Education Economics
Vol. 14, No. 3, 281–295, September 2006

Using Data Envelopment Analysis to Improve Estimates of Higher Education Institution's Per-student Education Costs[1]

CARLO SALERNO

Center for Higher Education Policy Studies, University of Twente, The Netherlands

ABSTRACT *This paper puts forth a data envelopment analysis (DEA) approach to estimating higher education institutions' per-student education costs (PSCs) in an effort to redress a number of methodological problems endemic to such estimations, particularly the allocation of shared expenditures between education and other institutional activities. An example is given using data for a sample of higher education institutions in The Netherlands and the results are compared with PSC estimates generated by a more traditional approach. Although several methodological concerns still persist, the use of DEA is argued to increase the likelihood of producing more realistic cost estimates for individual institutions.*

KEY WORDS: Higher education; resource use; costs; efficiency; data envelopment analysis

Introduction

As competition for scarce public funding intensifies, so too have tensions between institutions of higher education (IHEs) and the public they serve. Persistent annual increases in student tuition coupled with simultaneous efforts to pursue aggressive fundraising or secure increasingly greater levels of public financial support have both helped to mold the contemporary view that IHEs are, in general, inherently wasteful organizations. In response, numerous studies have been undertaken (e.g., de Groot *et al.*, 1991; J.M. Consulting, 2000; NCCHE, 1998; Toutkoushian, 1999) that seek a better understanding of what drives cost behavior in IHEs and, more specifically, to produce estimates of how much it actually costs IHEs to provide students with an education.

Unfortunately, developing institutional per-student cost (PSC) estimates that possess any degree of useful precision has historically faced several major obstacles. Non-uniform accounting procedures employed across institutions make it difficult to determine or even isolate relevant labor and non-labor costs (Winston, 2000) or to properly account for internal cost variations, such as those between

Correspondence Address: Carlo Salerno, Senior Research Associate, Center for Higher Education Policy Studies (CHEPS), University of Twente, Postbox 217, 7500 AE Enschede, The Netherlands. Email: c.s.salerno@utwente.nl

0964-5292 print/1469-5782 online/06/030281-15 © 2006 Taylor & Francis
DOI: 10.1080/09645290600777485

relatively expensive physical or life sciences programs and those in the social sciences or humanities. There are also difficult-to-measure cost differences between the *levels* of education provided, as the smaller course sections in graduate and upper-level undergraduate classes are clearly more expensive on a per-student basis than introductory courses taught in large lecture halls. Then there is the persistent omission of differences that are possibly explainable by productive and/or cost inefficiencies. Studies frequently estimate diverse ranges of education unit costs without recognizing that some of the variation is probably due to relatively inefficient production practices.

Yet, as Winston (2000) suggests, how to deal with joint or shared costs 'seems to be the most difficult problem facing the generation of meaningful estimates of the cost of undergraduate education … and it is the problem most in need of coordination of methodologies and assumptions among schools if their results are to be comparable' (p. 43). From faculty members' time to library and computing resources, practically all university resources are used in the production activities of several outputs, and determining where such shared expenditures should be allocated has direct consequences for both research on IHE costs and for policy-makers' efforts to design funding mechanisms that prudently allocate public funds. Estimates that place too much of an institution's shared expenditures into education can potentially lead to overly generous funding that may promote waste. On the other hand, underestimating institutional expenditures to education may lead to funding levels that are insufficient for supporting the provision of a quality education. While contemporary education cost estimates employ one of several different methods to allocate institutions' shared costs, as researchers generally agree, the limitations to current techniques suggest an alternative approach is warranted.

This paper puts forth such an alternative by way of data envelopment analysis (DEA). Briefly, I argue that by utilizing what has been called a 'shared resources DEA model' to first allocate shared expenditures between institutions' various activities, many problems associated with contemporary cost-estimation techniques can be simultaneously mitigated, thus increasing the likelihood of generating more realistic institutional PSC estimates. This is followed by an empirical example from another study completed by Jongbloed *et al.* (2003). Based on data collected for a subset of higher professional education institutions in The Netherlands, two different sets of PSC estimates are presented; one based on a more traditional approach and the other using DEA. These are then compared and contrasted to magnify the potential estimate variations arising from using the different estimation techniques.

Background

An institution's per-student education cost can be defined as the ratio of all expenditures attributable to education production relative to the total number of students educated. This gives a measure of the *average total cost*[2] to provide an education, and can be specified mathematically as:

$$PSC = \frac{\sum_{i=1}^{n} w_i E_i}{\sum_{j=1}^{m} v_j S_j} \tag{1}$$

where *PSC* is the per-student cost, E_i is the *i*th category of institution expenditures, w_i is the percentage of expenditure *i* for education services, S_j is the *jth* category of students, and v_j is the weight or importance attached to the *jth* student category.

