Decision Sciences
Volume 31 Number 4
Fall 2000
Printed in the U.S.A.

Cost Efficiency Benchmarking for Operational Units with Multiple Cost Drivers

Marvin D. Troutt
Department of Management & Information Systems, Kent State University, Kent, OH 44242-0001, email: mtroutt@bsa3.kent.edu

Donald W. Gribbin
School of Accountancy, Southern Illinois University at Carbondale, Carbondale, IL 62901-4631, email: dgribbin@cba.siu.edu

Murali Shanker
Department of Management & Information Systems, Kent State University, Kent, OH 44242-0001, email: mshanker@kent.edu

Aimao Zhang
Management Department, Southern Illinois University at Carbondale, Carbondale, IL 62901-4627, email: aimao@siu.edu

ABSTRACT

We consider the activity-based costing situation, in which for each of several comparable operational units, multiple cost drivers generate a single cost pool. Our study focuses on published data from a set of property tax collection offices, called rates departments, for the London metropolitan area. We define what may be called benchmark or most efficient costs per unit of driver. A principle of maximum performance efficiency is proposed, and an approach to estimating the benchmark unit costs is derived from this principle. A validation approach for this estimation method is developed in terms of what we call normal-like-or-better performance effectiveness. Application to longitudinal data on a single unit is briefly discussed. We also consider some implications for the more routine case when costs are disaggregated to subpools associated with individual cost drivers.

Subject Areas: Activity-based Costing, Data Envelopment Analysis, Mathematical Programming, Performance Evaluation, and Statistics.

INTRODUCTION

This paper considers an area of activity-based costing (ABC) that has received little attention in the literature. In what may be called routine ABC, the general procedure for obtaining unit cost rates is as follows (Horngren, Foster, & Datar, 2000).

Costs for an activity are accumulated in cost pools. A variable, called a cost driver, is identified that measures the amount or extent of the activity performed, and that varies proportionally with the cost pool level. Then, the cost per unit of driver is found by dividing the cost pool by the associated driver level. The respective cost drivers generally are used as the basis for allocation of costs to specific cost objects (e.g., products or services). This procedure requires that costs can be disaggregated in such a way so as to be associated to a single cost driver. In practice, it is quite often possible to carry out ABC in this way by sufficient information refinement. In some cases, however, more than one cost driver may drive a cost pool simultaneously. The data set considered in this paper, attributed to Dyson and Thanassoulis (DT) (1988), and Thanassoulis, Dyson, and Foster (TDF) (1987), provides an example and is discussed below. Also, in the comprehensive study of the airline industry by Banker and Johnston (1993), two or more cost drivers were found to drive the major cost pools with high statistical significance.

As shown in this paper, the multiple cost driver case can be incorporated into a cost efficiency model useful for benchmarking comparable operational units in a firm, industry, or other comparable grouping. In this paper we propose a model-based method for benchmarking comparable operational units with multiple cost drivers. The proposed method yields relative cost efficiencies of the units and also provides estimates of what we call *benchmark costs*.

To clarify these concepts, suppose that x is a cost pool associated with the two simultaneous cost driver levels, y_1 and y_2. We wish to determine what unit cost rates or costs per unit (we call these simply costs when the meaning is clear) should be associated with these two drivers. Let a_1 and a_2 be the cost rates per unit of y_1 and y_2, respectively. Then, under the assumptions of linearity (additivity and proportionality) and constant returns to scale of the total cost function, these costs must satisfy the total cost function equation, $x = a_1 y_1 + a_2 y_2$. Of course, there is no unique solution in this one observation case. However, if several observations of x_j, y_{1j}, and y_{2j} are available, then modeling possibilities exist for estimation of the costs. In particular, regression through the origin with a model of the form:

$$x_j = a_1 y_{1j} + a_2 y_{2j} + \varepsilon_j \tag{1}$$

may be used to estimate what Dyson and Thanassoulis (1988) called “average costs.” This type of regression model was also the basis of the approach used in Banker and Johnston (1993). Thus, if data for several comparable units, or several observations of the same unit over time are available, then regression through the origin may be used to estimate costs (average unit cost rates) in the multiple simultaneous cost driver case.

However, there is a difficulty in the regression approach when the goal is to compare the units for what we call *cost efficiency.* That is, for benchmarking the efficient cost performance of the units it is desired to estimate the cost rates of the most efficient unit(s). The specific difficulty is that for one or more of the units, the x_j value may be larger than necessary for the associated y_{1j} and y_{2j} due to cost inefficiency. Namely, let $a_1°$ and $a_2°$ be the costs for the most efficient unit(s). We call these the *benchmark costs*. Thus, in general we must have:

$$a_1^\circ y_{1j} + a_2^\circ y_{2j} = x_j^\circ \leq x_j \qquad \text{for all } j, \tag{2}$$

where x_j° is the (unobserved) total cost associated with full efficiency, had it been achieved by unit j. Furthermore, if a_{1j} and a_{2j} are the actual cost rates for unit j (also unobserved), then the total cost function for unit j is:

$$a_{1j} y_{1j} + a_{2j} y_{2j} = x_j, \tag{3}$$

so that the inefficiency of unit j, if any, may be decomposed as:

$$s_j = x_j - x_j^\circ = (a_{1j} - a_1^\circ)y_{1j} + (a_{2j} - a_2^\circ)y_{2j}. \tag{4}$$

This decomposition suggests that it is possible to attain both benchmark costs simultaneously. Also, a measure of cost efficiency for the j-th unit is given by:

$$v_j = x_j^\circ / x_j. \tag{5}$$

Hence, inefficiency, as considered in this paper, occurs for such units by their having one or more unit cost rates higher than the corresponding benchmark cost rates. While it might also be the case that similar inefficiencies could exist with respect to the cost driver levels as well, we leave that case beyond the scope of the present paper. In the data set considered here the cost driver levels are not assumed to be under control of the units being compared.

