Lancaster University Management School
Working Paper
2009/013

Cost Structure, Efficiency and Heterogeneity in US Higher Education: An Empirical Analysis

T Agasisti and Geraint Johnes

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

COST STRUCTURE, EFFICIENCY AND HETEROGENEITY IN US HIGHER EDUCATION: AN EMPIRICAL ANALYSIS

Tommaso Agasisti

Politecnico di Milano
Department of Management, Economics and Industrial Engineering,
p.za Leonardo da Vinci, 32, 20133 Milano
Italy
T:+39 02 2399 3963
F: +39 02 2399 2720
E: tommaso.agasisti@polimi.it

Geraint Johnes

Lancaster University Management School
Lancaster LA1 4YX
United Kingdom
T: +44 1524 594215
F: +44 1524 594244
E: G.Johnes@lancs.ac.uk

This version - March 2009

Abstract. We estimate a variety of models to evaluate costs in US higher education institutions. A novel feature of our approach involves the estimation of latent class and random parameter stochastic frontier models of the multiproduct cost function. This allows us fully to accommodate both the heterogeneity of institutions and the presence of technical inefficiencies. Our findings suggest that global economies could be achieved by effecting a reduction in the number of institutions providing undergraduate instruction, while increasing the number of institutions engaged in postgraduate activity.

Keywords: costs, efficiency, stochastic frontier, latent class, random parameter.

JEL Codes: C14, C23, H52, I21, I22.

1. Introduction

Multi-product organizations have been the subject of much research since the seminal contribution of Baumol *et al*., (1982). Like countries, such organizations enjoy different resource endowments and so specialize in the production of different outputs. The structure of costs in these organizations likely differs across firms, and this heterogeneity gives rise to some interesting issues surrounding the evaluation of organizational performance.

In this paper, we focus on one industry in which data on multi-product organizations are readily available, namely higher education. We ask the question: do similar universities have the same cost structure? This is a crucial question for both theoretical and practical reasons. The policy implications of the answer are dramatically important for policy makers, especially in periods when pressures on public funds call for reflection on the organization and financing of such strategic sectors as higher education. For instance, different cost functions imply different scale and scope effects, which can justify the application of different policies to clusters of institutions in the sector, and hence the application of different funding models. For instance, the efficiency of institutions in using available resources matters for decisions about the distribution of public funds. Moreover, if the cost functions attached to different institutions are different, then this has implications for the evaluation of their efficiency.

This topic is particularly relevant with reference to the higher education (HE) system in US, where there is a multitude of institutions of higher education (IHEs), with different missions, activities, strategies, and organizational structure. For instance, De Groot *et al.* (1991) state that they *"account for research productivity differences between institutions"* (p.424). Harter *et al.* (2005) study the costs of 4-years US colleges, and acknowledge that *"expenditures per student showed significant variance across institutions having the same mission and control"* (p. 371). This issue is so evident that Sav (2004) has conducted a specific analysis to check whether there is a difference between the cost structure of public and private universities, finding that actually there are some quite remarkable differences.

Heterogeneity across IHEs has been a a major focus of research also in Europe. In this context, universities are often deemed to be similar in character to one another, and a strong differentiation across institutions (typical of the US system) does not occur. In a break with this tradition, Johnes *et al.* (2005) study the higher education sector in UK, and they detect significant differences across different subsectors. However, it is only recently that the problem of correctly identifying the heterogeneity across universities' cost functions has been systematically tackled using newly developed estimation technology. The idea proposed by Johnes and Johnes (2009) is that heterogeneity occurs not only (or not necessarily) between universities in different subsectors, but also it is possible between universities that belong to the same subsector of higher education. This analysis has been made possible by using the new techniques proposed by Tsionas (2002) and Greene (2005); these involve a random parameter specification of models, which will be discussed later in this paper. After this study, similar analyses were conducted for other European countries: Italy (Agasisti and Johnes, forthcoming), Spain (Johnes and Salas-Velasco, 2006), and Germany (Johnes and Schwarzenberger, 2007). In all these countries, the higher education institutions are conventionally thought of as being very similar to one another, while the analysis showed that relevant differences actually exist.

Recent empirical work conducted by Bonaccorsi and Daraio (2007) focused on different strategies pursued by European universities, and they exploit several differences across institutions. They conceptualized the concept of university strategy as *"(…) an emergent pattern of configuration of university outputs that depend on (relatively) autonomous decisions making by universities, supported by appropriate combinations of resources (inputs)"* (Bonaccorsi and Daraio, 2007; p. 11).

Coming back to the main focus of this paper, the HE sector in the US is traditionally characterized by a perception of strong differences across institutions. Part of the difference is due to different unit costs. Harter *et al*. (2005) conducted an analysis on four-year public colleges for the period 1989-1998, showing a great variation of cost per student - where students have been measured on a FTE (full-time equivalent) basis. Their analysis showed that, for colleges with an enrolment of between 3,000 and 10,000 FTEs, the range in real unit costs in 1989 was between \$8,144 for comprehensive public colleges without a medical school and \$17,538 for research colleges with medical schools. Another source of difference is in the output mix, that is the vector of outputs produced: some universities are more focused on research, others on undergraduate teaching, others on postgraduate teaching. Since 1970 the

Carnegie Foundation, aware of such heterogeneity of IHEs' mission, has proposed a classification of colleges according to different dimensions (http://www.carnegiefoundation.org/classifications/): Undergraduate Instructional Program, Graduate Instructional Program, Enrollment Profile, Undergraduate Profile, Size and Setting. A basic classification is also proposed considering the following categories of colleges: associate's colleges, doctorate-granting universities, masters colleges and universities, baccalaureate universities, special focus institutions, tribal colleges. Recently, differences across institutions (in terms of performance, output levels and reputation) have become key in determining the choice of students and families, not least because of the proliferation of public rankings of institutions such as the US News and World Report. The 2009 Report states: *"(…) the rankings provide an excellent starting point because they offer the opportunity to judge the relative quality of institutions based on widely accepted indicators of excellence. (…) Schools are categorized by mission, based on the basic Carnegie classification and, in some cases, by region"* (p. 82). It is reasonable to assume that such an accepted heterogeneity across institutions could impact also on their cost structure – as a pure analysis of unit costs indeed confirms.

The goal of this paper is to analyze the cost structure of HE sector in the United States. It represents a particularly interesting case study, given the high number of universities and colleges that are playing in the market and the huge differences amongst them – in terms of dimension, types of education and research provided, ownership, subject mix, and so on. We propose to employ here empirical approaches that exploit institutional heterogeneity in the calculation of the parameters of the cost function. Thus, the empirical analysis which follows uses and compares three different estimators of the universities' cost functions: (1) the traditional stochastic frontier cost model, (2) the latent class stochastic frontier model and the (3) random parameter stochastic frontiers specification. The methodology adopted in this paper allows us jointly to analyze costs and efficiency. This is the straightforward way to proceed, as the cost function represents the boundary that describes the lowest cost at which it is possible to produce a given vector of outputs. In this context, frontier estimations permit us to describe the cost structure and to compute efficiency scores simultaneously. The estimation is focused upon the cost function, rather than the production function, because IHEs potentially produce several outputs at the same time.

The paper is organized as follows. In the next section, a literature review is provided. Section 3 describes the background for the subsequent analysis. Section 4 concerns the data and methodology used for this study. Section 5 presents the results and a discussion. Section 6 concludes.