At the outset, the likelihood of obtaining precise PSC estimates will be driven by the available data. Under ideal circumstances all data would be reported as direct expenditures; w reduces to a vector of 1 values, equation (1) reduces to summing all of the E_i terms and the joint cost problem disappears. In the 'worst' case, all expenditures data would be reported as shared, meaning that estimates will depend heavily on the value each w_i takes. Between these two extremes lies a continuum of scenarios. As the ratio of direct to total expenditures rises, so does the overall reliability of PSC estimates, because only a smaller percentage of the overall expenditures will need to be allocated between education and other activities.

In practice it is unlikely that any expenditure will be reported directly. When one considers the typical professor working in his or her office it becomes readily apparent how cost prohibitive it would be to determine, among other things: how much of his salary (based on time spent in the office on education-related activities) should be regarded as an education expenditure; how much electricity is used in the office when he is doing education related activities (e.g., the amount of electricity consumed by the computer and lights); and how much of the rental costs for the office space should be allocated to education costs (e.g., based on the time spent holding office hours and doing other education-related activities). In fact, joint expenditures prevail in nearly all major cost centers within an institution, from human resources to physical plant and capital to academic computing. Hence the crucial task is finding an appropriate method for determining what values w should take. In the following we look at three contemporary approaches to solving this problem.

The simplest method has been to ignore them altogether (i.e., let $w = 1$) and include all education and research expenditures in calculations of the cost of providing undergraduate education. Proponents usually justify such an approach by appealing to Nerlove's (1972) theory that education and research are complementary outputs and to a number of empirical studies that find evidence for scope economies (e.g., Cohn *et al.*, 1989; Dundar and Lewis, 1995). Faculty members teach the findings from their research in the classroom and, at the same time, both students and faculty members use the process of research to broaden their education. This 'full costing' methodology was applied as recently as 2002 in the National Association of College and University Business Officers' (NACUBO) three-year study to 'create a uniform methodology that any college or university in the nation could use to explain and present the costs of providing one year of undergraduate education' (NACUBO, 2002, p. 2). However, while full costing greatly simplifies the estimation procedure, it puts an upward bias on the results. While most would agree that graduate education certainly has synergies with academic research, faculty members' incremental research contributions to narrowly tailored disciplinary areas rarely filter down into introductory or even intermediate undergraduate course curricula that largely underpin instructional costs.[3]

The most widely used method for determining w has been to use the amount of time faculty members report spending on education-related activities (in workload studies) as a proxy for all other resource use (e.g., Enders and Teichler, 1997;

Goudriaan *et al.*, 1998; James 1978; Jongbloed and Vink, 1994). The rationale behind this approach is that higher education production is predominantly a labor-intensive activity. If faculty direct how other inputs are employed then it seems reasonable to conclude that surveys of how faculty members spend their time should also reflect the proportion of other inputs' use. This way of allocating shared resources is intuitively appealing yet several problems emerge. Data collection is a costly undertaking in terms of both money and time. As such, faculty workload studies tend to be conducted infrequently, which means that researchers must often base current estimates on past data. Although surveys generally show average faculty workloads to be relatively stable over time, even a measured shift of 5–10% away from or towards education activities would have a substantive impact on PSC estimates. Second, there is considerable evidence to suggest that faculty members may not accurately report *how* they spend their time (Jordan, 1994) as they are frequently known to put little effort into filling out time sheets and in some cases to simply fill in hours based on what they believe their superiors expect. Research by Teichler (1996) has shown that even when faculty members do take workload studies seriously, results are still likely to be inconsistent as individuals often have different definitions about what is regarded as a teaching or research activity. Third, and perhaps most important, it does not consider substitution effects. As faculty resources decline, one would intuitively expect that students would have to substitute (i.e., add more) other university inputs in their place.

The last approach to accounting for shared expenditures, which is in some respects similar to that already discussed, was first put forth by To (1987) and later given wider attention by Winston and Yen (1995). Here the objective is to base shared expenditure allocations on the proportion of an institution's clearly identifiable instructional expenditures relative to its total expenditures. For example, if such instructional expenditures constitute 70% of an institution's total expenditures then it would be assumed that education production commands 70% of all shared resources as well. This method has limited applicability as it implies that certain expenditures can be clearly regarded as 'education only', which depends largely on how data are reported in accounting statements or to national statistical agencies. It is also conditional on the extent to which the proportionality assumption is valid, and even Winston cautions that such an approach is of dubious value when making institutional comparisons. Finally, it does not account for the possibility that significant cost variations arise because of inefficient resource employment. An institution may very well allocate 70% of its shared resources to education, but if 20% of this allocation is due to waste then PSC estimates will not reflect the true PSC but an unadjusted measure of what is spent.

To summarize, there are at least three problems with contemporary approaches to apportioning shared expenditures that seriously affect contemporary estimates of institution's PSCs. The first is bias stemming from the assumption that all institutions utilize shared inputs proportionally to instructional expenditures or faculty workload. The second is that estimates are likely to be overstated either because no accounting is made for inefficient resource use or because research expenditures are fully accounted for in the cost of education. Third, where faculty workload surveys are concerned, not only is there potential bias from inaccurate reporting but, since data are so costly to collect, researchers generally must derive shares based on past, rather than current, results.

