Accounting for Background Variables in Stochastic Frontier Analysis

Philip Andrew Stevens[1]

National Institute of Economic and Social Research
2, Dean Trench Street, Smith Square, London, SW1 3HE, UK.

Abstract

The question of how to model background variables in stochastic frontier analysis is seldom discussed in the literature. Some studies, following on from a long history of two-step TFP studies, include them in the determinants of efficiency. Others include them in the frontier itself. We propose a statistical method to obtain the appropriate specification of the influences on costs and inefficiency. We provide the example of a stochastic cost frontier for a panel of English and Welsh universities.

May 2004

Discussion Paper Number 239

[1] Correspondence to: Philip Andrew Stevens, National Institute of Economic and Social Research, 2, Dean Trench Street, Smith Square, London, SW1 3HE. e-mail: philip.stevens@niesr.ac.uk. Tel: 020 7654 1927. Fax: 020 7654 1900

Acknowledgements: The research reported here was funded by the Leverhulme Trust. Data was supplied by the Higher Education Statistical Agency. Thanks go to Tim Coelli, Geriant Johnes, Martin Weale, John Rust, anonymous referees and participants at the NIESR/University of Birmingham Higher Education Conference in June 2002 and NIESR Public Sector Productivity Conference at the British Academy in November 2002 for helpful comments, although the usual caveats apply.

1. Introduction

Stochastic frontier analysis has become a popular tool for production analysis. One particular advantage is the ability to model the production relationship and the determinants of inefficiency in one stage. However, few studies discuss the issue of how best to model the effect of exogenous, or background variables. Do these factors affect production directly, through the production or cost frontier, or indirectly, though their effect on efficiency? Economic theory is relatively silent on this matter. In this paper we suggest a statistical criterion by which to choose the most appropriate specification of our empirical model.

Previous studies (e.g. Pitt and Lee, 1981, and Kalirajan, 1981) investigated the determinants of technical efficiency of firms by performing a second stage regression of the predicted efficiency scores on firm-specific factors. However, as Kumbhakar, Gosh and McGulkin (1991) and Reifschneider and Stevenson (1991) have noted, there is a significant problem with this approach. In the first stage, when the production or cost relationship is estimated, the efficiency terms are assumed to be independently and *identically* distributed, but in the second stage they are assumed to be a function of these firm-specific factors, implying that they are not identically distributed[2]. A popular method that overcomes this problem is that of Battese and Coelli (1995) whereby one estimates both the frontier and efficiency terms in one stage[3]. This formulation brings a particular issue into the open: should the exogenous influences on costs (or output) be included in the frontier itself or in the determinants of inefficiency.

[2] unless all the coefficients of the factors are simultaneously equal to zero
[3] The Battese and Coelli model is an extension of models such as Kumbhakar, Gosh and McGulkin (1991), Reifschneider and Stevenson (1991) and Huang and Liu (1994) into the panel context.

This question has largely been ignored in the literature (an exception is Coelli, Perelman and Romano, 1999), 'In the absence of a convincing theoretical argument with regard to the appropriate modelling of exogenous factors, we suggest a statistical one, based on the log likelihood of the estimated specifications. We provide a statistical criterion by which to choose the most appropriate specification of our empirical model, in terms of whether additional variables enter the cost function itself or the set of efficiency determinants.

2. Stochastic Frontier Analysis

Stochastic frontier models date back to Aigner, Lovell and Schmidt (1977) and Meesen and van den Broek (1977), who independently proposed a stochastic frontier production function with a two-part 'composed' error term. In the production context, where its use is most common[4], this error is composed of a standard random error term, representing measurement error and other random factors, and a one-sided random variable representing what Farrell (1957) called 'technical inefficiency', i.e. the distance of the observation from the production frontier. This notion of technical efficiency reflects the ability of a firm, country or university to obtain maximal output from a given set of inputs. It is measured by the output of the firm relative to that which it could attain if it were 100 % efficient, i.e. if it lay on the frontier itself, and is therefore bound between zero and one. When one combines this with allocative efficiency, the ability of the firm etc to use the inputs in optimal proportions, given their respective prices, one has a measure of total *economic efficiency*.

[4] See Lovell and Schmidt (1993) and Coelli, Rao and Battese (1999) for examples.

Using duality in production, one can consider cost efficiency[5]. This is particularly useful in the context of higher education institutions as it allows for the unit of observation to produce more than one product (Baumol, Panzar and Willig, 1982), 'A typical stochastic cost frontier would be

$$c_i = c(\boldsymbol{X}_i, \boldsymbol{\beta}) + \varepsilon_i \tag{1}$$

where (the logarithm of) costs for university i, c_i, depend upon a vector of variables, $\boldsymbol{X}$, and parameters, $\boldsymbol{\beta}$, and a composite error term $\varepsilon_i = \eta_i + \upsilon_i$, which is made up of a non-negative random variable, η_i, and a random error term, υ_i. It is this η term that represents inefficiency. Note that the inefficiency effect is *added* in the cost frontier, rather than subtracted, as is the case for the SFA production frontier. This is because the cost function represents minimum cost, whereas the production function represents maximum output. Also unlike the production frontier SFA approach, this inefficiency represents total economic inefficiency, i.e. technical inefficiency (not getting enough output from the inputs) plus allocative inefficiency (not using the inputs or producing the outputs in the correct proportions), 'Such a cost frontier is estimated for the UK HE sector by Izadi, Johnes, Oscrochi, and Crouchley (2002), although they do not investigate the determinants of inefficiency.

One question that arises from inspection of (1) is how one models the effects of exogenous factors on costs. These factors, which are beyond the control of the firm, at least in the short-run, are often called 'background' or 'environmental' factors. Firms operating in certain environments may incur higher costs, perhaps due to government regulations or problems with transport infrastructure etc. In our example of university costs, the standing

[5] E.g. Kumbhakar, Gosh and McGulkin (1991) and Reifschneider and Stevenson (1991)

of a university will affect its ability to attract pupils, staff and research funding. Also, if wages do not precisely reflect the marginal productivity of staff, this may increase or reduce costs *ceteris paribus*. One example of this is staff on permanent contracts. It is possible that their lifetime wage and productivity profiles will diverge (Disney, 1999), 'Those with high productivity may be rewarded by promotions, but this is a blunt instrument with which to equate marginal costs and productivity.

More generally, therefore, we can rewrite (1) as

$$c_i = c(\boldsymbol{X}_i, \boldsymbol{Z}_i, \boldsymbol{\beta}) + \eta(\boldsymbol{E}_i, \boldsymbol{\Delta}) + \upsilon_i \tag{1*}$$

where $\boldsymbol{Z}_i$ and $\boldsymbol{E}_i$ represent background variables affecting costs directly through the cost frontier and indirectly through inefficiency, respectively and $\boldsymbol{\Delta}$ is a vector of parameters to be estimated. The empirical counterpart to the question of how one decides whether an exogenous or background factor affects costs directly or indirectly is: how does one parse background variables into $\boldsymbol{Z}$ and $\boldsymbol{E}$?

The question of what determines inefficiency was one that was raised soon after stochastic frontiers were developed. Technical and allocative inefficiency can be caused by all sorts of productive and organisational inefficiency, from ineffective management to problems with inputs. In the context of our empirical example, we might ask: Are certain types of HEIs more likely to be inefficient than others? Does it depend upon the personal characteristics of the staff employed by an institution? or does a student body with particular characteristics make the efficient provision of HE more difficult? In the production context, work such as Pitt and Lee (1981) and Kalirajan (1981) investigated the determinants of technical efficiency of firms by performing a second stage regression of the predicted ηs on firm-specific factors. However, as Kumbhakar, Gosh and McGulkin

(1991) and Reifschneider and Stevenson (1991) have noted, there is a significant problem with this approach. In the first stage, the η terms are assumed to be independently and *identically* distributed, but in the second stage they are assumed to be a function of these firm-specific factors, implying that they are not identically distributed, unless all the coefficients of the factors are simultaneously equal to zero. Kumbhakar, Gosh and McGulkin (1991), Reifschneider and Stevenson (1991) and Huang and Liu (1994) presented models to overcome this problem by estimating the both the frontier and efficiency terms in one stage. These models were extended by Battese and Coelli (1995) to allow for panel data estimation. It is the Battese and Coelli (1995) method that we employ in this study.

Nevertheless, these models specify a relatively simple frontier and attribute all of the effect of background variables to inefficiency. Whether this is an appropriate modelling strategy is seldom discussed, although by necessity it has been discussed in the linear programming literature[6]. One paper which has investigated this issue is Coelli, Perelman and Romano (1999), who examined the effect of including the background factors either all in $\boldsymbol{Z}$ or all in $\boldsymbol{E}$. Our method extends this method to include intermediate cases where certain variables are considered to be direct cost-drivers ($\boldsymbol{Z}$) and other efficiency determinants ($\boldsymbol{E}$).