Benchmarking and Its Importance

Many firms have voluntarily formed cooperative arrangements whereby they agree to share benchmarking information (Elnathan & Kim, 1995). Other firms have collaborated (at an expense to their individual firms) and hired a management consulting firm to study their costs. For example, in 1986, 16 major U.S. companies (subsequent additions increased the total number of companies to 26) hired A.T. Kearney to study their in-house accounting costs (Shank, 1993). Other firms have used databases of external management consulting firms for benchmarking purposes. For example, Lucent Technologies recently benchmarked various finance processes to 22 other large companies (the source of the benchmarking data was an outside consultant) in various industries (Francesconi, 1998). This benchmarking study revealed that Lucent's costs were significantly greater than those of the other benchmarked firms. A second source of benchmarking data is that of trade and industry associations (Kirby, 1988; Elnathan, Lin, & Young, 1996).

Many organizations that use benchmarking may incur significant related costs. For example, the American Accounting Association (AAA) recently initiated a benchmarking effort in which participating university accounting programs provide data and pay an $800 annual fee to receive benchmarking data in order to compare their program to peer programs (American Accounting Association, 2000). Likewise, those firms that hire a management consulting firm incur significant costs. In addition to the out-of-pocket costs incurred to participate in a benchmarking effort, there may be intangible costs associated with the sharing of proprietary information.

While there may be significant costs associated with benchmarking, there are also many benefits. Elnathan and Kim (1995) discussed three sources of changes in profits that may accrue when firms collaborate in cooperative benchmarking. First, firm profits may increase due to improvements in operations (increased productivity or reduced production costs). Second, a firm's competitive advantage within its industry may change firm profits due to information sharing. Third, there may be other political, social, or control-related effects of benchmarking.

In short, even though there may be significant costs associated with benchmarking, firms undertake benchmarking efforts because they view the benefits to be gained as outweighing the costs. Ideally, the benchmarking process should: (1) identify organizational units whose practices or procedures can be improved in terms of efficiency; (2) identify outstanding performers for emulation; (3) provide an operational measure of efficiency; and (4) provide performance targets in terms of unit costs of activities or similar figures to which management can relate. The method proposed here addresses these needs.

The Rates Departments' Data

The data for this study are attributed to Dyson and Thanassoulis (1988) and Thanassoulis, Dyson, and Foster (1987). The reader is referred to those papers for a detailed description. These data were collected for a set of 62 property tax collection offices, called *rates departments*, in the London Boroughs and Metropolitan Districts. Total annual costs, measured in units of £100,000 for these offices (units), were collected along with activity driver levels, called *outputs* in Dyson and Thanassoulis and Thanassoulis et al., for four activities. The first three activities—collection of non-council hereditaments, rates rebates generated, and summonses issued and distress warrants obtained—were measured in units of £10,000, £1,000, and £1,000, respectively. The fourth activity, net present value of non-council rates collected, was measured in units of £10,000. This last activity was included to reflect the additional administrative effort exerted to ensure the timely payment of large revenue-producing accounts. Thus, this data set gives total costs and cost drivers for four activities. Based on this data set we wish to determine the cost efficiencies of the units and estimate the activity units costs for the most efficient unit(s).

BENCHMARK MODELING APPROACHES

What we have called routine ABC appears to be the basis of most of the benchmarking studies. Two model-oriented approaches were found in the literature for benchmarking analysis. One of these is credited to Dyson and Thanassoulis (1988), and was based on modified data envelopment analysis (DEA, Charnes, Cooper, Lewin, & Seiford, 1994). Dopuch and Gupta (1997) recently proposed the use of stochastic frontier estimation (SFE) for estimating benchmark standards in a public education setting. Both of these approaches are reviewed in this section. Further discussion of the routine ABC approach is given below.

Thanassoulis et al. (1987) and Dyson and Thanassoulis (1988) considered the application of DEA to the data set considered here. The weights estimated in DEA models correspond to the cost estimates needed in this study. A shortcoming of DEA, from the viewpoint of this paper's goal, is that the particular weights that render one unit efficient may differ from those for another unit. This is called *weights flexibility* by Dyson and Thanassoulis, and is discussed further below. In addition, a relatively large number of units are typically declared fully efficient. Thus, no consensus on weights is achieved in DEA. Dyson and Thanassoulis proposed a heuristic modification of DEA. They first estimated average costs by regression through the origin as discussed above. Then, half of the respective average costs were used as reasonable lower bounds for values of the DEA weights. Although this approach reduces the weights flexibility, it is not removed entirely. Also, it may be argued, especially for benchmarking, that obtaining reasonable lower bounds on the weights (benchmark costs) is a central part of the problem. From that perspective, the Dyson and Thanassoulis method may be considered as somewhat ad hoc, even if reasonable. However, primary interest in that work was computation of efficiency scores with weights that do not vary as widely as those in unmodified DEA. While the present paper is also concerned with estimating efficiency scores, we seek consensus on the benchmark cost estimates as well.

Dopuch and Gupta (1997) proposed a benchmarking model using a Stochastic Frontier Estimation (SFE) method owing to Aigner, Lovell, and Schmidt (1977). They applied their model to evaluating the cost efficiency of a segment of the Missouri public school system in a data set similar to the one used here. In general, SFE models first define a parametric frontier model, which represents best possible performance, minimum or maximum depending on context. Then, actual performance is modeled as the frontier model plus an error term composed of two parts. The first error term is assumed to be normally distributed with mean zero. It is usually regarded as accounting for uncertainty in the frontier model. The second error term is a nonnegative one representing a measure of inefficiency or deviation from the efficient frontier. The Aigner et al. (1977) method assumes that such nonnegative inefficiencies are distributed as half-normal. The method proposed here does not require a preliminary assumption on the form of this density.

Other SFE approaches have also been proposed but do not appear to have been applied to cost efficiency as defined here. Green (1990) considered a model that assumes a gamma distribution for the inefficiency error terms. However, Ritter and Léopold (1997) have found that such models are difficult to accurately estimate. Recently, van den Broeck, Koop, Osiewalski, and Steel (1994) have also considered Bayesian SFE models.