Using DEA to Allocate Joint Costs

DEA was first developed in the late 1970s as a non-parametric way to assess productive inefficiency, and since that time it has gained a remarkable degree of popularity. While several factors have driven its widespread adoption its most appealing feature is the ability to estimate efficiency in complex multi-input/multi-output firms where the underlying production process is not well understood (Cooper *et al.*, 2000). As will be seen, this property also lends itself nicely to objectively allocating joint costs.

One of the most common ways to define a firm's productivity is by the ratio of total output it produces to the amount of inputs it uses. A good higher education example might be the number of students (educated) per faculty member. However, in the case of firms such as IHEs that produce a number of different education outputs (n) using a bundle of different inputs (m), summing partial measures like that above to obtain a single aggregate measure of the institution's productivity is not sufficient. In order to develop an aggregate or total factor productivity (TFP) measure it is necessary to attach some relative importance or weight to each input and output:

$$TFP = \frac{\sum_{i=1}^{n} d_i y_i}{\sum_{j=1}^{m} w_j x_j} \tag{2}$$

where y_i is the ith output, d_i is the weight attached to the ith output, x_j is the jth input, and w_j is the weight attached to the jth input.

Equation (2) shows that productivity is nothing more than a rank-free indicator of the rate at which inputs are translated into outputs. To this point nothing has been said about the units in which inputs and/or outputs are expressed, but it is evident that if inputs are measured in terms of costs then the productivity measure is the inverse of cost per unit of output. Using the earlier higher education example, replacing 'faculty' with 'expenditures on faculty members' and putting inputs in the numerator gives the more familiar measure of 'faculty expenditures per student'. If instead of just faculty salaries one includes all institutional expenditures, then the inverse of equation (2) captures an IHE's per-student education cost.

Recall that the joint cost problem is determining what values w should take. This is where DEA's usefulness becomes evident. Through linear programming, DEA 'solves' for values of w that maximize each institution's ratio of weighted output to weighted input or to maximize, for each institution being studied (equation (2)). Rather than rely on expert opinion or educated guesses about what values different weights should take, DEA derives results by letting the data 'speak for itself'. A more formal specification of these ideas is as follows. Given a set of Q institutions, the optimization problem facing the kth one is:

$$\max \theta_k = \frac{\sum_{i=1}^{n} d_i y_i}{\sum_{j=1}^{m} w_j x_j} \tag{3}$$

subject to:

$$\theta_b = \frac{\sum_{i=1}^{n} d_i y_i}{\sum_{j=1}^{m} w_j x_j} \leq 1 \qquad (b \in 1, \ldots, k, \ldots, B) \tag{4}$$

$$w_j, d_i \geq \varepsilon \qquad \forall i \text{ and } j \tag{5}$$

Together, equations (3)–(5) represent the most basic of all DEA models.[4] The objective is to select, for institution k, the optimal values of w and d that maximize k's ratio of weighted output to weighted input subject to two constraints: (1) that the selected values of w and d cannot produce a larger value than 1 if applied to any other institution against which k is compared (equation (4)); and (2) that all w and d have a positive value (i.e., all inputs and outputs have at the least some marginal value to the institution). In short, DEA seeks to maximize each institution's productivity by identifying a set of weights that puts it in the best possible light relative to all other institutions being evaluated.[5]

This approach to determining weights has several appealing characteristics that lend themselves nicely to estimating institutional PSCs. First, equations (1) and (3) are identical in form. Second, the selection of weights is unique to individual institutions,[6] which is much more realistic than applying a set of uniform weights across the entire sample. Third, because it is based on observed data, which is generally available from year to year, it is not necessary to rely on dated faculty workload findings. Fourth, it is not based on what faculty members *report* doing but instead on what was actually observed, which helps to counter the bias problems endemic to survey research methodologies. Fifth, because DEA is really designed to estimate inefficiency it is also possible, to a certain extent, to distinguish between what it actually costs institutions per student and what they spend on a per-student basis.

No technique is without its share of limitations. One of the main criticisms leveled against DEA is that because it is a deterministic approach to assessing efficiency the findings can be highly sensitive to outlying observations in the sample and to omitting relevant inputs or outputs (Mettas *et al.*, 2001). As a result, misjudgment on behalf of the researcher as to what constitutes a relevant variable or simply the presence of random 'noise' in the data can heavily influence which weights are selected. Another major concern that is particularly relevant here is that DEA may be too flexible in the way it assigns weights. Because it attempts to paint each institution in the best possible light, the model may put too much weight on what are deemed 'unimportant' inputs or outputs and too little weight on critical ones. To deal with such problems, researchers have developed more sophisticated DEA models—for example, Dyson & Thanassoulis' (1988) assurance region or Charnes *et al.*'s (1989) cone ratio weight restriction models—that allow researchers to impose limitations on the extent to which weights for individual inputs and outputs can vary.