From a theoretical point of view, the decision over whether a variable should be included in $\boldsymbol{Z}$ or $\boldsymbol{E}$ boils down to ones *a priori* belief over whether such factors should

[6] In the linear programming literature, it was quickly realised that it was inappropriate for methods such as data envelopment to include environmental factors directly in the linear program itself for two reasons: First, this assumes that the firm had direct control over the factors. Since these environmental factors are, by definition, beyond the control of the firm, such an assumption is untenable. Second, since the inclusion of more inputs and outputs can never decrease efficiency and usually increase them quite considerably, the sensitivity-to-specification problem (Johnes, 1992) is magnified. This led to the development of the concept of 'non-discretionary inputs', with a more sophisticated linear program, but this only partially alleviated the problems (Banker and Morey, 1986; Staat, 1999).

affect the cost technology itself, or whether all universities share the same technology and these factors influence only the distance of each university from the best practice cost function. In the former case, the background variables are considered to be hedonic controls or quasi-fixed factors in the short-run cost function. These variables might be factors which impede universities ability to produce their bundle of outputs at the cost level associated with current best practice.

Since we cannot find a convincing theoretical argument for one particular specification over the others, *a priori*, we discriminate on statistical grounds. We estimate a number of different specifications with different sets of variables in $\boldsymbol{Z}$ and $\boldsymbol{E}$. We can create an artificial nested model that includes all background variables both in $\boldsymbol{Z}$ and in $\boldsymbol{E}$. We call this the 'general nested model'. Using likelihood-ratio tests, we can test the null hypothesis associated with each of the models against the alternative general nested model. Our selection criterion is that our preferred model passes the null hypothesis associated with it (i.e. that coefficients attached to the excluded variables are simultaneously equal to zero) against the alternative general nested model. In the event of more than one specification passing this test, our preferred specification will be that with the highest likelihood, although we will examine the effect of such a decision on the measure of efficiency obtained.

3. University Costs

There have been many investigations of university costs and their implications for the scale and scope of the provision of the various aspects of higher education institutions (for an overview, see Stevens, 2001), 'However, these studies almost exclusively assume that the university produces on the minimum-cost frontier.

There are many problems associated with the analysis of higher education institutions (HEIs) as producers. Rating their output is difficult without reference to a set of policy preferences. Is the primary purpose of the higher education sector merely to produce as many graduates as possible, or is it to safeguard the value of certification for the select few that undertake it? HEIs and society as a whole may have a number of objectives regarding the quality and quantity of both teaching and research. In the UK this problem has become more acute because of the blurring of the lines between the former polytechnics, with their emphasis on teaching and the vocational aspects in particular, and the 'old' universities with their accent on the more generalist and academic aspects. Tied up with this policy question is the analytical problem of how best to consider such institutions. Although the output of a HEI is generally considered to be the teaching and research it undertakes, the majority of the empirical work in this area before the 1980s ignored the research aspect of HEIs' activities. The view of a university as a producer in recent years has been overhauled in light of the pioneering work of Baumol, Panzar and Willig (1982) in industrial economics. HEIs are now considered as producers of multiple outputs.

Few would disagree with the assertion of Hare and Wyatt (1992) that 'The principal output of the higher education system is knowledge ... produced by ... research and teaching' (p. 48), 'However, it is difficult to be more specific when defining the precise *raison d'être* of a 'typical' higher education institution, or indeed whether such a typical institution exists. The higher education sector in the UK is now made up of new and old universities that were once polytechnics and universities with different functions, but are now practically identical entities[7]. Any model of university production must be able to

[7] Prior to 1992, institutions in the UK higher education sector were divided into independent 'universities' and 'polytechnics'. Polytechnics were originally set up to enable working-class men and women to advance

account for the fact that universities attach different levels of importance to the various aspects of higher education.

Early empirical work in this area concentrated on the cost per student of HEIs and the nature of the economies of scale that might exist[8]. However, an important factor to consider when investigating the performance of HEIs is their multi-dimensional nature. Over the last decade or so, work has recognised the multi-product nature of HEIs by including a measure of research output (typically research funding attracted) and graduate instruction (e.g. Cohn, Rhine and Santos, 1989; Glass, McKillop and Hyndman, 1995: Hashimoto and Cohn, 1997; Koshal and Koshal, 1999), 'Other studies have tried to account for the fact that the production of science and arts graduates both have very different cost implications. This is done by either including a dummy variable for lab-based subjects (Nelson and Hevert, 1992) or including numbers of both separately (Johnes, 1997; Izadi, Johnes, Oscrochi, and Crouchley, 2002), 'All of the previous studies suffer from two problems: the assumption of technical and allocative efficiency[9] and the omission of qualitative measures of inputs and outputs. Such work therefore is likely to give a seriously misleading picture of the effectiveness of the HE sector.

Previous studies have rarely taken into account the quality of the students enrolling at the HEI. An exception to this is Koshal and Koshal (1999) who included the average Student Aptitude Test scores of students. For the UK, work such as Bee and Dolton (1985), Dolton (1986), Johnes and Taylor (1987), and Naylor and Smith (2001) have found a strong positive relationship between degree results and A-level scores. Therefore, a failure

their general knowledge and industrial skills. The former polytechnics were maintained and regulated by local authorities until the Education Act 1988 made them autonomous and they gained university status following the Further and Higher Education Act 1992.

[8] For a review of this literature from a US perspective see Brinkman and Leslie (1986).

[9] We discuss these concepts in more detail in Section 3 below.

to account for input quality would provide an imprecise measure of university teaching output. The value of higher education to both the student and society as a whole is the ‘value added’ by the university (Johnes, 1992; Cave and Weale, 1992), ‘In this study we use the average A-Level/Highers score of students on entering the university as a measure of the quality of students as an input.

Another factor of which none of the previous studies mentioned above has taken into account is the *quality* of HEIs’ output. Whilst it is difficult, if not impossible, to account fully for any such variation, one can account for differences in degree classification. In the presence of external examiners, visits by assessors such as the Quality Assurance Agency (QAA) and the pressure of the maintenance of reputation, we believe that the proportion of first and upper second class degrees represents a relatively consistent measure of degree quality, and certainly the best that is readily available.

4. Empirical Model

In order to investigate empirically cost inefficiency in UK HEIs, we employ a multi-dimensional cost function model. Our empirical model is a translog cost function of the form

$$\begin{aligned}\ln\left(\frac{C_{it}}{P_{kit}}\right)_{it} &= \alpha_0 + \sum_{j=1}^{4}\beta_j \ln Q_{ji} + \tfrac{1}{2}\sum_{j=1}^{4}\sum_{k=1}^{4}\beta_{jk}\left(\ln Q_{ji}\cdot \ln Q_{ki}\right) + \phi_1 \ln\left(\frac{P_{sit}}{P_{kit}}\right) + \phi_2 \ln\left(\frac{P_{sit}}{P_{kit}}\right)^2 \\ &\quad + \sum_{j=1}^{4}\psi_j\left(\ln\left(\frac{P_{sit}}{P_{kit}}\right)\cdot \ln Q_{jt}\right) + \sum_{l=1}^{L}\varphi_l \ln Z_{it} + \varepsilon_{it}\end{aligned} \tag{2}$$

$$\varepsilon_{it} = \eta_{it} + \upsilon_{it}$$

where Q_j is the output of the j-th product (undergraduate teaching in the arts and in the sciences, postgraduate teaching and research), P_s and P_k are the prices of staff and capital respectively (average staff and capital costs deflated using the sector average in 1995/6[10]), 'Finally, $\boldsymbol{Z}$ is a vector of other factors that influence costs directly, η_{it} represents inefficiency, and υ_{it} are random errors, all of which we discuss in greater detail below. By normalising total cost and the other input price by the price of capital, we impose the theoretical condition that the cost function is linearly homogenous in input prices (Jorgenson, 1986), '

The outputs considered in this study are the numbers of arts and science undergraduates (*AUG* and *SUG*, respectively)[11], the number of post-graduate students (*PG*) and the total research funding attracted (*RES*), 'The use of what is apparently an input to measure an output may at first appear strange. Johnes (1997) argues that 'research grants are in general awarded to meritorious groups of researchers on the basis of the quality and quantity of their previous work (and … the weights assigned to quantity and quality in this measure are precisely those assigned by the 'market' for research)' (p. 728), 'Moreover, Koshal and Koshal (1999) support its use by the fact that there is be a high correlation between research output and grant support.' They find a correlation of 0.7 between this and faculty publications (based on data compiled from the National Academy of Sciences), for the institutions in the US where the latter data is available. There is a correlation of 0.65 between the research funding a university attracts and the average score across all departments in the 1996 Research Assessment Exercise[12]. The RAE is entirely

[10] See below for a more thorough description of these and the other variables.