MODEL DEVELOPMENT

Suppose we have $j = 1, \ldots, N$ comparable business units, achieving y_{rj} units of driver $r = 1, \ldots, R$, respectively, and with associated cost pools, x_j. In addition to having the same activities and cost drivers, we further require that comparable units be similar in the sense that the practices, policies, technologies, employee competence levels and managerial actions of any one should be transferable, in principle, to any other. Define a_r° as the cost rates associated with the most efficient unit or units under comparison. Then, in the equation

$$\sum_{r=1}^{R} a_r^\circ y_{rj} + s_j = x_j \quad \text{for all } j, \tag{6}$$

s_j may be interpreted as an inefficiency error term as in (4) and in the SFE models. The ratio $v_j = \sum a_r^\circ y_{rj} / x_j$, is an efficiency measure for the jth unit when the a_r° are the true benchmark cost rates, and $\sum a_r^\circ y_{rj} \le x_j$ holds for all units.

A technique for estimating parameters in efficiency ratios of the above form was proposed in Troutt (1995). In that paper the primary data were specific values of decision variables. Here, a more general form of that approach called *maximum performance efficiency* (MPE) is proposed and applied to estimate the benchmark a_r° values. Assume that each unit $j = 1, \ldots, N$, seeks to achieve maximum (1.0) efficiency. Then, the whole set of units may be regarded as attempting to maximize the sum of these efficiency ratios, namely, $\sum\sum a_r^\circ y_{rj} / x_j$.

The maximum performance efficiency estimation principle proposes estimates of the a_r° as those that render the total or, equivalently, the average of these efficiencies a maximum.

> **Maximum Performance Efficiency (MPE) Estimation Principle**: In a performance model depending on an unknown parameter vector, select as the estimate of the parameter vector that value for which the average performance efficiency ratio is greatest.

This estimation criterion is a variation of the maximum decisional efficiency (MDE) principle (Troutt, 1995). The MPE approach is technically the same as MDE but is applied to general performance measures rather than values of decision variables. The MDE principle assumes that decisions are made, that is, decision values are realized, in such a way as to maximize average efficiency relative to the model in question. MPE is a restatement in that performance vector values are realized in such a way as to maximize average efficiency relative to the model in question again, but where the model is expressed in terms of the performance vectors.

Define the data elements Y_{rj} by $Y_{rj} = y_{rj} / x_j$, where y_{rj} is the driver level (output value in DT) for the rth activity in operational unit j, and x_j is the total cost for unit j. For the rates departments' data we have $N = 62$ operational units and $R = 4$ cost drivers. The Y_{rj} are the performance vectors to which the MPE estimation method applies. We computed these values based on the data of DT. Then, the estimation model for the benchmark a_r° values is given by:

$$\text{MPE: max} \sum_{j=1}^{N} \sum_{r=1}^{R} a_r Y_{rj} \tag{7}$$

subject to

$$\sum_{r=1}^{R} a_r Y_{rj} \le 1 \quad \text{for all } j \tag{8}$$

$$a_r \geq 0 \text{ for all } r. \tag{9}$$

The MPE model is a linear programming (LP) problem whose unknown variables are the benchmark unit costs, the $a_r°$ values. Solution of this model provides values for the $a_r°$ as well as the unit efficiencies, v_j. The model was applied to the rates departments' data using the LP option of the Solver Tool in Microsoft Excel ™. The resulting estimates were:

$$a_1° = 0.00000,\ a_2° = 0.08820,\ a_3° = 0.2671,\ a_4° = 0.0664. \tag{10}$$

Unfortunately, there are two concerns about this raw MPE model solution. First, as part of their analysis, DT obtained an estimate of the average costs by regression through the origin as described above. Their results were:

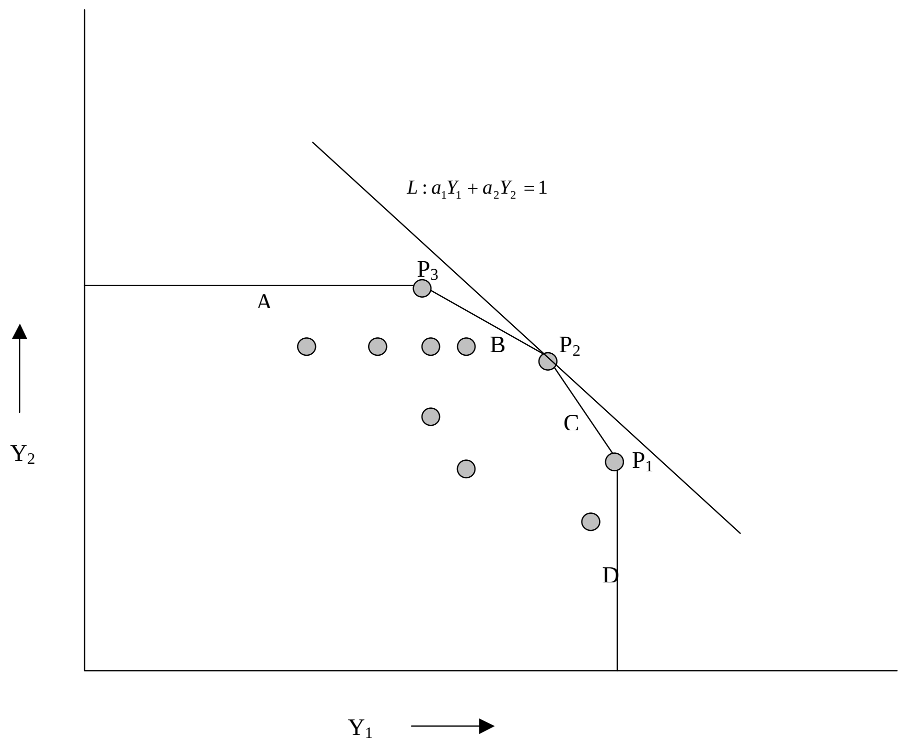
$$\bar{a}_1 = 0.05042,\ \bar{a}_2 = 0.07845,\ \bar{a}_3 = 0.1765,\ \bar{a}_4 = 0.1940. \tag{11}$$

Clearly, minimum (i.e., benchmark) activity unit cost estimates should not exceed average ones. However, it can be noted that the estimates for both $a_2°$ and $a_3°$ from the raw MPE model exceed the corresponding average cost estimates from the DT regression model.