2. Literature review

Recent theoretical work has shown that universities, in a context of limited resources and competition, tend to specialize (Del Rey, 2001), and that the equilibrium of the HE market converges towards different steady states, where mass production of both research and learning is just one of four possible strategies of universities. The others are: full-time teaching, full-time research, selective teaching and research. Similar results have been obtained by De Fraja and Valbonesi (2008). In light of these findings, it is natural to suppose that a number of different types of HEI might exist, and that these should be characterized by distinct cost structures.

The empirical investigation of the cost structure of universities is not a new theme in the United States. Early work is surveyed by Brinkman and Leslie (1986), and points to the existence of widespread scale effects. The first paper to introduce the concept of universities as multiproduct organizations, and hence to focus upon more sophisticated measures of economies of scale and scope, is due to Cohn *et al.*(1989). After this pioneering work, several contributions aimed to provide further empirical evidence about the cost structure of the higher education sector.

In general terms, the approach adopted in this field is to estimate a cost function of the following type:

$$TC_i = f(\mathbf{x}_i) \qquad (1)$$

where TC_i is the total cost for the i_{th} university, and $\mathbf{x}_i$ is a vector of outputs. It is widely recognized that universities typically produce three kinds of output: teaching, research and social service (the last of these typically taking the form of knowledge transfer, organization of cultural events, consultancy, and the like). Nevertheless, given the difficulty of identifying proper measures for the "third mission" work of this kind, estimated cost functions usually include only measures of teaching and research; indeed only a very few attempts have been made to include proxy for the "third mission" (an exception is Johnes *et al*., 2005).

The contribution of Cohn *et al*., (1989 employs a multi-output cost function measuring both teaching and research outputs. In practice, they measure full-time-equivalent enrolments as teaching output, and use measures of research grant income as a proxy for research output. Faculty salaries are included as a price factor. Their sample comprises 1,887 IHEs. The results of this pioneering study are that: (i) institutions can benefit from scope economies by raising the output level; (ii) comprehensive institutions are less costly than specialized ones; and (iii) very-small institutions are more costly than average ones. All of these findings suggest that an important role is played by scale economies in the higher education sector.

De Groot *et al*. (1991) considered 147 doctorate-granting universities. The variables considered are FTE enrolments of undergraduate and graduate students, and the number of research publications. Their results indicate the presence of sizeable economies of scale for the average institution, as well as scope economies associated with the joint production of undergraduate and graduate education. At the same time, they did not find evidence of any significant impact of ownership and state regulation on costs. The results have been validated also through a sensitivity analysis by substituting enrolments with degrees awarded.

Koshal and Koshal (1999) focus on comprehensive institutions. Their sample contains 158 private and 171 public IHEs. FTE enrolments and dollars spent for research activity are used as indicators for teaching and research output. Their analysis is unusual as they included a measure of quality for the teaching and learning domain, that is the average total scores on the Scholastic Aptitude Test (SAT) of entering freshmen.[1] Product-specific economies of scale are detected for undergraduate education but are absent in the case of graduate education; the authors also find that economies of scope remain unexploited.

Laband and Lentz (2003) work on the Cohn *et al*. (1989) framework, estimating cost functions for 1,492 private and 1,450 public IHEs. They use enrolments and

---

[1] Johnes et al. (2005) also include a quality measure in their work. Collinearity problems often preclude this.

externally funded research as proxies for teaching and research output, respectively. The results report a significantly different cost structure for private and public institutions. There are two common features across these two sectors: the presence of economies of scale, and diseconomies of scope. The straightforward policy implication of such findings is that global unit costs could be reduced by increased specialization of institutions, while raising output levels in each university.

Sav (2004) conducted a study on an extensive sample of 2,189 universities and colleges. In his cost function, he includes *inter alia* a wage variable as a factor price, and a dummy for the presence of a medical school. Teaching output is captured by a measure of teaching hours, for undergraduate, graduate and professional courses; research output has been measured through the research grants. The author reports ray economies of scale in private, but not in the public, sector. The private sector also benefits from economies of scope. Moreover, both private and public universities show product-specific economies of scale attached to research activity. The paper derives some strong and important policy implications: that *"(…) instead of trying to be everything to everyone or all taxpayers, large state-supported research institutions might move to more specialized production"* (p. 613). Also regional differences are detected, by suggesting a possible differentiation of cost functions due to the socio-economic differences of the regions (in particular, the wage factor could be affected by such differences).

Laband and Lentz (2004) devote further attention to the differences between private and public sectors, by analyzing a sample of more than 2,700 universities and colleges. Also this paper uses FTE enrolments as teaching output and research grants as research output, and a measure of faculty compensation has been included as a price factor. The results confirm the hypothesis that the two subsectors actually are quite different in terms of the cost structure of the institutions. More specifically, *"(…) the public IHEs produce more cheaply than the private, not-for-profit IHEs at almost all scales of output"* (p. 438).

3. Background

The use of random parameters for modeling the cost function of universities is quite recent, even tough it was already utilized for empirical analyses in UK (Johnes and Johnes, 2009), Italy (Agasisti and Johnes, forthcoming), Spain (Johnes and Salas-Velasco, 2006), and Germany (Johnes and Schwarzenberger, 2007). The idea behind these studies is that colleges tend to be different, and so they each face a cost function that is distinct. In the same spirit, recent work by Bonaccorsi and Daraio (2007) demonstrates that, even if is there such presumption of homogeneity, universities already are very different in terms of productivity, performance and efficiency. More specifically, they developed the idea of “strategy” of different universities, that is each university tends to specialize (voluntarily or not) towards the activity in which it has a competitive advantage.

In the US popular perceptions on topic are quite different, as it is widely accepted that universities behave in a different manner from one another, with some well defined segments of the market leading to a number of ‘types’ of institution, each of which specializes to a greater or lesser extent on the provision of certain types of output. A clear source of heterogeneity in institutions’ behavior is in the specialization towards teaching or research, and, within teaching, they might specialize in undergraduate or graduate activities.

While technology might explain how specialization can lead to allocative efficiencies, it is difficult in a world where IHEs face homogeneous cost structures to see how some institutions choose to specialize in the production of some outputs while others do otherwise. A possible way out of this conundrum is to acknowledge the possible existence of heterogeneity in the cost structure of universities. In other words, it would be possible that institutions differ not only in terms of size – with the implied differences in scale and scope effects – but also because of inter-institutional differences in structural factors that affect their performance.

In much of the received literature, cost functions have been estimated in a parametric setting that *assumes* that all institutions face the same cost function. Yet we know from theory, and from studies that have started from the identification of prescribed categories of institution, that they do not. Put another way, the economic literature has pointed out that we should expect there to be heterogeneity across institutions, and that this should be captured in the institution-specific effects associated with their operations. So our models should allow for the possibility that different institutions have different cost functions. A fairly crude way of representing this is to use fixed effects models. However, the recent development of estimators proposed by Greene (2005) and Tsionas (2002) allow us to identify variation not only in the constant, but also in the parameters of the cost equation, so that each HEI faces a different cost function – all within the context of a frontier model. This approach resembles in key respects the nonparametric approach of data envelopment analysis, where the weights attached by an institution to each input and output are allowed to differ from those applied by other institutions. The random parameters approach is well explained by Tsionas (2002, p. 128): *"(…) production possibilities are expected to differ in a cross-section of firms, and a set of different technologies may simultaneously coexist at any given time. If that is the case, efficiency measurement cannot proceed under the assumption of common technology. (…) The relative difference in output reflects technological differences, not inferior practice"*.