[11] The science subjects are defined as those falling into the following categories: clinical medicine an dentistry, veterinary science, anatomy and physiology, nursing and paramedical studies, health and community studies, psychology and behavioural sciences, pharmacy, pharmacology, biosciences, chemistry, physics, agriculture and forestry, earth, marine and environmental sciences, general sciences, general engineering, chemical engineering, mineral, metallurgy and materials engineering, civil engineering, electrical, electronic and computer engineering, mechanical, aero- and production engineering, other technologies, architecture, built environment and planning, mathematics, information technology and systems sciences, sports science, and computer software engineering.

[12] UK HEIs are publicly-funded institutions and their research is assessed by the Research Assessment Exercise (RAE), whose purpose is to enable the higher education funding bodies to distribute public funds

retrospective and, moreover, also does not vary over the time-scale of this study. The research finding attracted is likely to be more representative of *current* research output as it is paying for current research. It does in part depend upon the outcome of the RAE, but also depends on more recent output. It also reflects research outside of traditional publications and so reflects universities' more general research role.

We also consider the effect of the quality of student intake, as measured by the average A-Level/Highers score. To our knowledge, the only other study of HEI costs which takes the quality of students at matriculation into account is Koshal and Koshal (1999), who found a positive correlation between costs and SAT scores at private universities in the US, although they found no significant effect in public institutions. This is consistent with the view that only the private universities have the financial means with which to bid for the higher quality students[13]. However, this result is likely to be the product of their omission of an output quality measure. Indeed in an earlier version of this paper (Stevens, 2001) we investigated the effects of the omission of output quality and found similar results to Koshal and Koshal (1999) when output quality was omitted, but a negative correlation when output quality is accounted for. It is clear that a correctly-specified model is important when making policy recommendations.

The main difference between our model and the majority of the literature on university cost functions is the inclusion of the efficiency effect, η_{it}. The only other model to include an effect of this nature is, to our knowledge, Izadi, Johnes, Oscrochi, and

---

for research selectively on the basis of quality. The RAE provides quality ratings for research across all disciplines using a scale of ratings range from 1 to 5*, according to how much of the work is judged to reach national or international levels of excellence. Note that our measure of research funding attracted also includes funding from the private sector.

[13] The only potentially similar split that the UK has is between the 'old' and 'new' (i.e. former polytechnic) universities. However, we found no evidence that costs or efficiency are significantly different between these two types of institution, or that they were differently trended. In what follows, therefore, they are considered as a whole.

Crouchley (2002), 'Our work represents an advance over Izadi, Johnes, Oscrochi, and Crouchley (2002) in four areas: First, Izadi *et al* look only at one period, whereas we have a panel of four years; second, we investigate the influence of a broader range of factors that may affect costs, either directly or indirectly; third we consider the possibility that inefficiency itself might also depend such factors[14]; fourth we utilise a statistical method to determine which of these additional variables affect costs directly, through their effect on the cost frontier, and indirectly, through their affect on the distance an HEI produces from that frontier (inefficiency), '

We describe the other direct influences on cost and the determinants of the efficiency effects, η_{it}, below. The υ_{it} terms are random errors, assumed to be i.i.d. and have $N(0, \sigma_{\varepsilon}^2)$-distribution, independent of the non-deterministic part of the η_{it} terms. Following Battese and Coelli (1995), the inefficiency effect is obtained by a truncation of the $N(\mu_{it}, \sigma^2)$-distribution, where

$$\mu_{it} = \delta \mathbf{E}_{it} \tag{3}$$

where $\boldsymbol{E}_{it}$ is a ($M\times 1$) vector of observable explanatory variables representing the characteristics of the student and staff bodies (discussed below), and δ is a ($1\times M$) vector of unknown scalar parameters to be estimated (which includes an intercept parameter), '

[14] Note that Izadi *et al* find that inefficiency exists, but discus only the implications of their results for economies of scale in any depth.

The realisations of η_{it} are not observable – we only observe $\varepsilon_{it} = \eta_{it} + \upsilon_{it}$. We may however define the efficiency predictor using the conditional expectation of $\exp(\eta_{it})$, given the random variable ε_{it}:[15]

$$
\begin{aligned}
EE_{it} &= E\left[\exp(-\eta_{it})\middle|\varepsilon_{it}\right] \\
&= \left[\exp\left(-\mu_{it} + \tfrac{1}{2}\tilde{\sigma}^2\right)\right] \times \left[\Phi\left(\frac{\mu_{it}}{\tilde{\sigma}} - \tilde{\sigma}\right) \middle/ \Phi\left(\frac{\mu_{it}}{\tilde{\sigma}}\right)\right]
\end{aligned}
\tag{4}
$$

where $\Phi(.)$ denotes the distribution function of the standard normal variable,

$$
\mu_{it} = (1-\gamma)\left[\delta_0 + \sum_{m=1}^{M} \delta_m E_m\right] - \gamma\varepsilon_{it}, \qquad \tilde{\sigma}^2 = \gamma(1-\gamma)\sigma^2, \text{ and } \quad \gamma = \frac{\sigma^2}{\sigma_\varepsilon^2 + \sigma^2}.
$$

We can obtain an operational predictor for the efficiency of university i at time t by replacing the unknown parameters in equation (4) with the maximum likelihood predictors. Note that in what follows we will be discussing *inefficiency*, that is $1/EE$. This is intended to clarify discussion, since anything that increases inefficiency also increases costs.

The log-likelihood function for this model is presented in Battese and Coelli (1993), as are the first partial derivatives of the log-likelihood function with respect to the different parameters of the model[16]. The generalised likelihood-ratio test for the null hypothesis that the γ parameter and the δ parameters are jointly equal to zero is calculated by using the values of the log-likelihood function for estimating the full frontier model and that obtained by using OLS regression to estimate the parameters of the cost function only.

[15] See Battese and Coelli (1993) and Coelli, Perelman and Romano (1999), '

[16] This parameterisation originates in Battese and Corra (1977).

This statistic has a mixed chi-square distribution, as indeed will any generalised likelihood-ratio statistic associated with the null hypothesis involving the γ parameter[17].

The γ-parameter lies between zero and one, and its value provides a useful test of the relative size of the inefficiency effects. If $\gamma = 0$, this would indicate that deviations from the frontier are due entirely to noise and that previous studies that use a standard (i.e. non-stochastic frontier) econometric methodology are entirely correct in their implicit assumption of technical and allocative efficiency. If $\gamma = 1$, however, this would indicate that all deviations are due entirely to economic inefficiency and hence the stochastic frontier model is not significantly different from the deterministic frontier model with no random error. Note however that γ is not exactly equal to the proportion of the total error term explained (except at values of $\gamma = 0$ and $\gamma = 1$)[18].

The inefficiency measures obtained by this method are gross measures of inefficiency, since they depend on factors that are beyond the control of the university, at least in the short run. In order to calculate net efficiency measures we recalculate the conditional expectation of η_{it}, replacing $\boldsymbol{\delta E}_{it}$ with $\min(\boldsymbol{\delta E}_{it})$[19]. This gives us measures of efficiency levels when all universities are assumed to face the most favourable conditions.

Because (2) is not defined for zero outputs, many studies have used flexible fixed cost quadratic form for (2) (e.g. Cohn, Rhine and Santos, 1989; Hashimoto and Cohn, 1997; and Koshal and Koshal, 1999), 'This is because multiple output concepts such as average incremental costs require the calculation of costs when one of the outputs is set to zero (Baumol, Panzar and Willig, 1982), 'Since none of the universities in our sample

[17] See Coelli and Battese (1999).
[18] C.f. footnote 7, page 188 of Coelli, Rao and Battese (1999); and Coelli (1995)
[19] C.f. Coelli, Perelman and Romano (1999)

produce zero of any of the outputs and our focus is on inefficiency, we do not face this problem and can use the translog specification. The translog function has the advantage that it incorporates constant returns to scale and approximates more sophisticated functional forms, such as the CES. What is important about this specification is that it takes account of the fact that universities make a quantity-quality trade-off. This is done by the inclusion of the proportion of students obtaining first or upper second class degrees in the ***Z*** vector. Costs per student are undoubtedly lower at universities outside of the premier league, but this does not make them necessarily more efficient. Previous studies have not included measures of output quality and so cannot account for the fact that there are a number of production strategies undertaken within the university sector. Again, this has serious implications for policy recommendations made which are based upon such analysis.