A second solution quality issue is the zero value estimated for $a_1°$. In this paper, we consider an estimated benchmark cost of zero for an activity to be unreasonable. In the present data, activity 1 is a major activity that does, in fact, have positive average costs. While it may be conceivable that some unit can achieve one or more activities as cost-free by-products of one or more others, we believe that positive identification of such cases requires further research.

Thus, it was necessary to modify the raw MPE model to improve the solution quality of its estimates. To understand the modification, consider Figure 1, which depicts a two-dimensional version of the present data set but uses hypothetical data points. This figure indicates the efficient frontier formed by the three indicated DEA efficient units with data vectors at points P_1, P_2, and P_3. The linear functions defining the efficient frontier facets *A*–*D*, respectively, are the candidate basic feasible solutions to the raw MPE model. Suppose the solution associated with the frontier segment *D* is the raw MPE solution. This would be a vertical line with $a°_2 = 0$, namely, a zero parameter estimate. Also, this solution would assign only unit P_1 as having full efficiency. However, if the model is constrained to require that the unit at P_2 be efficient, then Line *L* depicts feasible values of a_1 and a_2 for the MPE model. The set of such feasible values has extreme points associated with rotation of *L* until it is coincident with frontier segments *B* or *C*, respectively. Both these segments have negative slope and, therefore, both parameter estimates will be positive. Thus, the modified MPE model is as follows. A separate run of the MPE model was made for each DEA efficient unit, requiring that unit to be fully efficient. The seven DEA efficient units were previously identified in Dyson and Thanassoulis (1988) and are shown in Table 1, Rows 1-7.

Figure 1: Hypothetical two-dimensional version of the rates departments data.



Note: Segments A, B, C, and D make up the efficient frontier found by DEA. Line L depicts feasible values of a_1 and a_2 for the MPE model when P_2 is required to be efficient. The set of such feasible values has extreme points associated with rotation of L until it is coincident with frontier segments B or C, respectively.

In each such run it was only necessary to change the inequality constraint for that unit to an equality constraint. The best of these seven solutions was identified by having all positive cost estimates and otherwise the maximal value of the objective function. Thus, the modified MPE model preemptively requires a maximal number of positive cost estimates. The efficiency values obtained are shown in Table 1, Column 5. The corresponding cost estimates were:

$$a^\circ_1 = 0.261807,\ a^\circ_2 = 0.049353,\ a^\circ_3 = 0.139833,\ a^\circ_4 = 0.127998. \qquad (12)$$

This solution passes the two tests of reasonableness above. Namely, all parameter estimates are positive, and no parameter value exceeds its average counterpart from the regression model. In addition, this model identifies four of the units as having maximal efficiency of unity. Four is the largest number of units that can be declared fully efficient since four linearly independent efficient points define a facet hyperplane of the efficient frontier for the dimensions of this data set.

We regard this as an additional measure of solution quality from the viewpoint of consensus. Four units that agree on the cost estimates is the best consensus achievable in this case. Thus, this solution has a number of desirable features. However, it is subjected to a further statistical test of reasonableness below.

Figure 1 also helps to highlight the differences between DEA and the proposed model. Application of DEA to this data set was fully discussed in Thanassoulis et al. (1987) and Dyson and Thanassoulis (1988), with resulting DEA efficiency scores as shown in Table 1. The seven fully DEA efficient units are in the first seven rows of the table. In terms of the hypothetical data set of Figure 1, DEA would identify the three fully DEA efficient units at P_1, P_2, and P_3. That is, DEA identifies the efficient frontier extreme points. Optimal DEA output weights would also be obtained for each unit. These correspond to the cost estimates and describe the coefficients in the lines that define the four facets, *A*–*D*, respectively. The solution of the modified MPE model corresponds to facet *C*. The coefficients of the linear function defining that facet are both positive and otherwise maximize the average efficiency of all the units. Thus, DEA is useful prior to applying the present model in order to determine the efficient frontier extreme points.

In the MPE model we have assumed only inefficiency error terms. This appears to be justified since we are able to specify the benchmark total cost function exactly, up to the benchmark cost parameters. Hence, there is no need to consider the additional normal error terms of the SFE approach.

It appears possible that there could exist data sets for which no solution of the modified MPE model yields all positive weights. For example, if the Y_{rj} data for some operational unit dominates all the other data, then the efficient frontier boundary of Figure 1 might resemble a rectangle. It that case, we would accept the estimates obtained from the solution of the unmodified MPE model, provided such solution is unique. Another possibility would be to consider the dominating operational unit for removal as a possible outlier. Failing these possibilities, it may not be possible to obtain estimates of benchmark costs along the present lines.

A TEST OF MODEL APTNESS

A basic assumption underlying the MPE estimation principle's applicability is that the sample of units under analysis do in fact have the goal of achieving maximum (1.0) efficiency. This is a model aptness issue that parallels the requirement of $N(0,\sigma^2)$ residuals in OLS regression theory. In the present MPE case, the corresponding issue is to specify a distribution characteristic of the v_j that indicates consistency with a goal or target of unity (1.0) efficiency. In this section we propose what may be called the *normal-like-or-better* effectiveness criterion for these fitted efficiency scores.

As a model for appropriate concentration on a target, we begin with an interpretation of the multivariate normal distribution, $N(\mu,\Sigma)$, on R^n. If a distribution of attempts has the $N(\mu,\Sigma)$ or even higher concentration of density at the mode μ, then we propose this as evidence that μ is indeed a plausible target of the attempts. This is exemplified by considering a distribution model for the results of throwing darts at a bull's-eye target. Common experience suggests that a bivariate normal density represents such data reasonably well. Steeper or flatter densities would still be

Table 1: Comparison of efficiency scores obtained by Data Envelopment Analysis (DEA), the modified DEA method of Dyson and Thanassoulis (1988) (DT), and the modified MPE model.