This argument is particularly convincing for the HE sector. While it is reasonable to assume the same functional form for the cost function of all institutions, it is less reasonable to impose the assumption that the parameters are constant across institutions. Use of panel data allows a random parameter specification to be modeled. From a methodological point of view, such an approach benefits from generality in the ability of modeling heterogeneity – in other words, *"many of the models already considered (such as fixed-effects of random-effects specifications) are special cases"* (Greene, 2005; p. 288). The assumption of the same technology for all institutions results in a *"(…) confusion between technological differences and technology-specific inefficiency"* (Tsionas, 2002; p. 128).

As we expect that there actually are relevant technologies differences and behavioral differences across colleges in their production processes of teaching and research, we estimate cost functions using a random parameter approach, and we compare the derived results with other more classical approaches (the traditional frontier model and latent class models).

4. Data and methodology

All the data come from the Integrated Postsecondary Education Data System (IPEDS) dataset, provided by the National Center for Educational Statistics (NCES), and refer to three academic years: 2003-04, 2004-05, and 2005-06. All the financial variables are collected on a financial year basis, and they are matched with the academic year that starts in this year. Hence costs in year t have been matched with students in the academic year that straddles t and t+1.

Our sample comprises all the universities classified as four (or more) year degree-granting institutions, both public and private, making a total of some 2,318 institutions. For the purposes of empirical analysis, however, we drop all the observations for which there are missing data for any variables in any one of the three years considered. This results in dropping 1,364 institutions, and the final sample includes 954 institutions for which we have three complete years of data, that is a total of 2,862 observations.

The dependent variable (COSTS) is defined as total amount of expenses, in the sense of "*outflow or other using up of assets or incurrence of liabilities*" (IPEDS definition).[2]

The independent variables related to the teaching and learning activities are defined as follows:

- the number of bachelor (b) degrees. A qualification at bachelor level represents the first completed level of higher education, which includes degrees obtained after 4 or (less usually) 5 years of study;
- the number of postgraduate (p) degrees. It includes the number of students who obtained a first professional degree (in, for example, medicine), a masters degree, or a doctorate.

These qualifications are clearly different from one another – first professional degrees are somewhat akin to unusually long bachelors programmes, taking students from entry to higher education through exit at higher degree level. Masters programmes are typically of relatively short duration – one or two years. Doctorates, meanwhile, combine a taught component with a requirement for substantive research activity, and typically take several years to complete. In early work, we separated out these various types of qualification, but as a result of multicollinearity problems we have merged them into a single 'postgraduate' variable for the purposes of the present paper. While many studies use the number of students (a stock) as an indicator for teaching activity, here we use the number of graduates (an outflow) because it better represents the final output of the teaching. In comparing the results provided by our study with other previous contributes, this difference should be kept in mind.

[2] The definition is slightly different, albeit equivalent, for public universities on the one hand and private universities on the other. This is because accounting principles for the former follow the guidelines of the Governmental Accounting Standards Board while those for the latter follow the guidelines of the Financial Accounting Standards Board.

As a proxy for research activities we use the value of grants received for conducting research. More precisely, this variable is the sum of *"all operating expenses associated with activities specifically organized to produce research outcomes, and commissioned by an agency either external to the institution or separately budgeted by an organizational unit within the institution"* (IPEDS website). The issues surrounding the use of this type of indicator as a proxy for research output are known and debated (De Groot et al., 1991; Johnes and Johnes, 1993); while grants may be regarded as an input, they offer several advantages as a measure of research activity. Notably they provide a quality adjusted measure of the volume of research that is done, and they offer a contemporaneous measure that has advantages over more retrospective alternatives such as citations and publication counts. Grants have therefore been used to proxy research in a number of earlier studies, amongst which are Cohn *et al*. (1989), Laband and Lentz (2004) and Sav (2004). Summary statistics for our sample of institutions are reported in Table 1, and clearly show strong heterogeneity among universities – the standard deviation is higher than the mean for all the variables considered.

<Table 1> around here

This heterogeneity can also be observed when some simple bivariate scatterplots are constructed. Figure 1 plots the number of bachelor students against the value of research grants earned by institutions in the sample. At least three patterns emerge: (1) a group of colleges focused on teaching activities, with a well below average level of research output; (2) a group of colleges with high levels of both teaching and research activity; and (3) a group of colleges strongly focused on research, with low levels of bachelor enrolments. Similar evidence of institutions with heterogeneity of missions derives from the plot of the number of bachelor students against postgraduates show in Figure 2. Again, three groups can be identified: (1) a group of colleges with a "balanced" level of bachelor and postgraduate students; (2) a group with above-average levels of bachelors and below-average levels of postgraduates; (3) a group of colleges focused on postgraduate teaching.

<Figures 1,2> around here

Noting this heterogeneity, we proceed to employ an appropriate modeling technology, specifically, using new panel data techniques developed by Tsionas (2002) and Greene (2005) in the context of the evaluation of a stochastic cost frontier. The traditional frontier estimation, as developed by Aigner et al., (1977), specifies an equation like the following (in a panel setting):

$$y_{it} = \alpha_i + \beta' x_{it} + v_{it} + u_{it} \qquad (2)$$

where v_i denotes a normally distributed residual and u_i is a non-normal residual which is supposed to capture technical inefficiency. The distribution of u_i must be specified *a priori*, and is usually assumed to be half-normal or exponential.

An elementary way to consider heterogeneity across institutions is to cluster observations into different groups or classes, which have some common characteristics that might explain the cost differentials. Where there are no strong *a priori* grounds on which to base the construction of these classes, or when there is a preference to 'let the data speak', the approach should be to estimate a latent class model (LCM) (Orea and Kumbhakar, 2004). This approach divides observations, on the basis of maximum likelihood, into *m* classes (where *m* is prescribed by the analyst), and estimates distinct parameter vectors for each of the *m* classes. The specification of a LCM is:

$$y_{it} = \alpha_{it} + \beta_m x_{it} + v_{it,m} + u_{it,m} \qquad (3)$$

It is important to note that the calculation of efficiency is not conducted with respect to the whole sample, but conditionally on the basis of the class to which the unit belongs.

An extreme case of the LCM is one in which *m* equals N, the number of institutions; in this case, each institution has its own distinct cost function. Here, the vector of coefficients is random across institutions – the case of a random parameters model (RPM). In this case, parameter heterogeneity is modeled as follows:

$$\left.\begin{array}{l}(\alpha_i,\beta_i) = (\overline{\alpha},\overline{\beta}) + \Gamma_{\alpha,\beta} w_{\alpha,\beta i} \\ \mu_i = \overline{\mu} + \Gamma_{\mu} w_{\theta i} \\ \theta i = \overline{\theta} + \Gamma_{\theta} w_{\theta i}\end{array}\right\} \qquad (4)$$

The random variation appears as the vector $w_{j,i}$ where j denotes the constant or the slope parameter. The terms μ and θ represent the moments of the inefficiency distribution. It is possible to model the RPM either with half-normal residuals for each institution constrained to be constant across time, or with these residuals unconstrained. In the work that we report below the residuals are unconstrained, but we note that this makes little difference to our results. In the sequel we estimate (i) a traditional frontier model, (ii) a LCM in which the number of classes is constrained to be 2, and (iii) a RPM. The results are reported and discussed in the next section.