The additional factors that affect costs and efficiency (the ***Z*** and ***E*** variables) can be divided into three groups: staff characteristics; student characteristics and the ubiquitous 'other' category. This last group includes a time trend as well as the two measures of 'university type': the arts/science mix and the average A-Level score of entrants. Since the model only requires the unexplained portion of η_{it}, what we call 'net' inefficiency, to be uncorrelated with the frontier variables, we can include such variables in the deterministic portion[20]. We include the arts/science mix to investigate the possibility that universities that concentrate on education in arts subjects are more or less efficient than others. The second aims to proxy overall university 'quality'. It may be the case that some universities will attract high-attaining students, come-what-may (at least in the short term, in the longer term one would expect such running-down of social capital to reduce this ability), 'Such

[20] I am indebted to Tim Coelli for pointing this out.

universities can use this market power to attract students that are relatively cheap to educate, *ceteris paribus*, and use some of the savings to invest in social capital. One extreme example of this is the college tutorial system employed by Oxford and Cambridge. It may well be more cost effective – in terms of cost per student – to educate these high quality students in a more standard classroom format, but this may reduce the attractiveness of these institutions to future high-flying students.

The staff characteristics considered are: the proportion of staff that are aged over fifty years of age, the proportion of staff who are female, the proportion of staff who belong to non-white ethnic groups, the proportion of the teaching staff who are professors[21] or senior lecturers[22] and the proportion of staff who are research active[23]. The proportions are based on staff full time equivalents (FTE) rather than absolute numbers.

In addition to the direct effects of their A-level points, the student characteristics that we consider include the proportion of mature students (i.e. those over twenty five years of age on starting their course), the proportion of students who are female, the proportion of students from non-white ethnic groups, the proportion of students from lower social classes (i.e. SOC major groups 8 and 9, and those with unemployed parents/guardians), the proportion of students from EU countries outside of the UK and the proportion of non-EU students.

The effect of these variables may work in a number of ways. They could measure the determinants of inefficiency in the sense that certain types of staff are more efficient than

[21] As suggested by Earl (2001), who divides staff grades into five groupings: 'professors'; 'senior lecturers and researchers'; 'lecturers'; 'researchers'; and 'other grades'. The professor group includes: 'heads of departments (PCEF scale)'; 'professors (UAP minimum)'; 'research grade IV (UAP scale)'; 'clinical professors'; 'professors/heads of department (CSCFC scale)'; and 'professors in locally determined scales'.

[22] Again, defined according to the typology of Earl (2001) and including: 'principal lecturer (PCEF scale)'; 'senior lecturer (UAP scale)'; 'research grade III (UAP scale)'; 'Clinical senior lecturer', 'senior lecturer (CSCFC scale)'; and 'senior/principal lecturers in locally determined scales'.

others. These variables could also be considered as background or environmental variables in that the university has little control over them (at least over the short run) and are what is considered in the efficiency literature as 'non-discretionary'[24]. Even if the university *could* control the amount of particular types of student i.e. the number of students it attracts from ethnic minorities or other groups such as mature students, there is a question mark over whether they *should* do so. Such selection choices are often implicit rather than explicit, being the indirect result of other criteria, such as entry requirements.

In order to determine which factors affect costs directly, through the cost frontier itself, and which affect them indirectly, through their effect on efficiency, we discriminate on statistical grounds. We estimate a number of different specifications with different sets of variables in ***Z*** and ***E***. Since the set of potential specifications is very large, we divide the additional variables into six subsets:

1. Staff demographics – *STA51*, *STAF*, *STANW*
2. Staff grades – *STAPR* and *STASL*
3. Research-active staff – *STARA*
4. Student demographics – *STU25*, *STUFF*, *STUNW*, *STUC8*
5. Student domicile – *EUSTU* and *FORSTU*
6. Other efficiency drivers – *ARTSTU*, *ALEV* and *time-trend*

The final set of variables will be included in the efficiency determinants for all specifications.

[23] Defined as being active in the last Research Assessment Exercise.
[24] C.f. Worthington (1999), Stevens and Vecchi (2002).

5. Data

The data for the analysis come, directly and indirectly, from data collected by the Higher Education Statistics Agency (HESA) on behalf of the UK funding councils. The HESA data come from the Finance Record (1995/6 to 1998/9), the Individualised Staff Record (1995/6 to 1998/9) and the Individualised Student Record (1997/8), 'We also use annual data on entry requirements (Average A-Level/Highers points) and the percentage of students who achieve firsts and upper seconds published by the *Times Higher Education Supplement*. We confine our analysis to institutions of higher education in England and Wales, to avoid any problems caused by the differences in the system in Scotland. It is also argued that the universities in Northern Ireland also should be considered as separate entities from those on the mainland. This choice and various others of data availability give us a balanced panel of eighty institutions over four years. Definitions of the variables and descriptive statistics are presented in Table 1.

Table 1 Definitions of variables and descriptive statistics

NB descriptive statistics in table refers to original data (i.e. not logs)

		Mean	s.d.	Min	Max
TEXP	Total Expenditure (£m)	99.37	61.99	9.45	355.35
SUGN	Science Undergraduates (000s)	5.50	2.57	0.07	13.73
AUGN	Arts Undergraduates (000s)	6.32	2.67	0.49	16.75
PGN	Post-graduates (000s)	22.89	12.42	1.34	59.52
RES	Research Income (£m)	15.99	24.52	0.10	126.19
P_l	Average staff costs‡	27.25	4.48	15.32	41.52
P_k	Cost of capital§	149.60	76.07	2.16	389.50
ALEV	Average A-Level score	17.55	5.03	8.20	29.70
FIRST	Proportion 1st and 2:1s	0.54	0.10	0.28	0.87
STA51	Proportion of staff aged > 50	0.23	0.04	0.13	0.37
STAF	Prop^n of female staff	0.31	0.07	0.14	0.48
STANW	Prop^n of non-white staff	0.08	0.04	0.01	0.32
STAPR	Prop^n of professors	0.08	0.06	0.00	0.22
STASL	Prop^n of senior lecturers	0.19	0.08	0.00	0.74
STARA	Prop^n of RAE active staff	0.44	0.19	0.00	1.00
STU25	Prop^n of students aged ≥ 25*	0.10	0.03	0.05	0.18
STUF	Prop^n of female students*	0.51	0.06	0.29	0.65
STUNW	Prop^n of non-white students*	0.18	0.13	0.02	0.54
STUC8	Prop^n of students from lower classes†*	0.14	0.11	0.00	0.92
EUSTU	Prop^n of students from other EU*	0.06	0.03	0.02	0.19
FORSTU	Prop^n of non-EU students*	0.09	0.06	0.02	0.41
ARTSTU	Prop^n of arts students	0.48	0.03	0.40	0.58

* These come from the Student Record 1997/8, and, therefore, do not have a time dimension.

† Defined as having their prime parent/guardian who is employed in Standard Occupational Classification major groups 8 (plant and machine operatives) and 9 (other occupations), or is unemployed.

‡ Deflated by 1995/6 sector average.

§ Capital expenditure (including expenditure on equipment, furniture, land and buildings) divided by total net assets.

6. Results

The results of the maximum likelihood estimation of the stochastic frontier quadratic cost functions for the academic-years 1995/6 to 1998/9 are summarised in Table 2. In the table we present the number of variables included in the ***Z*** and ***E*** vectors, the value of the log-likelihood function and its ranking. In the final column of Table 2 we include the results of the likelihood ration test of the null hypothesis associated with the model against the general model.

Table 2 Summary of Results

Model	***Z***	***E***	*Log-likelihood*	*Rank*	χ^2_{12}
1	12	3	353.62	12	9.98
2	10	5	349.48	18	18.27
3	8	7	348.70	22	19.82
4	6	9	349.76	16	17.71
5	11	4	348.92	21	19.38
6	9	6	353.93	10	9.37
7	7	8	348.30	23	20.62
8	5	10	346.93	28	23.36
9	10	5	342.70	30	31.82
10	8	7	351.47	13	14.29
11	6	9	349.32	20	18.59
12	4	11	350.05	14	17.12
13	9	6	354.43	7	8.36*
14	7	8	353.95	9	9.32
15	5	10	341.93	32	33.36
16	3	12	342.25	31	32.73
17	9	6	354.40	8	8.43*
18	7	8	353.70	11	9.83
19	5	10	348.22	25	20.79
20	3	12	345.97	29	25.29
21	8	7	356.02	3	5.18***
22	6	9	354.55	6	8.13*
23	4	11	349.78	15	17.67
24	2	13	348.27	24	20.68
25	7	8	356.33	2	4.57***
26	5	10	354.87	4	7.49*
27	3	12	349.38	19	18.47
28	1	14	347.58	27	22.06
29	6	9	357.14	1	2.94***
30	4	11	354.70	5	7.82*
31	2	13	349.68	17	17.87
32	0	15	347.63	26	21.97
General	11	15	358.61		

Notes:

- χ^2_{12} is the value of the likelihood ratio test of the null hypothesis associated with each of the models against the alternative general model. This test has twelve degrees of freedom.
- *, **, *** signify that the constraint associated with the model can be accepted at the 10%, 5% and 1% levels respectively.