No.	Rates Department	DEA Efficiency	DT Efficiency	Modified MPE Efficiency
1	Lewisham	1.000	.827	.790
2	Brent	1.000	.743	.695
3	Stockport	1.000	1.000	1.000
4	Bradford	1.000	.999	1.000
5	Leeds	1.000	1.000	1.000
6	City of London	1.000	1.000	1.000
7	Liverpool	1.000	.796	.760
8	Walsall	.996	.861	.840
9	Rotherham	.994	.849	.795
10	Wakefield	.993	.890	.866
11	Lambeth	.961	.834	.816
12	Sunderland	.942	.801	.753
13	Solihull	.931	.917	.899
14	Redbridge	.847	.827	.814
15	Calderdale	.842	.818	.802
16	Haringey	.822	.710	.690
17	Barking and Dagenham	.801	.644	.610
18	Newcastle-upon-Tyne	.798	.713	.703
19	Manchester	.789	.641	.626
20	Wolverhampton	.782	.686	.667
21	Trafford	.761	.756	.751
22	Tameside	.759	.705	.683
23	St. Helens	.757	.694	.670
24	Sutton	.746	.692	.659
25	Rochdale	.745	.718	.696
26	Barnsley	.714	.617	.599
27	Kirklees	.713	.697	.690
28	Oldham	.702	.687	.679
29	Sheffield	.702	.702	.695
30	Havering	.700	.698	.695
31	Dudley	.700	.672	.659
32	Sefton	.690	.677	.664
33	Bexley	.688	.682	.669
34	Gateshead	.686	.621	.605
35	Wigan	.683	.652	.639
36	Kensington and Chelsea	.676	.587	.570
37	Coventry	.674	.645	.631
38	Sandwell	.644	.604	.593

Table 1: (continued) Comparison of efficiency scores obtained by Data Envelopment Analysis (DEA), the modified DEA method of Dyson and Thanassoulis (1988) (DT), and the modified MPE model.

No.	Rates Department	DEA Efficiency	DT Efficiency	Modified MPE Efficiency
39	Bury	.639	.638	.632
40	South Tyneside	.635	.526	.483
41	Salford	.629	.590	.581
42	Hackney	.614	.468	.445
43	Camden	.597	.562	.556
44	Hillingdon	.588	.587	.587
45	Tower Hamlets	.568	.529	.523
46	Barnet	.568	.567	.563
47	Bolton	.557	.549	.543
48	Ealing	.556	.545	.542
49	Bromley	.548	.520	.506
50	Wandsworth	.543	.524	.511
51	Birmingham	.535	.500	.491
52	Enfield	.516	.512	.505
53	Southwark	.509	.470	.464
54	Knowsley	.500	.487	.481
55	Islington	.496	.420	.411
56	North Tyneside	.465	.465	.461
57	Kingston-upon-Thames	.442	.426	.413
58	Hounslow	.435	.433	.430
59	Richmond-upon-Thames	.431	.410	.396
60	Hammersmith and Fulham	.424	.373	.364
61	Newham	.333	.331	.329
62	Merton	.329	.302	.286
Mean		.705	.652	.637
Standard Deviation		.187	.168	.166

indicative of effective attempts, but densities whose modes do not coincide with the target would cause doubts about whether the attempts have been effective or whether another target better explains the data. We call this normal-like-or-better (NLOB) performance effectiveness. It is next necessary to obtain the analog of this criterion for the efficiency performance data, Y_{rj}, relevant to the present context.

If x is distributed as $N(\mu, \Sigma)$ on R^n, then it is well known that the quadratic form, $w(x) = (x - \mu)' \Sigma^{-1}(x - \mu)$ has the gamma distribution, $g(\alpha,\beta)$, where $\alpha = n/2$, and $\beta = 2$. This distribution is also called the Chi-square distribution with n degrees of freedom (see Law & Kelton, 1982). We may note that for this case, $w(x)$ is in the nature of a squared distance from the target, here the singleton set {m}. It is useful to derive this result by a different technique. Vertical density representation

(VDR) is a technique for representing a multivariate density by way of a univariate density called the ordinate or vertical density, and uniform distributions over the equidensity contours of the original multivariate density. VDR was introduced in Troutt (1993). See also Troutt (1991), Kotz and Troutt (1996), Kotz, Fang, and Liang (1997), and Troutt and Pang (1996). The version of VDR needed for the present purpose can be derived as follows. Let $w(x)$ be a continuous convex function on R^n with range $[0,\infty)$; and let $g(w)$ be a density on $[0,\infty)$. Suppose that for each value of $u \geq 0$, x is uniformly distributed on the set $\{x\colon w(x) = u\}$. Consider the process of sampling a value of u according to the $g(w)$ density and then sampling a vector, x, according to the uniform distribution on the set $\{x\colon w(x) = u\}$. Next, let $f(x)$ be the density of the resulting x variates on R^n. Finally, let $A(u)$ be the volume (Lebesgue measure) of the set $\{x\colon w(x) \leq u\}$. Then, we have the following VDR theorem that relates $g(w)$ and $f(x)$ in R^n. The proof is given in the Appendix.

Theorem 1: If $A(u)$ is differentiable on $[0, \infty)$ with $A'(u)$ strictly positive, then x is distributed according to the density $f(x)$, where:

$$f(x) = \phi(w(x)),$$

and

$$g(w) = \phi(w) \,/\, A'(w). \tag{13}$$

Theorem 1 can be applied to derive a very general density class for performance related to squared distance type error measures. The set $\{x\colon (x - \mu)'\,\Sigma^{-1}(x - \mu) \leq u\}$ has volume, $A(u)$, given by $A(u) = \alpha_n\,|\Sigma|^{1/2}\ u^{n/2}$, where $\alpha_n = \pi^{n/2}/{}^{n}/_{2}\,\Gamma({}^{n}/_{2})$ (Fleming, 1977), so that $A'(u) = {}^{n}/_{2}\,\alpha_n\,|\Sigma|^{1/2}\,u^{\,n/2\,-\,1}$. The gamma $g(\mathrm{a},\beta)$ density is given by:

$$g(u)\ =\ (\Gamma(\alpha)\beta^{\alpha})^{-1}\,u^{\alpha\,-\,1}\,exp\{-u^2/\beta\}. \tag{14}$$

Therefore, Theorem 1 implies that if $w(x) = (x - \mu)'\Sigma^{-1}(x - \mu)$ and $g(u) = g(\alpha,\beta)$, then the corresponding $f(x)$, which we now rename as $\psi(x) = \psi(x;\, n,\alpha,\beta)$, is given by:

$$\psi(x) = \Gamma({}^{n}/_{2})(\pi^{n/2}\Gamma(\alpha)\beta^{\alpha})^{-1}[(x - \mu)'\Sigma^{-1}(x - \mu)]^{\alpha\,-\,n/2}$$
$$exp\{-{}^{1}/_{\beta}(x - \mu)'\,\Sigma^{-1}(x - \mu)\}. \tag{15}$$

For this density class we have the following observations:

(i) If $\alpha = n/2$ and $\beta = 2$, then $\psi(x)$ is the multivariate normal density, $N(\mu,\Sigma)$.