A choice must be made about the functional form of the cost equation. The development of a theory about the functional form in the case of multiproduct organizations is due to Baumol *et al.* (1982) who argue that the cost function of a multi-product firm should meet a number of requirements. Firstly, it must be nonnegative, nondecreasing, concave and linearly homogenous in input prices. In our analysis, we do not use input prices as an explanatory variable since earlier contributions to the literature have shown that the influence of such factor prices is typically not or weakly significant (Cohn *et al*., 1989; Laband and Lentz, 2003, 2004). We argue that allowing heterogeneity across universities' cost structures better captures such differences given their limited impact. Secondly, cost functions must allow sensible predictions to be made for the costs of institutions that produce zero levels of some outputs. Thirdly, the function must not be linear, because it should allow for economies of scale or scope. Following these considerations, and in line with much of the literature in this area (Cohn and Cooper, 2004), we employ a quadratic form. More specifically, the cost function that has been estimated is:

$$y = \alpha + \Sigma\beta_i x_i + \Sigma_i\Sigma_j \varphi_{ij} x_i x_j + v + u \qquad (5)$$

where y is the IHE's cost, x_i and x_j are the outputs of type i and j, respectively. The quadratic terms allow for scale economies; the interaction terms allow for scope effects.

Finally, in our approach we use measures of average costs, marginal costs, scale and scope effects that are coherent with the multiproduct nature of the institutions. Baumol *et al*. (1982) define the average incremental cost (AIC) associated with product *k* as

$$AIC(y_k) = \left[C(y_K) - C(y_{K-k})\right] / y_k \qquad (6)$$

where $C(y_K)$ is the cost of producing the outturn output vector, and $C(y_{K-k})$ is the cost associated with producing the outturn values of all outputs other than the *k*th output, and where the output of type *k* is zero. Defining $C_k(y)$ as the marginal cost of the kth output, we can then define product-specific returns to scale associated with the *k*th output as

$$S_k(y) = AIC(y_k) / C_k(y) \qquad (7)$$

This definition is therefore analogous to the ratio of average to marginal costs that is often used as a measure of scale economies in single product contexts. A value of $S_k(y)$ that exceeds unity reflects product-specific returns to scale that are increasing, and *vice versa*.

Ray returns to scale (S_R), which capture scale effects associated with a simultaneous and proportional change in all outputs, may be calculated as

$$S_R = \frac{C(y)}{\sum_k y_k C_k(y)} \qquad (8)$$

A value of S_R exceeding unity indicates that a simultaneous proportional increase in the production of all output types results in economies of scale, while a value less than one indicates decreasing returns to scale.

Global economies of scope are calculated using the formula

$$S_G = \left[\sum_k C(y_k) - C(y) \right] \Big/ C(y) \qquad (9)$$

where $C(y_k)$ is the cost of producing only the outturn value of kth output, with zero output of all other types. This formula therefore compares, in the numerator, the cost of producing the outturn output vector in a single institution with that of producing the same output in several different, single-product, institutions. If S_G is positive, then it is cheaper to produce jointly than not, and so economies of scope are said to exist. Conversely, $S_G < 0$ implies diseconomies of scope.

Product-specific returns to scope associated with output of type k can analogously be defined as

$$SC_i = \left[C(y_k) + C(y_{N-k}) - C(y) \right] / C(y) \qquad (10)$$

5. Results and discussion

Table 2 contains the regression results. The table contains four columns: in the first, we report the estimates for the traditional frontier model; in the second and third, the estimates for the LCM (here, two sets of coefficients have been reported due to the assumption of two classes); in the final column, the mean parameters of the RPM model are presented.

<Table 2> here

The regression results are difficult to interpret owing to the nonlinear nature of the model. A more useful picture emerges from the analysis of Average Incremental Costs (AIC), Marginal Costs (MC) and product-specific scale effects calculated by using the formulae discussed in the previous section. These results are reported in tables 3, 4 and 5. The calculations have been made for an average institution – that is considering the means of outputs reported in table 1 as the hypothetic output vector – and for institutions that produce 75% and 125% mean output levels.

<Tables 3, 4 and 5> here

Several features of the results are worth commenting upon. First, the average incremental cost estimates for postgraduates are (with the exception of large institutions under the RPM) estimated to be lower than those for undergraduates. This may reflect in part the long duration of bachelors programs in comparison with masters programs, and also the fact that the cost attached to the production of doctoral graduates is mitigated by the fact that students at this level often contribute to a university by providing teaching assistance.

The costs attached to postgraduate education vary markedly with the estimation technology. They are lowest (implausibly low, we think) in the stochastic frontier model, which does not allow for any parameter variation across institutions, and are highest in the random parameter stochastic frontier model. As might be expected in view of the fact that it resembles a methodological halfway house, the results for the latent class model lie somewhere between the two extremes. The sensitivity of average incremental costs of postgraduate provision with respect to estimation method is likely due to the fact that many institutions produce few or no postgraduates, and these institutions bias the coefficients on postgraduates in the fully parametric specification.

Further results of interest include the findings that there are scale economies for undergraduate teaching (particularly in the RPM specification) while there are diseconomies of scale associated with postgraduate provision. Economies of scale associated with research activity, meanwhile, are virtually exhausted. A note of caution is needed here, however, in that all the calculations have been made considering an “average” institution, which is a college producing mean values of the output vector. Such an institution does not exist, because the reality is a sector very differentiated where colleges tend to be more specialized towards a certain output. This being the case, our results must be interpreted as illustrating the scale effects for an average institution, while recognizing that differences from the average are the norm.

A further interesting result concerns the estimation of efficiency scores. The traditional frontier model postulates the existence of one efficiency frontier, and the efficiency of each college is computed as the distance from it. The latent class specification employed here considers the presence of two frontiers, one for each group of institutions. We expect in this case that the average efficiency of the sector should be higher than is the case where a single cost function is assumed to apply to all institutions, since the reference frontier is, under the latent class method, calculated with respect to similar competitors. Finally, the RPM assumption is that colleges must be compared with their own (potential) cost performance, and then the efficiency mean should be even higher than in the latent class case. In figures 1, 2 and 3 the frequency distribution of efficiency scores are reported, and they confirm our expectations[3].

<Figures 3, 4 and 5> around here

The distribution is widespread across scores in the first case (traditional frontier model), suggesting that when considering a unique frontier many colleges result are deemed to be highly inefficient. The concentration of scores below 0.5 confirms this. Moreover, the number of highly efficient institutions is very limited. The picture dramatically changes when we turn to the latent class specification. Allowing each college to compare its efficiency with the frontier for its own group markedly improves the overall measured performance of institutions. Indeed, in this case efficiencies are concentrated between 0.6 and 0.9, suggesting a good efficiency level of the market. The latent class specification accounts for a structural difference between two subsectors of HE market. Looking at mean output in the two classes, the difference may be primarily due to different scales of operation. Table 6 helps in explaining this point: large universities (with outflows of more than 2,000 undergraduates and 1,000 postgraduates, and with research income in excess of $100m) belong to the one group (group 1), while the remaining universities belong to the other. Thus, estimated costs are not radically different across institutions within each group, and the scale effects are also similar. What is different is the scale of operation: both groups have now found their dimension for their activities – as the estimated scale effects demonstrate – but now the two segments of the market are

[3] Some colleges' efficiency scores have been estimated as <0, due to their distance from the frontier. Obviously, it is a traditional shortcoming for a parametric approach to these topics. We eliminated scores <0 to derive our main results.

consistently separated, so it makes sense to compare each institution's efficiency with its counterparts in the submarket.

<Table 6> around here

The last case to be analyzed is that of RPM specification. Here the distribution indicates a concentration of efficiencies above 0.8, though there remain significant numbers of institutions with lower levels of measured efficiency. It is worth noting here that the comparison of each college's cost efficiency is not conducted with its competitors in a submarket, but within the whole market.