The first thing to note from Table 2 is that for neither of the two extremes (all of the variables in $\boldsymbol{Z}$ or all in $\boldsymbol{E}$[25]) can we accept the null hypothesis associated with the specification, even at the 10% level. There are three specifications (21, 25 and 29) for which we accept the null at the 1% level, none at the 5% level and five more at the 10% level. In what follows, we will concentrate our discussion on specifications 21, 25 and 29 (although we will consider the specification with the lowest log-likelihood (model 16) later for comparative purposes).

The results our estimation for the three specifications with the highest log-likelihoods are presented in Table 3 and Table 4[26]. The implications of our cost function for economies of scale and scope in university production are not the main thrust of this study and so we do not, therefore, consider them in any depth. However, we do note that the β coefficients are consistent across specifications (this is also true of the other specifications, not reported here) and that our results suggest that there have been no trend in costs over the period.

[25] models 1 and 32, respectively

[26] Note that for the estimated relationship to be a cost function requires convexity and monotonicity. The estimated functions are concave in input prices and are increasing in input prices and the level of output at mean values of the variables.

Table 3 Results, Cost Function

	Model	*21*	*Model*	*25*	*Model*	*29*
	β_k	t	β_k	t	β_k	t
	2.677	(10.6)	2.822	(9.4)	2.831	(8.9)
SUGN	0.112	(1.9)	0.055	(0.8)	0.065	(1.1)
AUGN	0.004	(0.0)	0.013	(0.1)	-0.027	(0.2)
PGN	0.538	(4.7)	0.661	(4.9)	0.637	(5.8)
RES	0.074	(1.6)	0.057	(1.1)	0.043	(0.9)
$SUGN^2$	0.044	(2.7)	0.059	(3.1)	0.044	(2.7)
$AUGN^2$	0.086	(3.9)	0.075	(3.4)	0.082	(3.8)
PGN^2	0.072	(1.4)	0.166	(2.8)	0.151	(2.8)
RES^2	0.064	(7.3)	0.076	(7.6)	0.075	(8.2)
$SUGN \times AUGN$	0.044	(1.1)	0.058	(1.3)	0.061	(1.6)
$SUGN \times PGN$	-0.084	(2.2)	-0.129	(3.3)	-0.111	(2.2)
$SUGN \times RES$	0.007	(0.4)	0.015	(0.9)	0.019	(1.2)
$AUGN \times PGN$	-0.193	(2.8)	-0.196	(2.8)	-0.203	(2.0)
$AUGN \times RES$	0.014	(0.5)	0.007	(0.3)	0.019	(0.7)
$PGN \times RES$	-0.068	(1.7)	-0.126	(2.9)	-0.127	(3.1)
P_l	0.152	(1.2)	0.177	(1.1)	0.119	(1.4)
P_l^2	-0.153	(1.1)	-0.086	(0.6)	-0.166	(1.2)
$P_l \times SUGN$	-0.085	(1.0)	-0.006	(0.1)	-0.021	(0.2)
$P \times AUGN$	0.000	(0.0)	0.001	(0.0)	0.058	(0.6)
$P \times PGN$	-0.045	(0.4)	-0.202	(1.5)	-0.262	(1.5)
$P_l \times RESN$	0.071	(1.4)	0.109	(2.0)	0.120	(1.5)
ALEV	-1.528	(3.3)	-1.118	(2.4)	-1.232	(2.9)
FIRST	4.730	(3.5)	3.418	(2.6)	3.809	(2.9)
time trend	0.005	(0.9)	0.004	(0.6)	0.006	(1.0)
Log likelihood		356.02		356.33		357.14
LR test of the one-sided error		78.6		84.06		97.17

Unlike Glass, McKillop and Hyndman (1995), we do find evidence of a complimentarity between research and post-graduate teaching. One might expect this *a priori* since graduate instruction is much more closely related to research and one can see the potential for economies of scope from the viewpoint of academic staff more clearly.

Moreover, many students reading for taught post-graduate courses are preparing for research and those reading for degrees by research are actually undertaking research themselves. Certainly, Cohn, Rhine and Santos (1989) also find a statistically significant positive relationship between postgraduate enrolment and research in both public and private US universities. The absence of such an effect in Glass *et al* (1995) may be due to the omission of input and output quality variables and the new universities from their analysis[27]. It is certainly true that new universities have both lower numbers of post graduate students and also have lower research ratings (and thus attract much less research funding), 'We also find complementarities between both science and art undergraduate teaching with postgraduate output.

Of particular interest to this study are the results for our student input and output quality measures, since we are the first to our knowledge to include both. Our results confirm our expectation that having better quality inputs (as measured by their A-level scores) reduces costs, *ceteris paribus*, and producing better quality outputs (as measured by the proportion of firsts and upper seconds) costs more money. There are two, related, implications of this result. First, earlier results using only input quality that suggested that higher quality inputs lead to higher costs (Koshal and Koshal, 1999) are likely to be a product of conflating input and output quality. This is because they did not account for the fact that universities with high SAT scores also are likely to produce better quality students at the end. If an output quality measure is excluded, then the implicit assumption is that all HEIs are producing the same quality of graduates. The second implication is more general. The omission of a measure of output quality assumes that all of the HEIs are following the same policy of teaching as many students as possible, and that any variations in costs per student are due to scale, technical or allocative inefficiencies. This is clearly not the case

[27] At the time of Glass *et al*'s analysis, these were still polytechnics.

when comparing the top research universities, say, with the top 'new' universities, many of whom have successfully implemented strategies of mass undergraduate instruction. This is particularly pertinent when one considers the present UK government's desire to have 50% of 18-30 year olds experience higher education by 2010.

Turning our attention to the ***Z*** and ***E*** variables, an interesting pattern emerges. In all of the top three specifications, student characteristics appear in the frontier and staff demographics appear in the efficiency determinants. The only difference between the three specifications is which of the other two sets of staff characteristics go in ***Z*** and which in ***E***. Our preferred specification (model 29) all of the staff characteristics in the efficiency determinants, i.e. it has both the staff grades and percentage of research staff variables in ***E***. Specifications 21 and 25 have one of these sets of variables in ***Z*** – staff grade and research-active staff respectively. This consistency of specification enables us to be fairly confident in our results.

Table 4 Results, Cost Function (cont.)

	Model 21		*Model 25*		*Model 29*	
	β_k	t	β_k	t	β_k	t
STAPR	-0.556	(3.0)				
STASL	-0.052	(2.4)				
STARA			0.032	(0.7)		
STU25	-1.543	(4.8)	-1.676	(5.0)	-1.481	(3.1)
STUG	0.045	(0.3)	0.092	(0.5)	-0.033	(0.2)
STUNW	-0.431	(6.0)	-0.305	(3.4)	-0.306	(3.6)
STUC8	0.156	(2.7)	0.159	(2.5)	0.203	(3.2)
EUSTU	0.091	(0.3)	-0.072	(0.2)	-0.209	(0.7)
FORSTU	0.677	(3.1)	0.563	(1.7)	0.651	(3.0)

The effects of student characteristics are also consistent across the specifications. The three significant factors are the proportion of mature and non-white students and that of students from lower social classes. Universities with a larger proportion of mature or non-white students have lower costs *ceteris paribus*, and those with a larger proportion of students whose parents/guardians are from lower social classes (or are unemployed) higher costs. The idea that mature students appear cheaper may at first be counter-intuitive; it certainly goes against the formula whereby the Higher Education Funding Councils allocate their funding[28]. The reason for this result is likely to be related to their A-Level results. Mature students in general have lower A-Level scores than students who go to university straight from school (including those that have gap years), 'Indeed, many have no A-Level or equivalent qualifications at all. Thus HEIs with a high proportion of mature students have lower costs, for a given average level of entrance qualification. Someone who is under the age of twenty-five on entering higher education with a low A-level score is likely to do so because they did badly in their exams, whereas mature students are likely to have low A-level scores for a number of reasons. They may have left school without these qualifications and either entered university through some other route (such as a foundation course or using professional qualifications or experience) or taken them at night school. Students whose parents/guardians are from lower social classes are likely to be a different set of students. It is important to note that mature students are much less likely to report the social class of their parent or guardian[29]. Students from lower social classes are also likely to have lower A-level scores[30] but for different reasons than mature students.