While several parallels have been drawn between equation (1) and equations (3)–(5), the optimal weights derived from the DEA model are *not* resource allocations but instead weights that indicate how important each input (output) is to deriving that institution's efficiency rating. The reason why equations (3)–(5) do not suit our need to allocate shared resources is that the basic DEA model implicitly assumes that all of a firm's outputs are generated by a single, joint production function; it tells nothing about how a firm *internally* allocates resources to the production of different outputs. This is problematic in the current context because there is no reason to assume *a priori* that the technologies behind education and research production are identical. Moreover, practically all available university inputs can be used to produce either education or research and it may be the case that one output is being produced more efficiently than the other.

Beasley (1990, 1995) was the first to bring attention to this problem and outlined a way to separately assess the teaching and research efficiency of UK universities' Chemistry and Physics Departments using only aggregate expenditure data. Further research by Mar-Molinero (1996) and Tsai and Mar-Molinero (2002) demonstrated that Beasley's solution could be achieved by use of a modified DEA model. This is the approach used to allocate joint costs in this paper. Building on equations (3)–(5), let higher education institutions' outputs instead be categorized into two broad activities: teaching (y^r) and non-teaching (y^R). Let teaching be measured by n different education outputs and non-teaching be measured by p different outputs. For simplicity, assume that the available input data (x) are aggregated in such a way that each of the m available inputs can be used in either activity but not simultaneously for both activities.[7] This last assumption captures the pragmatic way in which higher education institutions' expenditures data are normally collected and reported (Beasely, 1995, p. 445). Given a set of B institutions, the optimization problem facing the k th one is:

$$\max \phi_k = \lambda\theta_{kt} + (1-\lambda)\theta_{kr} \tag{6}$$

subject to:

$$\theta_{bt} = \frac{\sum_{i=1}^{n} d_i y_i^T}{\sum_{j=1}^{m} q_j w_j x_j} \le 1 \qquad (b \in 1,\ldots,k,\ldots,B) \tag{7}$$

$$\theta_{br} = \frac{\sum_{i=1}^{p} d_i y_i^R}{\sum_{j=1}^{m} (1-q_j) w_j x_j} \le 1 \qquad (b \in 1,\ldots,k,\ldots,B) \tag{8}$$

$$\varepsilon \le q_j \le 1-\varepsilon \tag{9}$$

$$w_i, d_j \ge \varepsilon \qquad \forall i \text{ and } j \tag{10}$$

Two major features separate this model from that represented by equations (3)–(5). One, it no longer estimates just the institution's overall efficiency (θ_k) but instead a weighted average of the component efficiencies. The weights in the objective function (λ) reflect the relative value the institution places on each activity. A university may be an efficient producer of research but, if it is a teaching-oriented institution, then one should not place so much stock in the fact that it is efficient at a relatively unimportant task. The second is the introduction of the q weights, which identify the percentage of each shared resource (x) allocated to teaching and whose value is what we are actually seeking. Equations (7) and (8) are similar to equation (4) except they break equation (4) out into separate teaching and research components. They are still linked by q, which builds in the idea that expenditures allocated to one activity cannot be simultaneously allocated to the other activity.

One of the major benefits to this approach, as Beasley notes, 'is that it does not require an *a priori* split of expenditure into teaching/research. Instead, such a split is automatically decided …' (1995, p. 446).[8] In effect, DEA determines the optimal resource allocation split by appealing to the basic economic condition that firms attempt to efficiently allocate resources so that the marginal productivity of a dollar spent toward producing one good is equivalent across all other goods it produces. In other words, the same way that the basic DEA model solves for the weights that maximizes the ratio of weighted output to weighted input, and hence overall efficiency, here the model seeks the optimal weights and resource allocation combination of the shared inputs that jointly maximizes the ratio of outputs to inputs used for teaching and separately for research. In the shared resources variant, q can simply be regarded as one more set of weights to be estimated. Like the objectivity associated with assigning weights in the basic DEA model, a similar line of reasoning prevails for the distribution of shared resources. When determining component efficiencies, let us again give institutions the 'benefit of the doubt' and select the resource allocation that paints each institution's activities in the best light relative to other institutions. Here the 'best possible light' means maximizing the efficiency of each production process separately, albeit simultaneously.

Earlier it was stated that DEA makes it possible to redress other problems that arise when trying to allocate resources. Not only do those aspects also apply to the shared resources model, but there is an additional benefit. Rather than assuming, as the faculty workload or Winston approach do, that all inputs are internally allocated proportionally, each shared resource produces a unique *internal* allocation (Mar-Molinero, 1996). For example, if the data were sufficiently disaggregated so that one could identify four shared expenditure categories such as academic labor, non-academic labor, computing and other capital, it is possible to find out whether an institution allocates a less than proportional amount of non-academic labor to education and a more than proportional amount to, say, computing resources.

