.133 (2.53)	-0.133 (2.55)
research^2	-0.0002 (15.35)	-0.0002 (14.96)	-0.0002 (7.46)	0.0004 (13.96)
ugsci*ugnonsci	-0.017 (1.46)	-0.029 (2.54)	-0.004 (0.18)	-0.004 (0.18)
ugsci*pg	-0.281 (13.54)	-0.271 (13.14)	-0.165 (4.65)	-0.165 (4.82)
ugsci*research	0.006 (10.34)	0.007 (11.56)	0.002 (2.37)	0.002 (2.30)
ugnonsci*pg	0.210 (12.14)	0.214 (12.13)	0.034 (0.99)	0.034 (1.02)
ugnonsci*res	-0.003 (5.82)	-0.002 (4.94)	-0.002 (1.82)	-0.002 (1.95)
pg*research	0.019 (16.46)	0.018 (15.14)	0.021 (8.13)	0.021 (8.56)

Random parameters[2]:				
constant	1.867 (81.81)	1.473 (80.45)	6.700 (39.33)	6.700 (40.22)
ug science		0.030 (6.24)		1.900 (39.55)
σ	0.475 (50.47)	0.479 (50.17)	1.900 (31.93)	1.900 (31.74)
λ			6.700 (7.48)	6.700 (7.57)
log likelihood	-457.177	-432.66	-673.58	-710.58

Notes: (1) t statistics in parentheses; (2) coefficients reported here are estimates of standard deviation of normal distribution of random parameters.

Table 3 Efficiencies, Intercept Shifts and Slope Shifts

Institution	intercept shift: model 1	intercept shift: model 3	efficiency: model 3	intercept shift: model 4	slope shift: model 4	efficiency: model 4
Anglia Polytechnic University	3.13	1.33	0.861	0.47	1.13	0.850
Aston University	1.20	-0.97	0.749	-0.37	0.70	0.745
Bath Spa University College	0.02	-0.86	0.561	-0.89	0.84	0.534
University of Bath	1.75	-0.50	0.852	-0.43	0.95	0.863
Birkbeck College	0.02	-0.83	0.782	-0.73	0.72	0.773
University of Birmingham	0.69	0.36	0.946	-0.29	1.21	0.957
Bolton Institute of Higher Education	0.71	-1.12	0.601	-0.96	0.80	0.590
Bournemouth University	1.46	-0.53	0.767	0.14	0.64	0.783
University of Bradford	2.13	-0.35	0.871	-0.36	0.92	0.875
University of Brighton	1.91	-0.41	0.787	-0.86	1.03	0.787
University of Bristol	1.80	-0.18	0.869	-3.67	1.74	0.899
Brunel University	0.41	-1.62	0.845	-1.78	0.94	0.850
Buckinghamshire Chilterns University College	2.11	0.54	0.750	0.31	1.10	0.750
University of Cambridge	4.17	3.64	0.971	6.56	1.93	0.984
Institute of Cancer Research	1.72	0.43	0.742	0.66	0.92	0.739
Canterbury Christ Church University College	0.48	-0.59	0.708	-0.70	0.95	0.698
University of Central England in Birmingham	2.93	1.04	0.865	-0.45	1.31	0.879
University of Central Lancashire	3.20	0.28	0.929	1.68	0.58	0.895
Chester College of HE	1.04	-0.57	0.665	-0.59	0.84	0.633
University College Chichester	-0.41	-1.50	0.464	-1.35	0.61	0.432
City University	4.33	1.92	0.795	-0.79	1.77	0.834
Coventry University	2.18	-0.09	0.838	0.26	0.84	0.840
Cranfield University	4.48	2.56	0.769	3.06	0.75	0.784
De Montfort University	2.21	0.23	0.882	0.25	0.91	0.885
University of Derby	1.97	-0.35	0.821	-1.25	1.17	0.823
University of East London	1.72	-0.12	0.842	-0.68	1.04	0.836
Edge Hill College of Higher Education	0.62	-0.68	0.681	-0.63	0.88	0.688
University of Essex	0.50	-0.86	0.839	-0.79	0.89	0.837
University of Gloucestershire	0.80	-1.04	0.643	-1.07	0.90	0.634
Goldsmiths College	0.18	-0.75	0.778	-0.73	0.89	0.779
University of Greenwich	3.82	1.98	0.877	-0.22	1.42	0.878
Harper Adams University College	0.79	-0.69	0.191	-0.68	0.85	0.173