[28] The Higher Education Funding Council allocates an extra 5% of funding for mature students (over 25 years of age) in their first year of study.
[29] The data in the student record come from the students' UCAS (university application) forms. The percentage of mature students who answer 'do not know', or do not answer the question is over 90%, as opposed to around 30% for younger students.
[30] Students with parents from the top two major SOC groups have a mean A-level points score of 20, whereas those in the bottom groups have less than 15.

There is evidence to say that social class is a good predictor of exam success and that this continues throughout the students' life[31]. Our results suggest that the negative effects of low social class go beyond their A-level results. That is, universities with large numbers of students from lower social classes have higher costs over and above those one would expect because of their students' lower average entrance qualifications. If society believes that part of the value added by the state university system is one of enabling members of lower social classes to undertake university education when they would otherwise have not done so, then this coefficient reflects the cost of such redistribution of opportunity.

Universities with a high proportion of non-white students tend to have lower costs once we account for their social class, A-level grades and their final degree. Students from ethnic minorities tend to come from lower social classes, achieve lower grades at school and be more likely to attend lower achieving universities (Modood and Shiner, 1994), 'Our results suggest that, like mature students, once we take account of other aspects of their background non-white students do not impose any additional costs on HEIs and are in fact associated with lower costs.

We found no significant effect of the proportions of female students on expenditure in any of the three preferred specifications. We also found no statistically significant effect of non-UK EU students although non European-domiciled students appear to be more costly to teach.

[31] For an excellent overview of work in this area, see Haveman and Wolfe (1995).

Table 5 Results, Inefficiency Effects

	Model 21		*Model 25*		*Model 29*	
	δ_k	t	δ_k	t	δ_k	t
	0.456	(1.2)	0.134	(0.4)	0.321	(1.0)
STA51	1.016	(2.8)	1.183	(3.3)	1.078	(3.2)
STAF	0.941	(3.3)	0.832	(2.4)	0.887	(3.3)
STANW	0.002	(0.0)	-0.205	(0.4)	-0.195	(0.5)
STAPR			-0.933	(2.5)	-0.612	(2.7)
STASL			-0.036	(1.3)	-0.045	(1.3)
STARA	-0.038	(0.4)			-0.064	(0.6)
ARTSTU	1.041	(2.0)	1.193	(1.7)	0.891	(2.1)
ALEV	-0.461	(4.2)	-0.341	(2.8)	-0.333	(3.6)
time trend	-0.024	(1.4)	-0.020	(1.3)	-0.031	(2.1)
$\sigma_{\varepsilon}^2 + \sigma^2$	0.015	(10.8)	0.014	(9.5)	0.011	(7.5)
γ	0.994	(35.2)	0.987	(67.8)	0.995	(47.4)

Turning our attention to staff characteristics, we see that for all three of the specifications for which we can accept the associated restriction, some or all of the staff demographic variables appear in the efficiency determinants, ***E***. In our preferred specification, Model 29, all of the staff variables are included in ***E***. The difference between Models 21 and 25 is that the former has the staff grade variables in ***Z*** and the latter has the proportion of staff who are research active. The effect of the proportion of staff of professorial grade on cost is significantly negative, whether it appears in the frontier or in the efficiency terms. The effect of the proportion of staff of senior lecturer grade on costs is also negative, but this result is only statistically significant in Model 21. The proportion of staff who are research active does not appear to have a significant effect on costs, either directly or indirectly. Universities with high proportions of staff who are aged over fifty, or who are female, are generally less efficient, *ceteris paribus*. We find no relationship between efficiency and the proportion of staff who are non-white. These results are consistent across specifications. The fact that staff who are of a higher ability (i.e. those that reach the grade of professor or senior lecturer) are more efficient should not be a great surprise. Note that this is above and beyond the direct cost they incur in terms of higher wages, which operates through the cost function itself. Given that wage profiles of staff are determined by grade and age, the negative coefficients on the proportions of professors and senior lecturers and the positive one on staff age 51 and over could be related. Since promotions in grade are likely to be related both to innate ability and experience, those members of staff who are over fifty but are not senior lecturers or professors may well be less efficient than those who are. The negative effect of female academic staff on efficiency is more difficult to explain. One explanation of HEIs with a large number of female staff is that it could be a human capital effect, caused by absence from the labour market due to career breaks. However, studies have consistently found that women earn less than men for a given set of characteristics and so one might expect the effect to work in the opposite direction, since women's wages reflect their productivity less than that of men.

The relationship between 'university type' and efficiency on efficiency is consistent across the specifications. Universities with higher proportions of arts students are less efficient that those with more science students, *ceteris paribus*. Note that this result is relative to a frontier which includes the numbers of arts and science students separately. One inference that we can make from this result, therefore, is that the efficiency scores of arts-based universities are more dispersed than those of science-based institutions. Universities with better quality intake tend to be more efficient than their peers. Again, since the frontier takes account of the direct impact of student quality on costs, this suggests that there is less dispersion of efficiency scores amongst the premier league institutions.

Not only is there a consistency of specification – in terms of which variables appear in $\boldsymbol{Z}$ and $\boldsymbol{E}$ – but also a consistency of results. The signs, magnitude and significance of the student characteristics – all of which appear in $\boldsymbol{Z}$ in all three specifications – are all very similar. The same can be said for the university type and staff characteristics which appear in $\boldsymbol{E}$. The only difference between the three specifications with the highest log likelihoods is where the variables measuring the proportion of staff who are professors or senior lecturers or those who are research active. The implications of the coefficients on these variables on total costs are also very consistent, whether they affect costs directly or indirectly.

As have noted above, the γ-parameter gives an impression of the influence of the efficiency terms as the γ value shows the contribution of the η efficiency term to the whole of the dichotomous error term ($\eta + \upsilon$), 'The estimates of γ for models 21, 25 and 29 are 0.994, 0.987 and 0.995, respectively. Although as we have stated above, these do not correspond exactly to proportions of variation, they do suggest that the majority of the error is explained by inefficiency. The generalised likelihood ratio test of $\gamma = 0$ indicates that the inefficiency effect is significantly different from zero. Therefore, the 'average response' cost function is not an adequate representation of the data. This result mirrors Izadi *et al*'s (2002) findings using the simple stochastic frontier model and suggests that previous estimates of economies of scale and scope may be biased.

Table 6 Descriptive statistics of gross estimated scores

	1995/6	*1996/7*	*1997/8*	*1998/9*	*Average*
Model 21					
Mean	1.247	1.256	1.208	1.199	1.227
Median	1.17	1.2	1.156	1.134	1.169
Std Deviation	0.233	0.207	0.177	0.183	0.202
Minimum	1	1.003	1.001	1	1
Maximum	2.046	1.721	1.663	1.766	2.046
Model 25					
Mean	1.264	1.279	1.23	1.218	1.248
Median	1.189	1.247	1.175	1.17	1.19
Std Deviation	0.241	0.218	0.192	0.186	0.211
Minimum	1.002	1.004	1.006	1.003	1.002
Maximum	2.221	1.832	1.713	1.743	2.221
Model 29					
Mean	1.295	1.304	1.251	1.236	1.271
Median	1.23	1.257	1.189	1.188	1.218
Std Deviation	0.256	0.233	0.204	0.204	0.226
Minimum	1	1.001	1.001	1.002	1
Maximum	2.209	1.902	1.761	1.867	2.209
Correlations †	**1995/6**	**1996/7**	**1997/8**	**1998/9**	**Overall**
Pearson correlation					
Model 21 & 25	0.997	0.997	0.997	0.997	0.997
Model 25 & 29	1.000	1.000	1.000	1.000	1.000
Model 21 & 29	0.998	0.998	0.997	0.997	0.997
Spearman (rank) correlation					
Model 21 & 25	0.994	0.994	0.994	0.994	0.994
Model 25 & 29	0.994	0.994	0.994	0.994	0.994
Model 21 & 29	0.995	0.995	0.995	0.995	0.995

† All correlations significant at the 1 level.