(ii) If $\alpha = n/2$ and $\beta \neq 2$, then $\psi(x)$ is steeper or flatter than $N(\mu,\Sigma)$ according to whether $\beta < 2$ or $\beta > 2$, respectively. We call these densities the normal-like densities.

(iii) If $\alpha < n/2$, then $\psi(x)$ is unbounded at its mode, μ, but may be more or less steep according to the value of β. We call this class the better-than-normal-like density class.

(iv) If $\alpha > n/2$, then $\psi(x)$ has zero density at the target, μ, and low values throughout neighborhoods of μ. This suggests that attempts at the target are not effective. The data may have arisen in pursuit of a different target or simply not be effective for any target.

For densities in category (iii), the unbounded mode concentrates more probability near the target and suggests a higher level of expertise than that evidenced by the finite-at-mode $N(\mu,\Sigma)$ class. It seems reasonable to refer to α in this context as the expertise, mode, or target effectiveness parameter, while β is a scale or precision parameter. Thus, if $\alpha \le n/2$, we call $\psi(x)$ the normal-like-or-better performance density. To summarize, if attempts at a target set in R^n have a basic squared distance error measure, and this measure is distributed with the $g(\alpha,\beta)$ density with $\alpha \le n/2$, then the performance with respect to this target set is normal-like-or-better (NLOB).

We extend this target effectiveness criterion to the present context as follows. The target set is $\{Y \in R^4\colon \Sigma a_r Y_r = 1,\ Y_r \ge 0 \text{ for all } r\}$ rather than the point set $\{\mu\}$. If $\Sigma a_r Y_{rj} = v_j$, then the distance of Y_{rj} from the target set is $(1 - v)\,\|a\|^{-1}$. Since $0 \le v \le 1$, we employ the transformation $w = (-\ln v)^2 = (\ln v)^2$. This transformation has the properties that $w \cong (1 - v)^2$ near $v = 1$, and $w \in [0,\infty)$. Therefore, $w/\|a\|^2 = (\ln v)^2\|a\|^2$ is an approximate squared distance measure near the target set. Since the $\|a\|^2$ term is a scale factor, it can be absorbed into the β parameter of $g(\alpha,\beta)$. We therefore consider the NLOB effectiveness criterion to hold if w has the $g(\alpha,\beta)$ density with $\alpha \le 4/2 = 2$. That is, such performance is analogous to that of unbiased normal-like-or-better distributed attempts at a target in R^n.

There is one additional consideration before applying this effectiveness criterion to the present data. In the LP estimation model, MPE, at least one efficiency, v_j, must be unity (and, hence, $w_j = 0$). This is because at least one efficiency inequality constraint must be active in an optimal solution of the MPE model. We therefore consider the model for the w_j to be:

$$p\delta(0) + (1 - p)g(\alpha,\beta), \tag{16}$$

where p is the frequency of zero values beyond 1 (here, $p = 3/62 = 0.048$ from Table 1), and $\delta(0)$ is the degenerate density concentrated at $w = 0$. For this data we regard the NLOB criterion to hold if it holds for the gamma density after omitting the zeroes. Thus, when the $g(\alpha,\beta)$ density is fitted to the strictly positive w values, then NLOB requires that $\alpha \le 2$. For the data of $w_j = (\ln v_j)^2$ based on Table 1, column 5, the parameter value estimates obtained by the Method of Moments (see, e.g., Bickell & Doksum, 1977) are $\alpha = 1.07$ and $\beta = 0.32$. This method was chosen because the BestFit™ (1995) software experienced difficulty in convergence using its default Maximum Likelihood Estimation procedure. The Method of Moments estimates parameters by setting theoretical moments equal to sample moments. For the gamma density, $\mu = \alpha\beta$, and $\sigma^2 = \alpha\beta^2$. If $\bar{w}$ and

s^2 are the sample mean and variance of the positive w_j values, then the α and β estimates are given by:

$$\hat{a} = \bar{w}^2 / s^2$$

and

$$\hat{\beta} = s^2 / \bar{w} . \qquad (17)$$

Tests of fit of the w_j data to the $g(\alpha = 1.07, \beta = 0.32)$ density were carried out using BestFit™ (1995). All three tests provided there, the Chi-square, Kolmogorov-Smirnov, and the Anderson-Darling indicated acceptance of the gamma model with confidence levels greater than 0.95. In addition, for each of these tests, the gamma model was judged best fitting (rank one) among the densities in the library of BestFit™. We therefore conclude that the NLOB criterion was met. The NLOB criterion is important in establishing whether the estimated cost model is a plausible goal of the units being studied. The MPE model will produce estimates for any arbitrary set of Y_{rj} data. However, if the resulting v_j data were, for example, uniformly distributed on [0,1], there would be little confidence in the estimated model.

The requirement that $g(v)$ be a gamma density is not critical. This assures, together with an acceptable alpha parameter value, that the performance vectors have a normal-like distribution, with mode equal to the target set estimated by the model. A more general requirement with the same effect is that the performance vectors be distributed as unimodal, with mode coinciding with the target set. If, for example, $g(v)$ has a Weibull density, then the steps of Theorem 1 can be reversed to obtain the $f(x)$ density. The resulting density can be checked for unimodality, and its mode can, in principle, be directly compared to the target set.

DISCUSSION

In this section we discuss the managerial significance of the results. We also briefly discuss the weights flexibility issue in DEA, longitudinal use of the approach, and implications for the single driver-single cost pool case. In addition, we mention some limitations and questions needing further research.