University of Hertfordshire	3.49	1.21	0.888	0.45	1.03	0.885
University of Huddersfield	2.10	0.04	0.851	-0.07	0.92	0.846
University of Hull	2.27	0.40	0.883	0.48	0.87	0.879
Imperial College of Science, Technology & Medicine	4.38	2.86	0.903	0.42	2.47	0.928
Institute of Education	0.14	-1.08	0.710	-1.05	0.82	0.708
University of Kent at Canterbury	1.11	-0.52	0.860	-0.64	0.94	0.851
Kent Institute of Art & Design	0.78	-0.28	0.240	-0.30	0.93	0.235
King Alfred's College, Winchester	0.26	-0.81	0.421	-0.83	0.74	0.372
King's College London	5.99	4.63	0.958	0.71	1.91	0.978
Kingston University	2.41	0.17	0.901	0.55	0.80	0.880
University of Lancaster	0.80	-0.59	0.895	-0.40	0.85	0.891
Leeds Metropolitan University	3.55	0.92	0.877	-0.77	1.29	0.896
University of Leeds	0.50	0.53	0.921	0.12	1.13	0.928
University of Leicester	0.41	-1.09	0.827	-0.59	0.88	0.831
University of Lincoln	1.20	-0.58	0.887	-0.05	0.60	0.873
Liverpool Hope University College	0.30	-0.76	0.744	-0.73	0.53	0.589
Liverpool John Moores University	2.60	0.35	0.905	0.72	0.83	0.899
University of Liverpool	1.11	-0.46	0.897	-3.58	1.46	0.898
London Business School	3.78	2.62	0.790	2.62	0.99	0.787
University of London (Institutes and activities)	6.36	5.66	0.765	5.70	0.84	0.770
London Metropolitan University	2.02	0.13	0.854	-0.10	0.95	0.854
London South Bank University	3.69	1.37	0.916	0.95	0.98	0.911
London School of Economics and Political Science	0.60	0.99	0.893	1.09	0.82	0.891
London School of Hygiene & Tropical Medicine	0.96	-0.25	0.730	-0.06	0.84	0.726
Loughborough University	0.52	-1.53	0.859	-1.20	0.89	0.862
University of Luton	1.60	0.01	0.785	0.12	0.86	0.796
University of Manchester	2.39	2.73	0.966	0.29	1.40	0.945
University of Manchester Institute of Science & Technology	3.12	0.51	0.818	1.69	0.66	0.826
Manchester Metropolitan University	0.29	-1.33	0.947	1.01	0.54	0.932
Middlesex University	3.84	2.13	0.860	-0.09	1.45	0.855
University of Newcastle-upon-Tyne	3.62	1.87	0.915	-1.54	1.61	0.916
University College Northampton	1.10	-0.92	0.729	-0.78	0.80	0.714
University of Northumbria at	2.82	0.86	0.902	0.06	1.05	0.898