Turning our attention to the efficiency scores themselves, Table 6 shows the descriptive statistics for the gross estimated scores[32]. As one might expect, given the discussion above, the scores are similar in terms of their distribution and are highly correlated with each other. However, it is the net estimated scores (Table 7) that are of particular interest to us, since they represent unexplained inefficiency of an HEI, once we have netted out the effects of differences in the ***E*** variables. We can label this organisational or management inefficiency. As one would expect, these are much lower than the gross scores, but still very highly correlated. Indeed, this high degree of correlation is more indicative of the stability of the predicted efficiency scores than the details of their distribution. However, it is of interest to see how the predicted efficiency scores compare with some of the other models estimated.

[32] We do not include lists of the gross and net efficiency scores themselves for brevity. They are available on request from the author.

Table 7 Descriptive statistics of net estimated scores

	1995/6	*1996/7*	*1997/8*	*1998/9*	*Average*
Model 21					
Mean	1.167	1.175	1.14	1.133	1.152
Median	1.126	1.131	1.102	1.077	1.118
Std Deviation	0.196	0.174	0.148	0.153	0.17
Minimum	0.994	0.997	0.991	0.998	0.995
Maximum	1.834	1.57	1.525	1.59	1.831
Model 25					
Mean	1.205	1.217	1.173	1.168	1.192
Median	1.151	1.182	1.125	1.145	1.139
Std Deviation	0.209	0.189	0.165	0.161	0.183
Minimum	0.996	0.996	1	0.997	0.993
Maximum	2.068	1.689	1.603	1.608	2.076
Model 29					
Mean	1.237	1.244	1.195	1.187	1.214
Median	1.186	1.213	1.152	1.15	1.174
Std Deviation	0.224	0.201	0.177	0.177	0.196
Minimum	0.995	0.997	0.998	0.991	0.994
Maximum	2.067	1.761	1.644	1.715	2.068
Correlations †	**1995/6**	**1996/7**	**1997/8**	**1998/9**	**Overall**
Pearson correlation					
Model 21 & 25	0.997	0.997	0.997	0.997	0.997
Model 25 & 29	1.000	1.000	1.000	1.000	1.000
Model 21 & 29	0.998	0.998	0.997	0.997	0.997
Spearman (rank) correlation					
Model 21 & 25	1.000	0.999	1.000	1.000	1.000
Model 25 & 29	0.999	1.000	1.000	1.000	1.000
Model 21 & 29	0.995	0.995	0.995	0.995	0.995

† All correlations significant at the 1 level.

In Table 8 we present the correlations between the three models with the highest log likelihood as well at the general, encompassing model and the model with the lowest log likelihood, Model 16. We can see that the correlation between the efficiency scores estimated by the 'top three' models and the others are much lower than those between the top three models[33]. Interestingly, the rank correlations with the other models are slightly higher, but are still lower than the correlations between the top three.

Table 8 Correlations

	Gross efficiency scores					*Net efficiency scores*				
	Gen.	21	25	29	16	Gen.	21	25	29	16
Pearson										
Gross efficiency scores										
General	1	0.65	0.652	0.641	0.778	1.000	0.602	0.661	0.665	0.648
Model 21		1	0.997	0.997	0.696	0.617	0.999	0.997	0.998	0.464
Model 25			1	1.000	0.696	0.582	0.998	0.999	0.999	0.435
Model 29				1	0.662	0.643	0.997	0.999	0.999	0.465
Model 16					1	0.765	0.711	0.732	0.748	0.744
Net efficiency scores										
General						1	0.607	0.625	0.618	0.694
Model 21							1	0.997	0.997	0.471
Model 25								1	1.000	0.488
Model 29									1	0.485
Model 16										1
Spearman (Rank)										
Gross efficiency scores										
General	1	0.652	0.652	0.655	0.636	0.99	0.667	0.667	0.668	0.706
Model 21		1	0.994	0.994	0.734	0.632	1.000	0.994	0.995	0.637
Model 25			1	0.995	0.752	0.636	0.998	0.997	0.999	0.661
Model 29				1	0.757	0.637	0.996	0.999	0.999	0.663
Model 16					1	0.62	0.742	0.758	0.763	0.836
Net efficiency scores										
General						1	0.651	0.653	0.654	0.742
Model 21							1	1.000	0.995	0.662
Model 25								1	1.000	0.682
Model 29									1	0.683
Model 16										1

† All correlations significant at the 1 level.

[33] Although we do not include the correlations for all of the models for reasons of space, a similar story emerges when we consider the other models estimated.

Another question of interest is whether there has been any convergence in the efficiency of HEIs. Studies of economic growth that compare the per capita income levels of different countries distinguish two types of growth: β convergence and σ convergence[34]. The former case refers to a tendency for low income countries (inefficient HEIs in our case) to grow (decrease their costs) faster than more efficient ones. This is often referred to as 'regression toward the mean'. The second type of convergence, σ convergence, refers to a tendency for dispersion of inefficiency scores to decline over time. β convergence tends to generate σ convergence, but this process can be offset by new disturbances that tend to increase dispersion. We apply the same approach to our HEI inefficiency measures, looking to see first whether there is regression to the mean and secondly whether inefficiency scores have become less dispersed over time.

In Table 9 we present the results of regressing the change in inefficiency scores, $(\eta_{it} - \eta_{i,t-1})/\eta_{i,t-1}$, on the institution's initial score (i.e. for 1995/6) and a constant. For all three models, we find that the change in inefficiency does indeed depend negatively on initial inefficiency. That is, the greater the inefficiency in a HEI in 1995/6, the more inefficiency is expected to fall. The process is slow, with only 9-10% of the gap between the actual values and the mean being closed on average in any year. However, closer examination of Table 9 shows that this average reflects a fairly large adjustment in 1996/7, a smaller one in 1997/8 and little movement in 1998/9.

[34] See for example, Barro and Sala-i-Martin (1995).

Table 9 β Convergence Regression Results

	Constant		$\eta_{95/6}$		R^2
Model 21					
1996/7	0.217***	(4.046)	-0.163***	(3.850)	0.16
1997/8	0.096*	(1.977)	-0.102***	(2.673)	0.084
1998/9	0.028	(0.501)	-0.026	(0.583)	0.004
Average	0.114***	(5.289)	-0.097***	(5.721)	0.296
Model 25					
1996/7	0.204***	(3.794)	-0.147***	(3.510)	0.136
1997/8	0.077	(1.661)	-0.086**	(2.395)	0.068
1998/9	0.073	(1.379)	-0.062	(1.522)	0.029
Average	0.118***	(5.534)	-0.099***	(5.946)	0.312
Model 29					
1996/7	0.184***	(3.606)	-0.132***	(3.415)	0.13
1997/8	0.076*	(1.749)	-0.086**	(2.590)	0.079
1998/9	0.048	(0.910)	-0.043	(1.086)	0.015
Average	0.103***	(4.890)	-0.087***	(5.464)	0.277

* significant at 10%; ** significant at 5%; *** significant at 1%
Absolute value of t statistics in parentheses

There is also some evidence of σ convergence. As we can see from Table 6 and Table 7, there is a decline in the standard deviation of the inefficiency scores for the first two models model over the entire period and for the second model except for a small increase in academic year 1996/7. This supports the view that there has been convergence in the cost efficiency in English and Welsh HEIs. Moreover, any new disturbances to this process do not appear to have swamped the overall convergence. However, it should be borne in mind that these results are not very strong and only refer to a short time period.

7. Conclusions

We have examined the costs of higher education and found that there is inefficiency in production. Moreover, the costs and inefficiency in higher education institutions can be modelled as a function of their student and staff bodies. Our results suggest that a significant portion of the error term in the cost function is explained by the inefficiency effect. Therefore, any study that seeks to assess the costs of production (i.e. issues such as economies of scale and scope) are liable to produce biased results.

In the absence of a theoretical rationale for deciding whether staff or student characteristics affect costs directly, through the cost function, or indirectly, through inefficiency, we have proposed a selection criterion based on the log likelihood of the estimated model. We estimate 32 models with different combinations of variables appearing either in the frontier or the efficiency determinants and a general encompassing model. Our results point to three models which are preferred to the others. These three models are extremely similar in specification. All suggest that student characteristics should be included in the frontier and staff demographics should be included in the efficiency determinants. The only difference between the models is that one model suggests it most appropriate to include the proportion of staff of professorial or senior lecturer grade in the frontier, another model the proportion who are research active. Our preferred model, the one with the highest likelihood, suggests that it is appropriate to include all of the staff-related variables as determinants of inefficiency.

Although costs have generally risen, output has more so and efficiency does appear to have increased over the period 1995/6 to 1998/9. Finally, the results for our analysis of convergence issues suggest that there has been some convergence in efficiency over the short time period anglicised.

8. References

Aigne,r D.J., Lovell, C.A.K., and Schmidt, P., (1977), 'Formulation and Estimation of Stochastic Frontier Production Function Models', *Journal of Econometrics*, 6, pp. 21-37.