Importantly, the shared resources model also imposes an additional limitation. As both Beasley and Mar-Molinero found, under certain circumstances the shared resources model has a tendency to 'degenerate' and try to allocate all resources to one activity. This can be rectified by imposing an additional constraint, such as equation (9), which does not permit any resource to be allocated in its entirety to one activity. Yet the model can still generate unrealistic resource allocations. The only solution is to impose bounds on the extent to which resources can be allocated though this sacrifices some of the model's objectivity by incorporating researcher judgment. Overly generous bounds, those allowing resource allocation combinations not in the technology set, may produce unreasonably low expenditure estimates, while overly strict bounds may distort results because the optimal solution is no longer in the feasible set.

While the shared resources DEA model has to date only been applied to develop estimates of cost efficiency, its applicability to estimating institutional PSCs is readily apparent. A natural 'by-product' of computing efficiency scores in this way is that it determines the allocation of shared resources for education. To construct efficiency scores for separate processes within an institution, the model first allocates shared costs to the production of several outputs under the objective criterion of efficiency maximization. Provided that education-only outputs can be distinguished from an institution's non-education outputs, the model will allocate shared costs to each output. By focusing only on the proportion of shared costs allocated to education outputs, it is then a straightforward exercise to compute individual institutions' PSC estimates.

An Application

We demonstrate the technique by reporting results from a recently completed cost study. There, PSC estimates were developed for professional higher education institutions, *Hogerberoepsonderwijs instellingen* (HBOs), in The Netherlands. We do not expound on the methodology in this paper but instead refer the reader to Jongbloed *et al*. (2003). Briefly, the sample included 36 HBOs that were subdivided for analysis into five categories[9] reflecting the different technologies employed across different types of education. Separate DEA analyses were carried out on each category so that institutions were only compared with like institutions. Inputs included two shared expenditure categories; personnel and non-personnel. The objective was to identify each inputs use toward two activities: education, as measured by FTE enrollments; and research, measured by the amount of contract-based funding[10] each institution received. The range of feasible allocations to education for both inputs was restricted to no less than 80% and no more than 99% based on our own experience with analyzing such institutions in prior studies and through discussions with the Dutch Ministry of Education. The percentage of total institutional revenues from contract-based resources was used to proxy the priority each institution gave to research, $(1 - \lambda)$ in equation (16). All data were for the 2000 academic year and taken from the publication *Hogescholen Management Informatie*, which is produced by the Association of Universities of Professional – Education (*Hogerberoepsonderwijs Raad* or HBO Council).

For comparative purposes a set of 'traditional' PSC estimates similar in spirit to the To and Winston approaches identified earlier were computed based on the methodology used by the HBO Council. Each institution's education expenditures were calculated by identifying total expenditures and then subtracting the amount of contract-based funding received. PSC estimates were then obtained by dividing this expenditure measure by the same full-time equivalent enrollment data used in the DEA analyses. The findings for the different institutional categories are presented in Table 1 and 2: Table 1 presents the result for the three groups of multidisciplinary institutions and Table 2 presents single-cluster institutions. The final column in both tables is the difference between the traditional and DEA-based estimates; a positive (negative) value indicates that the DEA estimates were lower (higher) than the traditional ones.

It is evident from Table 1 that the DEA-based PSC estimates are both higher and lower than the traditional estimates, which one would expect if institutions do in fact pursue unique resource allocations. Note also that in the five-cluster

Table 1. PSC estimates using traditional and DEA-based approaches: multi-cluster institutions (€000)

Institution	Traditional approach	DEA	Difference
Five-cluster			
5HS1	5.868	5.666	0.201
5HS2	5.153	4.833	0.320
5HS3	4.949	4.688	0.261
5HS4	5.032	4.754	0.278
5HS5	5.701	5.394	0.308
5HS6	5.367	4.846	0.520
5HS7	8.103	5.362	2.742
Average	*5.757*	*5.178*	*0.661*
Four-cluster			
4HS1	4.610	4.600	0.010
4HS2	4.887	4.636	0.252
4HS3	4.943	4.488	0.455
4HS4	4.963	4.467	0.496
4HS5	5.065	4.558	0.507
4HS6	5.224	4.976	0.249
4HS7	4.687	4.390	0.298
4HS8	4.701	4.560	0.141
4HS9	5.009	5.232	–0.222
4HS10	5.393	5.054	0.340
Average	*4.952*	*4.704*	*0.252*
Three-cluster			
3HS1	4.630	4.698	–0.068
3HS2	4.389	4.338	0.051
3HS3	6.348	6.652	–0.304
3HS4	5.790	5.983	–0.193
3HS5	4.456	4.528	–0.072
3HS6	4.237	4.148	0.090
Average	*4.887*	*4.975*	*–0.083*

and four-cluster groupings, DEA estimates are consistently lower than those from the traditional approach. This suggests that institutions in these two groups allocate more resources to contract activities than they receive through contract-based income, which is the type of evidence that is consistent with other research suggesting that institutions cross-subsidize research with education resources (e.g., James, 1990). In contrast, the results in the three-cluster group show just the opposite (i.e., that contract activities seemingly subsidize education); however, the magnitude of the differences between the two estimation techniques is admittedly small and the non-parametric approach behind DEA limits the ability to draw statistical inferences about the differences.