Newcastle						
Nottingham Trent University	1.47	-0.19	0.944	1.65	0.55	0.931
University of Nottingham	2.73	1.69	0.897	0.26	1.28	0.914
Oxford Brookes University	2.36	0.46	0.868	1.52	0.67	0.880
University of Oxford	3.78	2.92	0.932	5.44	2.12	0.935
University of Portsmouth	1.26	-0.82	0.878	1.01	0.58	0.875
Queen Mary and Westfield College	3.83	1.79	0.891	-0.10	1.47	0.913
University of Reading	0.97	-0.41	0.869	-0.69	1.10	0.876
University of Surrey, Roehampton	0.30	-1.13	0.739	-1.04	0.86	0.751
Royal Academy of Music	0.67	-0.30	0.127	-0.32	0.95	0.116
Royal College of Art	0.78	-0.21	0.513	-0.22	0.87	0.510
Royal College of Music	0.86	-0.25	0.118	-0.26	0.84	0.108
Royal Holloway and Bedford New College	0.80	-0.58	0.751	-0.31	0.66	0.740
Royal Veterinary College	1.55	0.22	0.591	0.15	1.04	0.603
St George's Hospital Medical School	2.73	1.07	0.825	0.96	1.05	0.834
College of St Mark and St John	-0.08	-1.18	0.413	-1.23	0.60	0.268
St Martin's College	0.32	-0.89	0.624	-0.97	0.99	0.635
St Mary's College	0.28	-0.89	0.337	-0.84	0.56	0.204
University of Salford	4.04	1.69	0.901	2.88	0.66	0.893
School of Oriental and African Studies	0.76	0.02	0.690	0.03	0.98	0.690
School of Pharmacy	0.27	-1.08	0.170	-1.03	0.90	0.205
Sheffield Hallam University	1.18	-0.73	0.881	-2.43	1.16	0.898
University of Sheffield	0.40	-0.60	0.877	0.23	0.99	0.885
Southampton Institute	2.28	-0.01	0.848	0.84	0.37	0.825
University of Southampton	2.29	0.75	0.903	1.53	1.02	0.902
Staffordshire University	1.67	-0.80	0.817	-0.12	0.73	0.817
University of Sunderland	2.51	0.40	0.838	0.29	0.90	0.825
Surrey Institute of Art and Design, University College	0.99	-0.12	0.423	-0.12	0.92	0.422
University of Surrey	4.56	2.27	0.813	1.73	1.14	0.810
University of Sussex	0.63	-1.16	0.794	-0.08	0.58	0.801
University of Teesside	2.41	0.05	0.772	-0.42	1.00	0.762
Thames Valley University	3.13	1.72	0.794	1.67	0.90	0.787
Trinity And All Saints College	0.14	-1.01	0.066	-1.05	0.99	0.064
University College London	4.65	4.22	0.961	3.14	2.41	0.970
University of Warwick	3.87	2.12	0.837	-2.01	2.08	0.861
University of West of England, Bristol	1.86	-0.41	0.838	0.86	0.71	0.839

University of Westminster	2.67	0.94	0.853	-0.32	1.16	0.857
University of Wolverhampton	2.97	0.53	0.919	1.29	0.67	0.899
University College Worcester	0.49	-0.64	0.547	-0.63	0.86	0.541
Writtle College	1.04	-0.27	0.272	-0.33	0.98	0.269
York St John College	0.43	-0.91	0.651	-0.65	0.34	0.467
University of York	1.63	-0.27	0.854	0.32	0.80	0.855

Notes: The efficiency measures vary a little from year to year; those reported in this table refer to 2002-03. Estimates are not reported here for 15 small and specialist institutions for which the predicted value of deflated costs in the frontier equations is negative. These are: Bishop Grosseteste College; Central School of Speech and Drama; Cumbria Institute of Arts; Dartington College of Arts; Falmouth College of Arts; Homerton College; Institute of Advanced Nursing; Newman College of Higher Education; Northern School of Contemporary Dance; Norwich School of Art and Design; Ravensbourne College of Design and Communication; Rose Bruford College; Royal Northern College of Music; Trinity College of Music; and Wimbledon School of Art.

Table 4 Average Incremental Costs

	Model 2			Model 4		
	at 100% mean output	at 80% mean output	at 120% mean output	at 100% mean output	at 80% mean output	at 120% mean output
undergraduate science	5516	6262	4770	6452	6958	5946
undergraduate non-science	2869	2323	3416	3126	2923	3329
postgraduate	16215	16049	16382	10527	1141	10261

Table 5 Economies of Scale and Scope

	Model 2			Model 4		
	at 100% mean output	at 80% mean output	at 120% mean output	at 100% mean output	at 80% mean output	at 120% mean output
Product-specific returns to scale:						
undergraduate science	0.87	0.90	0.82	0.98	0.98	0.97
undergraduate non-science	0.96	0.96	0.96	0.79	0.81	0.77
postgraduate	1.22	1.17	1.28	1.30	1.22	1.40
research	1.04	1.04	1.05	1.08	1.07	1.09
Ray returns to scale	1.10	1.15	1.07	0.97	0.96	0.98
Returns to scope	0.30	0.40	0.23	-0.17	-0.20	-0.15