Few specifics on how to use the results of benchmark modeling analyses have been reported in the literature. It appears that routine activity-based costing has been the method used in most such studies. In that case, shared information consists of unit costs for various common activities and benchmarking amounts to seeing which operational unit is best on each activity unit cost. However, an operational unit indicated to be best on one cost might not, in general, be best on other activity unit costs. Hence, routine ABC may not be able to indicate the overall best performers and provide an overall rating of cost efficiency for the operational units being compared. With the present modeling approach, cost efficiency ratings are assigned to all the operational units, and best performers are identified for possible transferability of practices and procedures. In the present results, the four rates departments, Stockport, Bradford, Leeds, and City of London were identified as cost efficient. Hence, their practices and procedures should be considered for adoption, where possible, by the managements of less efficient operational units. It may happen that some operational unit other than these four could have a lower unit cost on activity 1, say, by direct comparison of results from routine ABC. However, managements of other operational units should be cautious to imitate such an operational unit. Assuming the validity of the model results, such an operational unit is not fully efficient and, therefore, may have performed poorly on one or more other activity unit costs.

A question of accuracy also arises in the routine ABC approach. As noted by Dopuch (1993), ABC is generally considered to result in "better" cost estimates simply because there are more cost pools and cost drivers. Such a conclusion may not be warranted without knowledge of the "true" cost numbers. In fact, Datar and Gupta (1994) have shown that disaggregation can increase errors. Thanassoulis et al. (1987) briefly discussed the possible disaggregation of the total costs for the rates departments' data. They indicated that this would have been possible to some extent but decided against this for several reasons. Among these, it was felt that the disaggregated data were less reliable. Therefore, additional tools like the proposed method may be useful as a means of signaling possible errors from disaggregation. Also, disaggregation may be expensive or impractical. We may also conjecture that firms may be more willing to share data at the aggregation level needed for the modeling approach than at a more detailed level. Therefore, the proposed multiple cost driver approach provides management with a benchmarking tool that may save costs as well as improve accuracy and promote wider sharing of data.

Weights Flexibility

From Table 1 it is clear that the proposed method is a more stringent measure of efficiency than DEA. Comparing the proposed method only to DEA, it can be seen that for all units, efficiency scores are largest for DEA and smallest for the proposed method. Also, in each case except Bradford, the Dyson and Thanassoulis (1988) efficiencies were between these bounds. This is as expected since DEA permits maximal weights flexibility, and the proposed method permits no weights flexibility. Thus, when used with DEA, a range estimate of the efficiency of each unit is obtained. The proposed method assumes that all units are comparable and, therefore, should have the same minimal unit cost goals. If this is not the case, then DEA results should be considered as possibly more valid for one or more operational units.

Longitudinal Data

Suppose the data Y_{rt} are given over time periods indexed by t for a single business unit. Then, the MPE model with index j replaced by t might be applied as a model for internal benchmarking. First, it would be necessary to adjust all the x_t cost pool figures and resulting Y_{rt} data to reflect current dollars using a cost index. This assumes that the estimated a_r° cost rates are in terms of current dollars. Then, the estimated a_r° may be interpreted to be the costs achieved by the unit during its most efficient observation period or periods. The resulting v_t suggest periods of more or less efficiency, and would be a useful source for self-study aimed at productivity and process improvements. The comparability issue for the units under comparison should be easier to accept in this case. However, process or technology and market or other environmental changes during the data time span could be problematical. A more complete discussion of limitations for this case is left for specific future application case studies.

Implications for the Single Driver-Single Cost Pool Case

The MPE model for this case simplifies to

$$\max \sum aY_j,$$

subject to

$$aY_j \leq 1, \text{ for all } j, \text{ and } a \geq 0.$$

The solution of this model is clearly $a° = \min Y_j^{-1}$. It may be verified that the NLOB criterion requires $\alpha° \leq 1/2$ in this case. If this condition fails to hold, then this minimum value may be unreasonably low, perhaps due to an outlier. Deletion of one or a few tentative outliers would be well supported if the remaining data do, in fact, pass the NLOB test. Otherwise, no credible $a°$ estimate is forthcoming from the present method.

Limitations and Further Research

In order to more fully parallel existing OLS theory for model aptness testing, attention should be given to potential outliers, independence of the v_j transformations, and constancy of the distribution of the v_j from trial to trial (analogous to homoscedasticity in OLS theory). See, for example, Madansky (1988) and Neter, Wasserman, and Kutner (1985). Theory developments for these issues are not yet available for the MPE approach and would be worthwhile for future research.

Thanassoulis et al. (1987) also discussed what we have called comparability of these units. A concern was noted relative to activity four, whose monetary driver level might have been affected by the prosperity of the community being served. That is, units with above average property values might be considered as being unfairly compared to the others. Other things being equal, units with an inappropriately inflated value of a driver level would be expected to exert a downward influence on the corresponding estimate in model MPE. We believe this kind of potential incomparability might be avoided by use of a property value index for future research.

We have assumed that the benchmark unit costs should be strictly positive in this paper. As noted above, it may be possible that for some situations and units that an activity can be achieved as a cost-free by-product of others. It was pointed out to us that this might not necessarily be inconsistent with the comparability assumption when the practice or process responsible for such an advantage is transferable to the other units. The unadjusted MPE model appears capable of indicating such situations. However, it should be validated on a known case of this type for further research.

CONCLUSIONS

This paper proposes a new method for estimating cost efficiencies and benchmark unit costs. The results provide a new tool for benchmarking studies in activity-based costing. The estimated costs provide plausible operational goals for the management of the units being compared. This method also provides efficiency measures and suggests which organizational units or time periods are more or less efficient, as well as an estimate of the degree of such inefficiency. Efficient units or time periods provide benchmarks for imitation by other units, or can be studied for continuous improvement possibilities. We also employ a new estimation technique that does not require prior specification of the distribution of the inefficiency errors as is necessary in the stochastic frontier approaches. A model aptness criterion was proposed for the new technique.