It is interesting to analyze whether the efficiency scores reflect subsectors of the HE market. For this purpose, we consider four different types of colleges: (1) private institutions with medical schools, (2) private institutions without medical schools, (3) public institutions with medical schools, and (4) public institutions without medical schools. The aim of such categorization is to investigate whether different institutional characteristics have an impact on the average efficiency. Table 7 reports, in the first row, some descriptive statistics for the entire sample, while the subsequent rows illustrate corresponding descriptive statistics for the subgroups. This analysis focuses only on efficiency scores calculated from the RPM, and these scores refer only to year 2005.

<Table 7> around here

The picture that emerges is that the presence of a medical school is associated with higher efficiency scores. Meanwhile, private ownership of the college is related to lower scores. The differences emerging from this analysis confirm the presence of heterogeneity in performance across subsectors of the HE market.
Johnes *et al*. (2005), when studying the cost structure and efficiency profile of the higher education sector in England, pointed out that efficiencies were highest in the top 5 universities and lowest in the colleges of higher education. Analogously, here we try to find out such a pattern in the HE market. We followed two strategies for this purpose: (1) a look at the efficiency scores for "high intensive research" universities – as defined by the Carnegie classification, and (2) an analysis of efficiency scores for

the best universities as classified by the US News and Report ranking (2009 edition). The last row of table 7 gives evidence that supports our hypothesis that research universities are relatively efficient. These universities are typically characterized not only by higher average efficiency, but also by a very small standard deviation – in other words, performances of this group of colleges are very similar across institutions. The distribution of efficiency scores, presented in figure 6, gives intuitive evidence of the structural difference that characterizes this group of universities with respect to the general distributions of scores (see figures 3-5) above.

<Figure 6> around here

Table 8 reports the efficiency scores for the first 20 universities listed in the US News and Report ranking (2009 edition). Here, the efficiency scores are those for all the three years considered (2003-05). Again, what emerges is a very high efficiency score, well above the average of the sector.

Several things could explain this finding. One possible interpretation is that efficiency, productivity and research and teaching excellence are highly correlated measures of performance.

<Table 8> around here

Ray returns to scale, derived using the RPM, are reported, for the average institution, in Table 9.

<Table 9> around here

The general picture that emerges is that there remain some scale economies to be exploited, albeit only at a modest level. There are, however, no unexploited economies of scope - especially at the higher levels of output. The interpretation of this story is quite straightforward and compelling: there are too many colleges overall (if the minimization of global costs is the objective), and, in particular, the provision of undergraduate education is not adequately concentrated. The other side of the coin is that there are too few colleges providing postgraduate education. Our findings on returns to scope serve to confirm this analysis – that increased specialization and concentration could lead to greater allocative efficiency across the higher education system.

6. Conclusion

The higher education system of the United States has been the subject of much research over the years, but the present paper represents the first attempt to apply recently developed econometric methods as a means of evaluating costs, returns to scale, returns to scope and efficiency in the context of a framework which allows for the heterogeneity of institutions. Our results confirm our view that American institutions of higher education are indeed heterogeneous. Those that have high profile as research institutions tend also to operate at high levels of efficiency. There is, however, a tail of less efficient institutions.

From the perspective of cost efficiency, there is evidence to suggest that there are too many higher education institutions in the US. To some extent, this may be due to geography – to cater for the needs of students who are unwilling or unable to travel far for tertiary education, there needs to be widespread geographical coverage. Nonetheless, global costs could be reduced by increased concentration of provision of undergraduate education. Meanwhile, scope economies in some institutions could be improved if postgraduate provision were extended to a greater number of institutions.

It is instructive to compare these results with those obtained for the higher education systems of some other countries. In Germany and Spain, studies have found evidence to suggest that product-specific economies of scale remain unexploited for all outputs of higher education institutions (Johnes and Schwarzenberger, 2007; Johnes and Salas-Velasco, 2007); in the case of Italy, meanwhile, Agasisti and Johnes (2008) have found the opposite. Perhaps the most intriguing comparison, though, is with England, where, in sharp contrast to the United States, Johnes and Johnes (2009) found increasing product-specific returns to scale for postgraduate education and decreasing returns to undergraduate provision. These findings suggest that there is much that can be learned from the comparison of systems across countries, and we suggest that further work on international data should be a priority for future research.

References

Agasisti, T. Dal Bianco, A., (2006) Data envelopment analysis to the Italian university system: theoretical issues and policy implications, International Journal of Business Performance Management, 8, 344-367.

Agasisti, T., Dal Bianco, A. (forthcoming), An analysis of the cost efficiency in Italian university system: non-parametric (data envelopment analysis) and parametric (stochastic frontier) approaches, International Journal of Management and Decision Making.

Agasisti, T., Johnes, G., (2008), Heterogeneity and the evaluation of efficiency: the case of Italian universities, Applied Economics.

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

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

Brinkman, P.T., Leslie, L.L., (1986), Economies of scale in higher education: sixty years of research, Review of Higher Education, 10, pp. 1–28.

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

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

Cohn, E., Cooper, S.T., (2004), Multi-product cost functions for universities: economies of scale and scope, in Johnes, G., Johnes J. (eds.), International Handbook on the Economics of Education, Edward Elgar Publishing Ld, Cheltenam, UK, pp. 579-612.

De Fraja, G., Valbonesi P., (2008), The design of university system, Centre for Economic Policy Research, Discussion Paper Series, No. 7038.

De Groot, H., McMahon, W.W., Volkwein J.F., (1991), The cost structure of American research universities, The Review of Economics and Statistics, 73(3), 424-431.

Del Rey, E., (2001), Teaching versus research: a model of state university competition, Journal of Urban Economics, 49, 356-373.

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

Getz, M., Siegfried, J.J., Zhang, H., (1991), Estimating economies of scale in higher education, Economics Letters, 37, pp. 203-208.

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

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

Harter, J.F.R., Wade, J.A., Watkins, T.G., (2005), An examination of costs at four-year public colleges and universities between 1989 and 1998, The Review of Higher Education, 28(3), 369-391.

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

Johnes, J., Taylor, J. (1990) Performance indicators in higher education, Buckingham: Open University Press.

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

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

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

Johnes, G. and Johnes, J. (1993) Measuring the research performance of UK economics departments, Oxford Economic Papers, 45, 332-347.

Johnes, G. and Johnes, J. (2009) Higher education institutions' costs and efficiency: taking the decomposition a further step, Economics of Education Review, 28, 107-113.

Johnes, G., Johnes, J., Thanassoulis, E., Lenton, P., Emrouznejad, A. (2005) An exploratory analysis of the cost structure of higher education in England, London: Department for Education and Skills Research Report 641.

Johnes, G. and Salas-Velasco M, (2007), The determinants of costs and efficiencies where producers are heterogeneous: the case of Spanish universities, Economics Bulletin, 4(15), 1-9.

Johnes, G., Schwarzenberger, A., (2007), Differences in cost structure and the evaluation of efficiency: the case of German universities, Lancaster University Management School.

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

Koshal, R.K., Koshal, M., (1999), Economies of scale and scope in higher education: a case of comprehensive universities, Education of Economics Review, 18, 269-277.

Laband, D.N., Lentz, B.F., (2003), New estimates of economies of scale and scope in higher education, Southern Economic Journal, 70(1), 172-183.

Laband, D.N., Lentz, B.F., (2004), Do costs differ between for-profit and not-for-profit producers of higher education?, Research in Higher Education, 45(4), 429-441.

Lenton, P., (2008), The cost structure of higher education in further education colleges in England, Economics of Education Review, 27, 471-482.