In institutions where only a single cluster of programs is offered (Table 2), again a similar pattern of DEA both overstating and understating traditional PSC estimates occurs. Estimates for teacher-training institutions are fairly similar between the two approaches, although DEA values are, for the most part, lower. A particularly interesting finding is that of the technical institutions, whose largely negative difference values point to contract work again partially subsidizing

Table 2. PSC estimates using traditional and DEA-based approaches: single-cluster institutions (€000)

Institution	Traditional approach	DEA	Difference
Technical			
T1	5.960	6.338	–0.377
T2	5.680	5.623	0.057
T3	5.831	5.198	0.633
T4	8.448	8.526	–0.078
T5	3.983	5.956	–1.973
T6	6.435	6.580	–0.145
Average	*6.336*	*6.656*	*–0.314*
Teacher-training			
ED1	4.558	4.376	0.182
ED2	5.768	5.575	0.193
ED3	5.464	5.476	–0.012
ED4	4.664	4.841	–0.177
ED5	5.348	5.175	0.173
ED6	5.288	5.235	0.053
ED7	4.669	4.596	0.073
Average	*5.112*	*5.050*	*0.069*

education. It should be noted, however, that several of the institutions in this group have strong agricultural slants. Not only are they likely to procure disproportionate contract research support, but also their more hands-on curricula would suggest cross-subsidies of this type would make sense.

Unfortunately there is no way of statistically ascertaining whether the DEA approach produces more reliable estimates, and several limiting factors are in evidence. For one, although the sample sizes are acceptable by DEA standards,[11] they are nonetheless still very small, which reduces the discriminatory power of the overall analysis. This is apparent in the relatively high number of institutions that were found to be relatively efficient in each institutional group (two on average). Second, because the DEA model attempts to equate marginal productivities of shared resources across different activities, the small number of hyperplanes formed by having so few relatively efficient institutions severely restricts the model's ability to identify unique resource allocations. This is evident in the findings reported in the original study showing a high number of cases where q values took the specified limits. If the imposed lower boundary was too low then some institutional PSC estimates may be lower than what was feasible. In the earlier study's analysis, this concern prompted us, in part, to place greater stock in the average cluster estimates.

In spite of the potential methodological concerns, the justification for DEA is intuitively appealing and logically sound. Not only does it impose fewer assumptions about an *individual* institutions' allocation behavior but it also links their behavior to that of similar institutions. This allows for evaluating each institution's behavior in a competitive framework rather than treating an individual institution's PSC estimates as being made independently of other institution's actions.

Overall, the differences between the traditional and DEA-based per-student expenditure estimates are, on average, small, which may raise concerns that policy-makers would find little benefit in using the more elaborate technique. Nevertheless, the fact that the DEA results are both lower and higher than the traditional estimates[12] depending on which institution is evaluated is important, particularly where equity and efficiency policies are concerned. Given that many governments utilize average PSC estimates in their funding formulas for annually appropriating funding for IHEs, the findings here suggest that formulas relying on traditional PSC estimates are inefficient resource allocation mechanisms; they probably provide some institutions with more resources than are necessary and, more importantly, may jeopardize the quality of other institutions' educational offerings through under-funding. A nice illustration of the former can be seen in our study's results for institution 5HS7 (a potential over-funding of €2742 per student) and for the latter in institution T5 (a potential under-funding of €1973 per student).[13]

Any concerns must also be tempered by the implications of magnifying per-student differences to an entire institution's, or even sector's, enrollment. At institution 4HS6, for example, whose enrollment is fairly large, the aggregate DEA-based estimate translates into approximately €0.5 million less in public funding *per year* when compared with the traditional estimate. Or, when the difference between the two estimation techniques is applied to the 104 000 students that were enrolled in just the seven five-cluster institutions, a funding formula based on the DEA estimates would annually allocate nearly €68.8 million *less* to these institutions. If achieving efficiency and equity are public priorities, and clearly they are, then any concerns about the small differences with student-level estimates dissipate when the economic reality from aggregating those findings at the system level is presented.

The final point addressed here is the difference between estimating PSCs and per-student expenditures. For example, in the five-cluster institutions both 5HS3 and 5HS4 were identified as relatively cost efficient in the provision of education, yet the estimate at 5HS3 was roughly €60 lower per student. To understand why involves going back to the optimal weights generated by the DEA model, where it can be seen that 5HS3's lower costs stem from using fewer personnel expenditures than 5HS4. An even more striking difference is evident in the four-cluster grouping where 4HS1 and 4HS8 were both found to be relatively efficient at education. The per-student expenditures at 4HS8 are approximately €100 lower than 4HS1's, and the reason for this again has to do with their minimization of personnel expenditures. The DEA weights indicate that 4HS8's relatively efficient rating is based largely on its minimizing personnel costs relative to other institutions while 4HS1's is the result of minimizing non-personnel expenditures. These findings are both interesting and highlight an important yet frequently overlooked aspect of efficient production; namely, that an institution producing above minimum cost does not necessarily imply technical inefficiency. As this application shows, how institutions deploy resources is just as important to explaining why different institutions generate different PSCs and, in some cases, it may unfair to infer that an institution is necessarily wasting resources simply because it has higher costs.