Athanassopoulos A.D., and Shale, E., (1997), 'Assessing the comparative efficiency of higher education institutions in the UK by means of data envelopment analysis', *Education Economics*, 5: 117-134.

Banke,r R.D., and Morey, RC., (1986), 'Efficiency Analysis for Exogenously Fixed Inputs and Outputs', *Operations Research*, 34, pp. 513-21.

Barro, R.J., and Sala-i-Martin, X., (1995), *Economic Growth*, McGraw-Hill: New York.

Battese, G.E., and Coelli, T.J., (1993), 'A Stochastic Frontier Production Function Incorporating a Model for Technical Inefficiency Effects', Working Papers in Econometrics and Applied Statistics, No. 69, Department of Econometrics, Armindale, NE: University of New England.

Battese G.E., and Coelli T.J., (1995), 'A Model for Technical Efficiency Effects in a Stochastic Frontier Production Function for Panel Data', *Empirical Economics*, 20, pp. 325-32.

Battese, G.E., and Corra, G.S., (1977), 'Estimation of a Production Frontier Model: With Empirical Applications in Agricultural Economics', *Agricultural Economics*, 7, pp. 185-208.

Baumol, W.J., Panzar, J.C., and Willig, D.G., (1982), *Contestable Markets and the Theory of Industrial Structure*, New York: Harcourt Brace Jovanovich.

Bee, M., and Dolton, P.J., (1985), 'Degree Class and Pass Rates: An Inter-University Comparison', *Higher Education Review*, 17, pp. 45-52.

Bowen, H.R., (1991), *The Cost of Higher Education: How Much do Colleges and Universities Spend Per Student and How Much Should The Spend?*, Josey Bass: San Francisco:

Brinkman, P.T, and Leslie, L.L., (1986), 'Economies of scale in higher education: Sixty years of research', *Review of Higher Education*, 10, pp. 1-28.

Cave M, Weale M., (1992), 'The Assessment: Higher Education: The State of Play', *Oxford Review of Economic Policy*, 8, pp. 1-18.

Cohn, E., Rhine, S.L.W., and Santos, C., (1989), 'Institutions of Higher Education as Multi-Product Firms: Economies of Scale and Scope', *Review of Economics and Statistics*, 71, pp. 248-90.

Coelli T.J., (1995), 'Estimators and hypothesis tests for a stochastic frontier function: A Monte Carlo analysis', *Journal of Productivity Analysis*, 6, pp. 247-68.

Coelli T.J., (1996), 'A Guide to FRONTIER Version 4.1: A Computer Program for Frontier Production Function Estimation', CEPA Working Paper 96/07, Department of Econometrics, University of New England, Armindale.

Coelli, T., Perelman, S., and Romano, E., (1999), 'Accounting for Environmental Influences in Stochastic Frontier Models: With Application to International Airlines', *Journal of Productivity Analysis*, 11, pp. 251-73.

Coelli, T.J., Rao, D.S., and Battese, G.E., (1999), 'An Introduction to Efficiency and Productivity Analysis', Kluwer Academic Publishers: North Holland.

DeGroot, H., McMahon, and W.W., Volkwein, J.F., (1991), 'The Cost Structure of American Research Universities', *Review of Economics and Statistics*, 73, pp. 424-31.

Disney, R., (1999), 'Why have older men stopped working?', in (P. Gregg and J. Wadsworth (eds.) *The State of Working Britain*, pp. 58-74.

Dolton, P.J., (1986), 'The Prediction of Degree Performance: A Statistical Analysis', University of Hull, Department of Economics and Commerce, Discussion Paper 133.

Earl, J., (2001), 'Further Guidance on the Staff Individualised Record 1999/00 (C99021)', Staff Circular, HESA.

Farrell, M.J., (1957), 'The Measurement of Productive Efficiency', *Journal of the Royal Statistical Society Series A*, 120, pp. 253-90.

Getz M, Siegfried JJ, Zhang H., (1991), 'Estimating economies of scale in higher education', *Economics Letters*, 37, pp. 203-8.

Glass, J.C., McKillop, D.G., and Hyndman, N.S., (1995a), 'The achievement of scale efficiency in UK universities: a multiple-input multiple-output analysis', *Education Economics*, 3, pp. 249-63.

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

Hare, P, and Wyatt, G., (1992), 'Economics of Academic Research and its Implications for Higher Education', *Oxford Review of Economic Policy*, 8, pp. 48-66.

Hashimoto, K., and Cohn, E., (1997), ‘Economies of scale and scope in Japanese private universities’, *Education Economics*, 5, p. 107-16.

Huang, C.J., Liu, J.T., (1994), ‘‘Estimation of a Non-Neutral Stochastic Frontier Production Function’, *Journal of Productivity Analysi*s, 5, pp. 171-80.

Izadi, H., Johnes, G., Oscrochi, 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, pp. 63-72.

Johnes G., (1992), ‘Performance Indicators in Higher Education: A Survey of Recent Work’, *Oxford Review of Economic Policy*, 8, pp. 19-34.

Johnes G., (1987), ‘Costs and Industrial Structure in Contemporary British Higher Education’, *Economic Journal*, 107: 727-37.

Johnes G., (1996), ‘Multiproduct cost functions and the funding of tuition in UK universities’, *Applied Economic Letters*, 3, pp. 557-61.

Johnes, G., (1998), ‘The costs of multi-product organisations and the heuristic evaluation of industrial structure’, *Socio-economic Planning Sciences*, 32, pp. 199-209.

Johnes, J., and Taylor, J., (1997), ‘Degree Quality: An Investigation into Differences between UK Universities’, *Higher Education*, 16, pp. 581-602.

Jorgenson DW., (1986), ‘Econometric Methods for Modelling Producer Behaviour’. In *Handbook of Econometrics*. Griliches Z, Intriligator MD (eds), ‘Elsevier, Amsterdam.

Kalirajan, K.P., (1981), ‘An Econometric Analysis of Yield Variability in Paddy Production’, *Canadian Journal of Agricultural Economics*, 29, pp. 283-94.

Koshal, R.K., and Koshal, M., (1999), ‘Economics of scale and scope in higher education: a case of comprehensive universities’, *Economics of Education Review*, 18, pp. 270-7.

Koshal, R..K., Koshal, M., and Gupta, A., (2001), ‘Multi-product total cost function for higher education: a case of bible colleges’, *Economics of Education Review*, 20, pp. 297-303.

Kumbhakar, S.C., Ghosh, S., and McGuckin, J.T., (1991), ‘A Generalized Production Frontier Approach for Estimating Determinants of Inefficiency in U.S. Dairy Farms’, *Journal of Business and Economic Statistics*, 9, pp. 249-86.

Lovell, C.A.K., and Schmidt, S.S., (eds.), (1993), *The Measurement of Productive Efficiency: Techniques and Applications*, Oxford University Press: Oxford.

Meeusen, W., and van den Broeck J., (1977), 'Efficiency Estimation from Cobb-Douglas Production Functions With Composed Error', *International Economic Review*, 18, pp. 435-44.

Naylor, R., and Smith, J., (2001), 'Determinants of Degree Performance in UK Universities: A Statistical Analysis of the 1993 Student Cohort', *Oxford Bulletin of Economics and Statistics*, 63, pp. 29-60.

Nelson, R., and Hevert, K.T., (1992), 'Effect of class size on economies of scale and marginal costs in higher education', *Applied Economics*, 24, pp. 473-82.

Pitt, M.M., and Lee, L-F., (1981), 'Measurement and Sources of Technical Inefficiency in the Indonesian Weaving Industry', *Journal of Development Economics*, 9, pp. 43-64.

Reifschnieder, D., and Stevenson, R., (1991), 'Systematic Departures from the Frontier: A Framework for the Analysis of Firm Inefficiency', *International Economic Review*, 32, pp. 715-23.

Staat, M., (1999), 'Treating non-discretionary variables one way or the other: Implications for efficiency scores and their interpretation', In G. Westermann (ed.), *Data Envelopment Analysis in the Service Sector*, 23-50, Gabler: Wiesbaden.

Stevens, P.A., (2001), 'The determinants of economic efficiency in English and Welsh universities', NIESR Discussion Paper No. 185.

Stevens, P.A., and Vecchi, M., (2002), 'The effect of environmental variables on public service efficiency', Presentation to the British Academy, 19th November 2002.

Verry, D, and Davies, B., (1976), *University Costs and Outputs*, Elsevier: Amsterdam.

Worthington, A.C., (1999), 'Performance indicators and efficiency measurement in public libraries', *Australian Economic Review*, 32, pp. 31-42.