Orea, L., Kumbhakar, S.C., (2004), Efficiency measurement using a latent class stochastic frontier model, Empirical Economics, 29, 169-183.

Sav, T.G., (2004), Higher education costs and scale and scope economies, Applied Economics, 36, 607-614.

Stevens, P.A. (2005) The determinants of economic efficiency in English and Welsh universities, Education Economics, 13, 355-374.

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

Table 1. Descriptive statistics

	Mean	St. deviation	Min	Max
Costs	259.006	482.855	.462	4,200.00
Bachelor	1,162.20	1,449.35	.000	9,840.00
Postgraduate	571.809	851.284	.000	6,985.00
Research	36.318	98.766	.186E-03	995.00
Private	.446	.497	.000	1.000
Medical	.164	.369	.000	1.000

Notes: Costs and Research have been reported in millions $.
Number of observations: 2,862.

Table 2. Regression results

	Frontier	LC Model (group 1)	LC Model (group 2)	RPM Model
Constant	-53.550	-101.497	-6.829	-17.998
	(4.309)	(19.492)	(.750)	(1.242)
Bachelor	.070	.065	.067	.059[a]
	(.004)	(.014)	(.001)	(.001)
Postgraduate	-.004	-.029	.025	-.015[a]
	(.008)	(.027)	(.002)	(.002)
Research	2.367	3.293	2.005	5.097[a]
	(.094)	(.234)	(.033)	(.018)
Bachelor^2	-.010	-.031	-.021	-.074
	(.009)	(.024)	(.003)	(.003)
Postgraduate^2	.133	.159	.050	.146
	(.019)	(.077)	(.006)	(.007)
Research^2	-10.648	-19.046	-9.190	-6.703
	(1.017)	(3.079)	(1.693)	(.306)
Bachelor x	-.074	.020	-.024	.334
Postgraduate	(.025)	(.062)	(.010)	(.008)
Research x	3.376	2.731	2.040	3.201
Postgraduate	(.296)	(.738)	(.356)	(.076)
Bachelor x	-.579	-.617	-.179	-5.879
research	(.137)	(.304)	(.130)	(.035)
Log-likelihood	-18,216.49	-15,299.56	-15,299.56	-15,788.64

Notes: [a] mean of random parameters.

Table 3. AIC, MC and product-specific scale effects – traditional frontier model

AIC	bachelor	postgraduate	research
1	62,432	7,125	2.45
0.75	48,244	3,243	1.82
1.25	75,672	12,408	3.10
MC	**bachelor**	**postgraduate**	**research**
1	61,224	14,739	2.42
0.75	47,565	7,525	1.80
1.25	73,785	24,304	3.03
Scale	**bachelor**	**postgraduate**	**research**
1	1.020	0.483	1.016
0.75	1.014	0.431	1.012
1.25	1.026	0.511	1.020

Table 4. AIC, MC and product-specific scale effects – latent class model

GROUP 1			
AIC	**bachelor**	**postgraduate**	**research**
1	53,871	28,489	3.28
0.75	42,567	10,442	2.46
1.25	63,733	53,820	4.10
MC	**bachelor**	**postgraduate**	**research**
1	47,262	49,116	3.05
0.75	38,850	22,045	2.33
1.25	53,406	86,049	3.74
Scale	**bachelor**	**postgraduate**	**research**
1	1.140	0.580	1.076
0.75	1.096	0.474	1.056
1.25	1.193	0.625	1.097
GROUP 2			
AIC	**bachelor**	**postgraduate**	**research**
1	63,914	27,264	2.05
0.75	48,466	20,096	1.53
1.25	79,008	34,667	2.58
MC	**bachelor**	**postgraduate**	**research**
1	62,112	29,031	2.04
0.75	47,452	21,090	1.52
1.25	76,192	37,428	2.56
Scale	**bachelor**	**postgraduate**	**research**
1	1.029	0.939	1.005
0.75	1.021	0.953	1.004
1.25	1.037	0.926	1.006

Table 5. AIC, MC and product specific scale effects – Random Parameter model

AIC	**bachelor**	**postgraduate**	**research**
1	48,086	43,708	4.57
0.75	38,128	21,745	3.53
1.25	56,668	73,030	5.55
MC	**bachelor**	**postgraduate**	**research**
1	39,314	52,086	4.55
0.75	33,194	26,457	3.51
1.25	42,963	86,120	5.51
Scale	**bachelor**	**postgraduate**	**research**
1	1.223	0.839	1.005
0.75	1.149	0.822	1.004
1.25	1.319	0.848	1.007

Table 6. Mean levels of output, LC model

	bachelor	postgraduate	research ($million)
Group 1	2,151	1,295	122
Group 2	864	353	10.6

Table 7. Efficiency scores (random parameters model) – descriptive statistics

College Type	**Mean**	**St. Dev.**	**Min**	**Max**	**#**
Overall	0.5776	0.2556	0.0013	0.9988	898
Medic Private	0.7728	0.2505	0.0330	0.9872	60
Medic Public	0.8384	0.1839	0.0043	0.9988	93
Nomedic Public	0.5997	0.2151	0.0013	0.9891	425
Nomedic Private	0.4359	0.2304	0.0022	0.9796	320
Research universities	0.9132	0.0480	0.7213	0.9988	95

Notes. All the efficiency scores refer to year 2003. Research universities included in the last row are those classified as "very high research" universities by the Carnegie Classification.

Table 8. The efficiency of best colleges – as classified by the US News and Report 2009 Edition

College	**2003**	**2004**	**2005**	**Mean**
Harvard University	0.9933	0.9722	0.9024	0.9560
Princeton University	0.9180	0.9416	0.9340	0.9312
Yale University	0.9846	0.9652	0.9446	0.9648
MIT	0.9959	0.9774	0.9413	0.9715
Stanford University	0.9590	0.8901	0.9571	0.9354
California Institute of Technology	0.8571	0.8545	0.8642	0.8586
University of Pennsylvania	0.9979	0.9821	0.9512	0.9771
Columbia University	0.9961	0.9067	0.9109	0.9379
Duke University	0.9710	0.9726	0.9833	0.9756
University of Chicago	0.9225	0.8474	0.8747	0.8815
Dartmouth College	0.9414	0.9561	0.9224	0.9400
Northwestern University	0.9900	0.9584	0.9585	0.9689
Washington University St. Louis	0.9594	0.9600	0.9673	0.9622
Cornell University	0.9879	0.9870	0.9489	0.9746
John Hopkins University	0.9531	0.9970	0.9613	0.9705
Brown University	0.9340	0.9130	0.8417	0.8962
Rice University	0.8935	0.8742	0.8307	0.8661
Emory University	0.9103	0.9056	0.9514	0.9225
University of Notre Dame	0.9297	0.9187	0.9060	0.9181
Vanderbilt University	0.9883	0.9472	0.9220	0.9525
University of California Berkeley	0.9733	0.9811	0.9264	0.9603