Conclusion

This paper put forth an alternative approach to allocating joint costs based on data envelopment analysis, in an effort to develop more realistic measures of

IME's PSCs. It was shown how a range of other factors including institutional priorities for education, notions of efficiency, and unique internal allocations can each be accounted for, and contributed to, the production of unique institutional PSC estimates. This technique was demonstrated using data for a sample of higher education institutions in The Netherlands and the findings compared with PSC estimates generated by a more traditional approach. While several methodological issues still persist we believe our findings are a marked improvement over contemporary estimation techniques and can provide both policy-makers and researchers with more accurate estimates on which intelligent policies or incentive systems can be based.

Acknowledgements

The author would particularly like to thank Ben Jongbloed and the two anonymous referees for their very useful comments. As usual, all mistakes and omissions are purely the author's own. This research was supported by the Netherlands Ministry of Education, Culture and Science (*Nederlands Ministerie van Onderwijs, Cultuur en Wetenschappen*).

Notes

1. An earlier draft of this paper was presented in September 2003 at the 16th Annual Consortium of Higher Education Researchers Conference in Porto, Portugal.
2. The most common alternative is to estimate the marginal or incremental cost incurred from enrolling an additional student or producing an additional unit of research. Although such measures are particularly useful for exploring the extent to which IHEs can realize economies of scale and scope, we restrict ourselves here to the body of research focusing on average total cost estimation.
3. Institutions' undergraduate populations are, in practically all cases, much larger than their graduate student populations. At the same time, it may be the case that laboratory equipment and other capital purchased initially for research activities through project funding may well be used in undergraduate laboratories at a later date.
4. Note also that the model here assumes that the production frontier can be characterized by constant returns to scale. This implies that if some input/output combination is feasible then, for any positive scalar, the larger input/output combination is also feasible.
5. Efficiency maximization does not imply that best-performing institutions are fully efficient, hence it is not possible to state that they are operating at minimum cost. Moreover, only a select number of decision-making units (DMUs) in a DEA analysis will ever be found to be relatively efficient. What the earlier statement means is that since inefficiency can already be seen as a penalty for an institution in a given analysis, let us attempt to minimize that penalty as much as possible.
6. Even where two institutions get projected back to the same facet of the efficient frontier, the weights will still differ unless the two institutions use exactly the same input proportions (i.e., one institution is a scaled down or up version of the other).
7. The more general formulation of the shared resources model takes into account that some inputs can only be used for certain activities (see Mar-Molinero, 1996). While this may be a more realistic reflection of the input/output relationships in higher education institutions, the focus here is on allocating 'reported' expenditures data, which implies that the simplifying assumption can be safely imposed.
8. The mathematical explanation for how the resource allocation decision is made is beyond the scope of this paper and the reader is referred to Mar-Molinero (1996) for a detailed explanation. Briefly, however, this is done to ensure that q takes the same value when computed by either the primal or dual linear program.
9. These groupings were produced by first sorting on the basis of whether each institution's education offerings could be grouped into a single cluster of like programs or whether it had a broad array of multi-cluster offerings. This gave rise to the following categories of institutions: (1) laboratory-based (e.g., technology); (2) teacher training and social science institutes; (3)

multi-disciplinary; (4) multi-disciplinary but without medical related programs; and (5) multi-disciplinary without medical and performance arts programs.
10. Contract-based funding, or third stream funding as it is commonly referred to, includes revenues for research activities but also for contracted education services as well.
11. Although there is no fixed decision rule for the minimum number of decision-making units in an analysis, a good rule of thumb is that the number should equal or exceed three times the product of the inputs and outputs specified (Cooper *et al.*, 2000). As was explained to me by Mar-Molinero, one particularly appealing characteristic of the shared resources model is that the minimum number of DMUs is less than that required by the more general DEA models. This can be seen in equations (6) and (7). Although two inputs and two outputs are employed, the shared resources model simultaneously estimates two models (one for each output) not one. This means that for each case here there is one output and two inputs. Using the rule of thumb above, the minimum number of institutions is $3 \times (2 \times 1) = 6$.
12. As can be seen in the tables, two-thirds of the DEA estimates were lower than traditional ones and one-third were higher.
13. These two results illustrate the point but are considerably extreme when compared with all of the other estimates in our study. This may suggest that there are unobserved problems with the underlying data for these two particular institutions.

References

Beasley, J. E. (1995) Determining teaching and research efficiencies, *Journal of the Operational Research Society*, 46(4), pp. 441–452.

Beasley, J. E. (1990) Comparing university departments, *Omega*, 18, pp. 171–183.

Charnes, A. *et al.* (1989) Cone ratio data envelopment analysis and multi-objective programming, *International Journal of Systems Science*, 20, pp. 1099–1118.