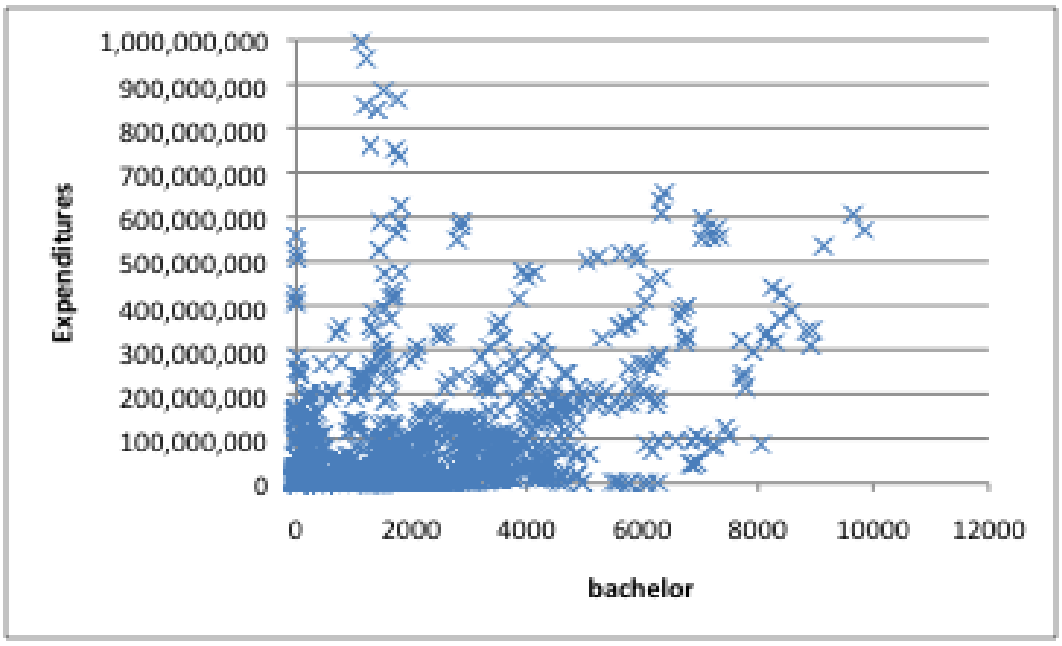
Notes. The efficiency scores have been derived using the random parameter modeling.

Table 9. Ray economies of scale and scope economies, random parameter stochastic frontier model

Ray economies of scale	
1	1.0716
0.75	1.0566
1.25	1.0851
Scope economies	
1	-0.3248
0.75	-0.0139
1.25	-0.5051

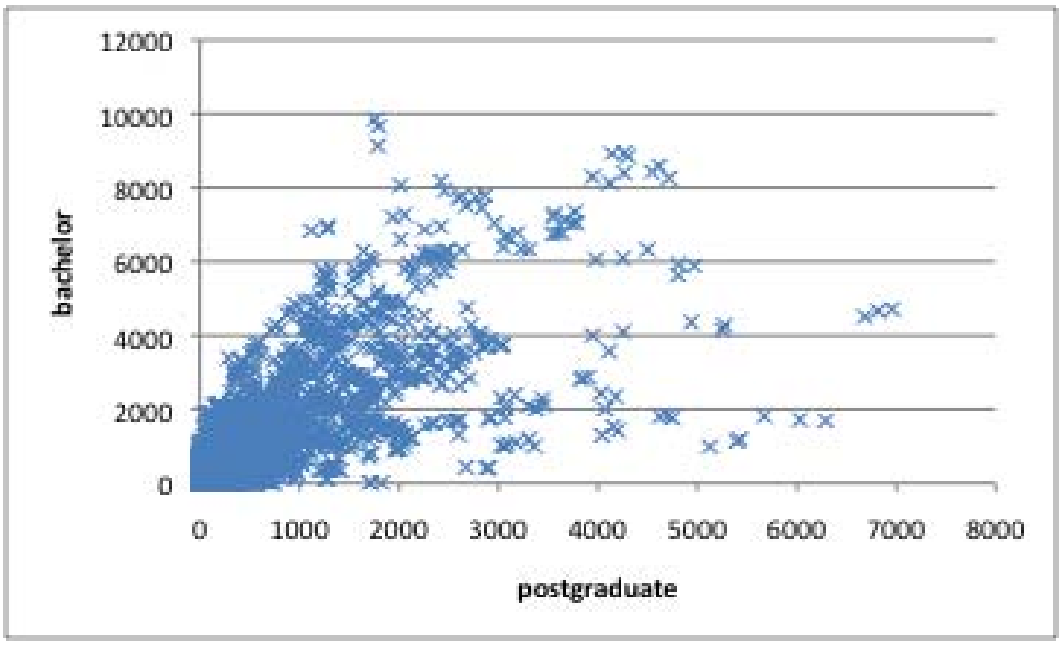
Notes. The mean values considered are the means of the output vector.

Figure 1. The output mix of US colleges' – bachelor vs research grants



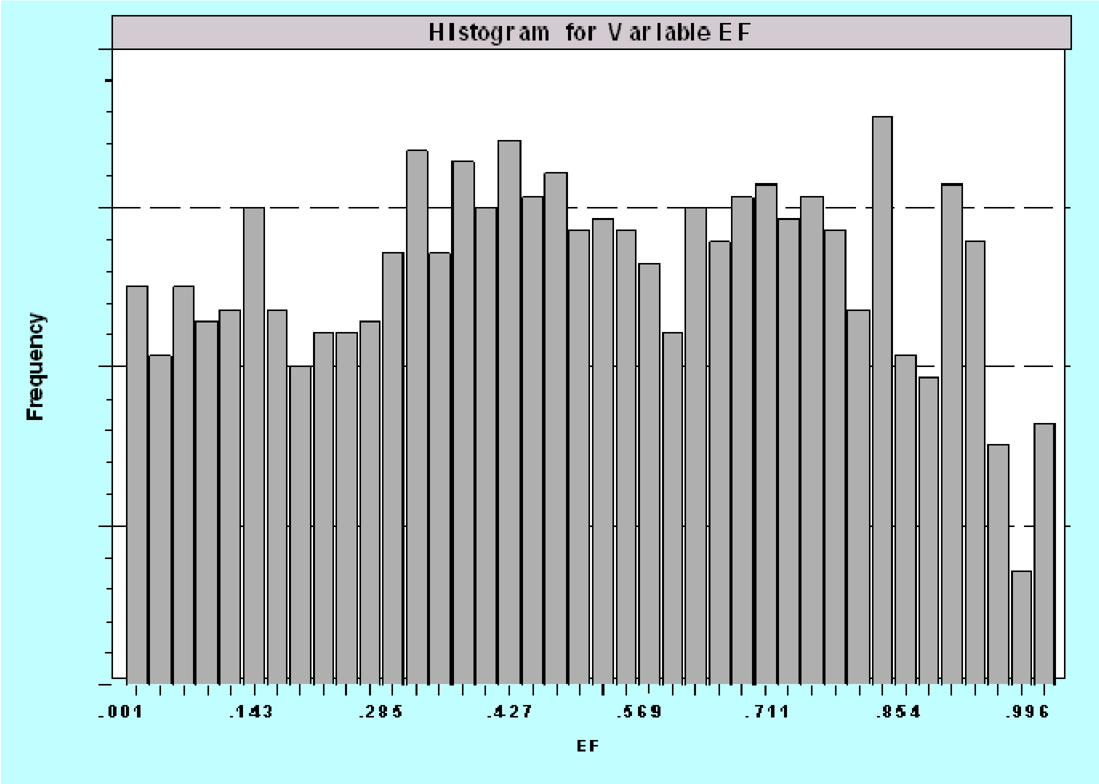
Notes. The observations refer to all the three years (each college is included three times in the figure).