Cohn, E. *et al.* (1989) Institutions of higher education as multi-product firms: economies of scale and scope, *Review of Economics and Statistics*, 71, pp. 284–290.

Cooper, W. W. *et al.* (2000) *Data Envelopment Analysis: A Comprehensive Text with Models, Applications, References and DEA-Solver Software* (Dordrecht: Kluwer).

Dundar, H. and Lewis, D. R. (1995) Departmental productivity in American universities: economies of scale and scope, *Economics of Education Review*, 14(2), pp. 119–144.

Dyson, R. G. and Thanassoulis, E. (1988) Reducing weight flexibility in data envelopment analysis, *Journal of the Operational Research Society*, 39, pp. 563–576.

Enders, J. and Teichler, U. (1997) A victim of their own success? Employment and working conditions of academic staff in comparative perspective, *Higher Education*, 34(3), pp. 347–372.

Goudriaan, R. *et al.* (1998) *Kostendeterminanten en Doelmatigheid van het Nederlandse Hoger Onderwijs:* Beleidsgerichte Studies Hoger Onderwijs en Wetenschappelijk Onderzoek 57, Ministerie van Onderwijs, Cultuur en Wetenschappen (Den Haag: SDU). [*Cost Determinants and Effectiveness in Dutch Higher Education*]

de Groot, H. *et al.* (1991) The cost structure of American research universities, *Review of Economics and Statistics*, 73(3), pp. 424–431.

James, E. (1990) Decision processes and priorities in higher education, in: S. A. Hoenack and E. I. Collins (Eds) *The Economics of American Universities*, pp. 77–106 (Buffalo, NY: State University of New York Press).

James, E. (1978) Product mix and cost disaggregation: a reinterpretation of the economics of higher education, *Journal of Human Resources*, 13(2), pp. 157–186.

J. M. Consulting (2000) Efficient cost of teaching: report for SHEFC Prices Working Group. Available at: http://www.shefc.ac.uk/library/shefc/consultation/2000 (accessed 15 September, 2004).

Jongbloed, B. W. A. *et al.* (2003) *Kosten per student: Methodologie, schattingen en een internationale vergelijking.* Beleidsgerichte Studies Hoger Onderwijs & Wetenschappelijk Onderzoek, 100 (Den Haag: Ministerie van Onderwijs, Cultuur en Wetenschapelijk). [*Per Student Costs: Methodology, Estimates and an International Comparison*]

Jongbloed, B. W. A. and Vink, M. (1994) Assessing efficiency in British, Dutch and German universities, in: L. Goedegebuure and F. A. van Vucht (Eds) *Comparative Policy Studies in Higher Education* (Utrecht: Lemma).

Jordan, S. M. (1994) What we have learned about faculty workload: the best evidence, in: J. F. Wergin (Ed.) *Analyzing Faculty Workload, New Directions for Institutional Research*, 83(Fall), pp. 15–24.

Mar Molinero, C. (1996) On the joint determination of efficiencies in a data envelopment analysis context, *Journal of the Operational Research Society*, 47(10), pp. 1273–1279.

Mettas, R. D. *et al.* (2001) An investigation of the sensitivity of DEA to data errors, *Computer and Industrial Engineering*, 41, pp. 163–171.

NACUBO (2002) Explaining college costs: NACUBO's methodology for identifying the costs of delivering an undergraduate education. Available at: www.nacubo.org

NCCHE (1998) *Straight Talk About College Costs and Prices* (Phoenix, AZ: Oryx Press).

Nerlove, M. (1972) On tuition and the costs of higher education: prolegomena to a conceptual framework, *Journal of Political Economy*, 80(2), pp. S178–S218.

Teichler, U. (1996) The conditions of the academic profession: an international, comparative analysis of the academic profession in Western Europe, Japan and the USA, in: P. A. M. Maassen and F. A. van Vught (Eds) *Inside Academia: New Challenges for the Academic Profession* (Enschiede: Center for Higher Education Policy).

To, D.-L. (1987) *Estimating the Cost of a Bachelor's Degree: An Institutional Cost Analysis* (Washington, DC: Office of Educational Research and Improvement, US Department of Education).

Toutkoushian, R. K. (1999) The value of cost functions for policymaking and institutional research, *Research in Higher Education*, 40(1), pp. 1–15.

Tsai, P. F. and Mar Molinero, C. (2002) A variable returns to scale data envelopment analysis model for the joint determination of efficiencies with an example of the UK health service, *European Journal of Operational Research*, 141, pp. 21–38.

Winston, G. C. (2000) A guide to measuring college costs, in: M. F. Middaugh (Ed.) *Analyzing Costs in Higher Education: What Institutional Researchers Need to Know, New Directions for Institutional Research*, 106(Summer), pp. 31–46.

Winston, G. C. and Yen, I. C. (1995) Cost, price, subsidies, and aid in U.S. higher education. Williams Project on the Economics of Higher Education, Discussion paper 32. Available at: www.williams.edu/wpehe