Figure 2. The output mix of US colleges – postgraduate vs undergraduate education



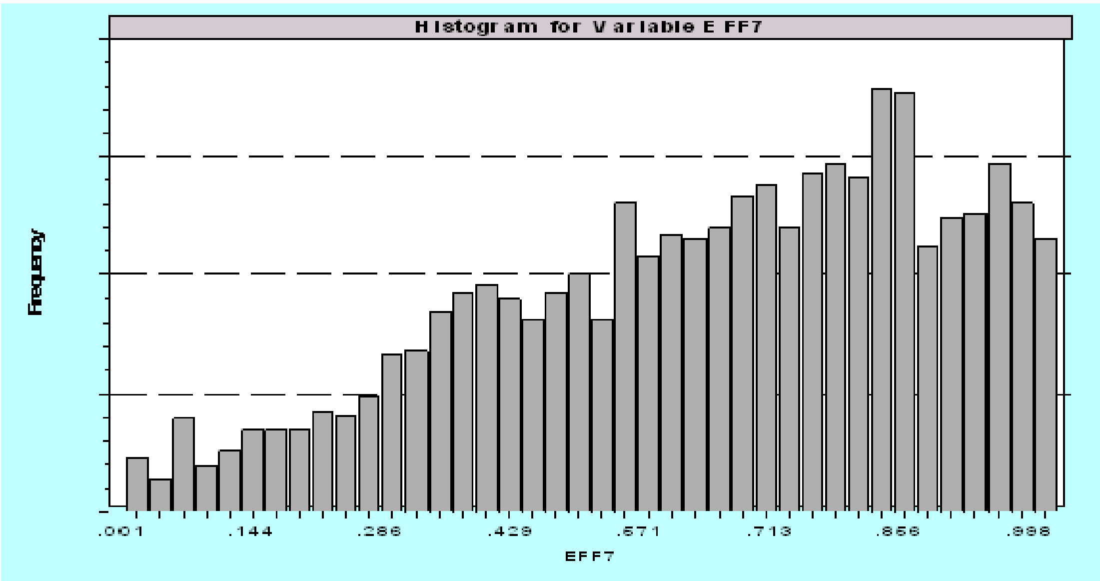
Notes. The observations refer to all the three years (each college is included three times in the figure).

Figure 3. Distribution of the efficiency scores – traditional frontier model



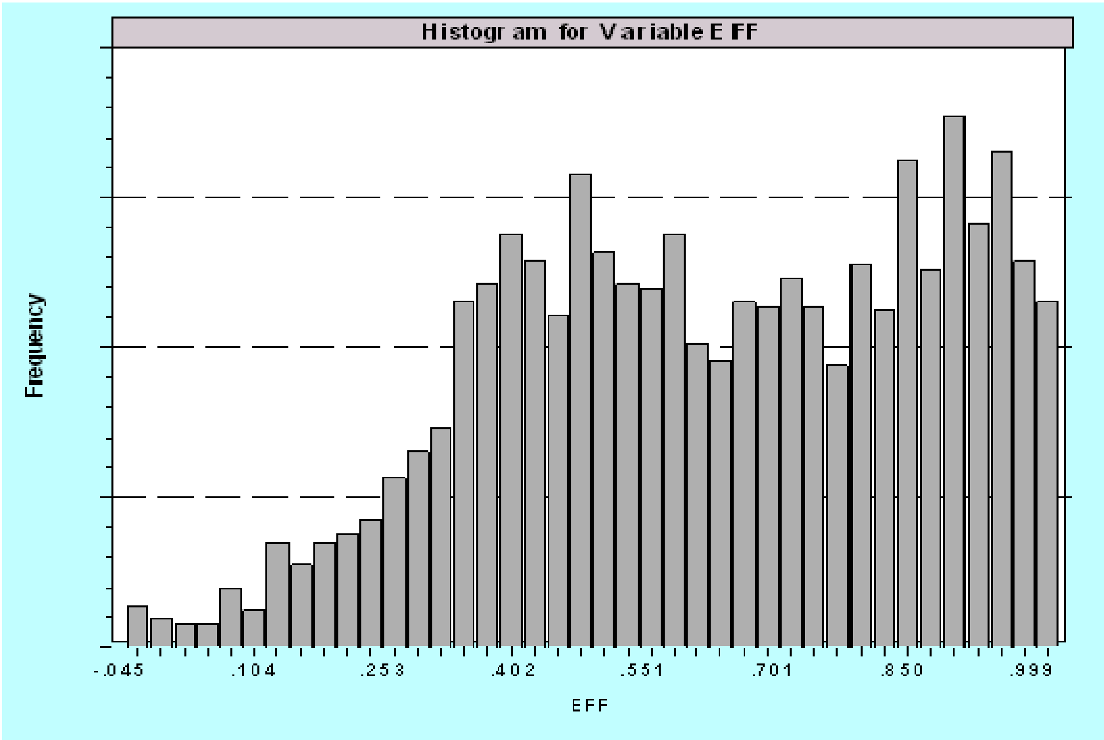
Notes. The efficiency scores refer to all the three years (each college is included three times in the figure).

Figure 4. Distribution of the efficiency scores –latent class frontier model



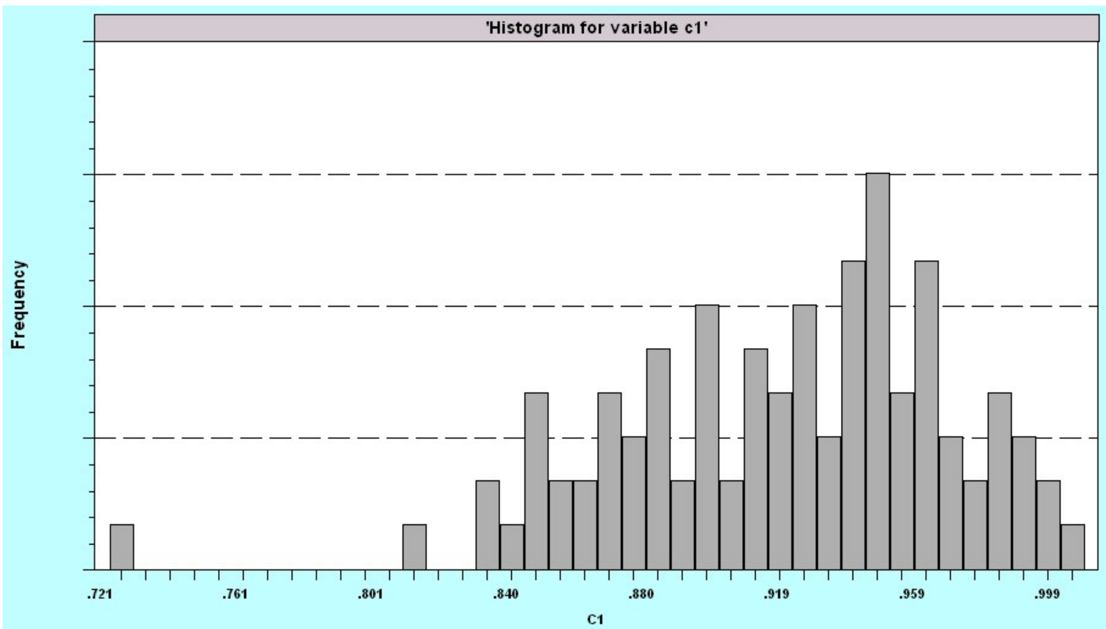
Notes. The efficiency scores refer to all the three years (each college is included three times in the figure).

Figure 5. Distribution of the efficiency scores – random parameter frontier model



Notes. The efficiency scores refer to all the three years (each college is included three times in the figure).

Figure 6. Distribution of efficiency scores for "research intensive" universities – random parameter frontier model



Notes. All the efficiency scores refer to year 2003. Research universities are classified as "very high research" universities by the Carnegie Classification.