The proposed estimation approach was applied to a real data set previously analyzed by a modified data envelopment analysis method. The resulting estimates were compared with the average costs obtained by the previous method. The estimated benchmark cost rates were uniformly and strictly lower than their average rate counterparts, consistent with their definitions, and providing a strong measure of face validity. [Received: September 27, 1999. Accepted: October 23, 2000.]

REFERENCES

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

American Accounting Association. (2000). [Online]. Available: http://AAA-edu.org, July 8, 2000.

Banker, R. D., & Johnston, H. H. (1993). An empirical study of cost drivers in the U.S. airline industry. *The Accounting Review, 68*(3), 576-601.

BestFit. (1995). *User's guide.* Newfield, NY: Palisade Corporation.

Bickell, P. J., & Doksum, K. A. (1977). *Mathematical statistics: Basic ideas and selected topics.* San Francisco: Holden Day.

Charnes, A., Cooper, W. W., Lewin, A., & Seiford, L. M. (1994). *Data envelopment analysis: Theory, methodology, and applications.* Boston: Kluwer Academic Publishers.

Datar, S., & Gupta, M. (1994). Aggregation, specification and measurement errors in product costing. *The Accounting Review, 69*(4), 567-591.

Dopuch, N. (1993). A perspective on cost drivers. *The Accounting Review, 66*(3), 615-620.

Dopuch, N., & Gupta, M. (1997). Estimation of benchmark performance standards: An application to public school expenditures. *Journal of Accountancy and Economics*, *23*, 141-161.

Dyson, R. G., & Thanassoulis, E. (1988). Reducing weight flexibility in data envelopment analysis. *Journal of the Operational Research Society, 39*(6), 563-576.

Elnathan, D., & Kim, O. (1995). Partner selection and group formation in cooperative benchmarking. *Journal of Accounting and Economics*, *19*, 345-364.

Elnathan, D., Lin, T. W., & Young, S. M. (1996). Benchmarking and management accounting: A framework for research. *Journal of Management Accounting Research*, *8*, 37-54.

Fleming, W. (1997). *Functions of several variables* (2nd ed.). New York: Springer-Verlag.

Francesconi, T. A. (1998). Transforming Lucent's CFO. *Management Accounting*, July, 22-30.

Green, W. H. (1990). A gamma-distributed stochastic frontier model. *Journal of Econometrics*, *46*, 141-163.

Horngren, C. T., Foster, G., & Datar, S. M. (2000). *Cost accounting: A managerial emphasis* (10th ed.). Upper Saddle River, NJ: Prentice-Hall.

Kirby, A. J. (1988). Trade associations as information exchange mechanisms. *Rand Journal of Economics*, *19*, 138-146.

Kotz, S., Fang, K. T., & Liang, J. T. (1997). On multivariate vertical density representation and its application to random number generation. *Statistics*, *30*, 163-180.

Kotz, S., & Troutt, M. D. (1996). On vertical density representation and ordering of distributions. *Statistics*, *28*, 241-247.

Law, A. M., & Kelton, W. D. (1982). *Simulation modeling and analysis*. New York: McGraw-Hill.

Madansky, A. (1988). *Prescriptions for working statisticians*. New York: Springer-Verlag.

Neter, J., Wasserman, W., & Kutner, M. H. (1985). *Applied linear statistical models* (2nd ed.). Homewood, IL: Richard E. Irwin.

Ritter, C., & Léopold, S. (1997). Pitfalls of normal-gamma stochastic frontier models. *Journal of Productivity Analysis, 8*, 167-182.

Shank, J. K. (1993). How safe is your job? *Journal of Accountancy*, October, 72-80.

Thanassoulis, E., Dyson, R. G., & Foster, M. J. (1987). Relative efficiency assessments using data envelopment analysis: An application to data on rates departments. *Journal of the Operational Research Society*, *38*(5), 397-411.

Troutt, M. D. (1991). A theorem on the density of the density ordinate and alternative derivation of the Box-Muller method. *Statistics*, *22*, 436-466.

Troutt, M. D. (1993). Vertical density representation and a further remark on the Box-Muller method. *Statistics, 24*, 81-83.

Troutt, M. D. (1995) A maximum decisional efficiency estimation principle. *Management Science, 41*(1), 76-82.

Troutt, M. D., & Pang, W. K. (1996). A further VDR-type density representation based on the Box-Muller method. *Statistics, 28*, 1-8.

van den Broeck, J., Koop, G., Osiewalski, J., & Steel, M.F.J. (1994). Stochastic frontier models: A Bayesian perspective. *Journal of Econometrics, 61*, 273-303.

APPENDIX

Proof of Theorem 1

This is a modification of a proof for a version of the theorem given in Troutt (1993). By the assumption that x is uniformly distributed on $\{x: w(x) = u\}$, $f(x)$ must be constant on these contours, so that $f(x) = \varphi(w(x))$ for some function, $\varphi(\cdot)$. Consider the probability $P(u \le w(x) \le u + \varepsilon)$ for a small positive number, ε. On the one hand this probability is $\varepsilon g(u)$ to a first-order approximation. On the other hand, it is also given by:

$$\int_{\{x:\, u \le w(x) \le u+\varepsilon\}} \!\!\cdots\! \int f(x)\Pi dx_i \;\cong\; \varphi(u) \int_{\{w:\, u \le w \le u+\varepsilon\}} \!\!\cdots\! \int \Pi dx_i$$

$$\cong \varphi(u)\{A(u+\varepsilon) - A(u)\}$$

Therefore,

$$\varepsilon\, g(u) \;\cong\; \varphi(u)\{A(u+\varepsilon) - A(u)\}.$$

Division by ε and passage to the limit as $\varepsilon \to 0$ yields the result.

Marvin D. Troutt is director of the Center for Information Systems, Department of Management & Information Systems, and professor in the Graduate School of Management at Kent State University. He received the PhD in mathematical statistics from The University of Illinois at Chicago. He is an associate editor of *Decision Sciences*. His publications have appeared in *Decision Sciences, Management Science, Operations Research, Journal of the Operational Research Society, European Journal of Operational Research, Decision Support Systems*, and elsewhere. His current research interests include the statistics of performance data, data mining, and decision support.

Donald W. Gribbin is an associate professor of accountancy at Southern Illinois University at Carbondale. He received the PhD in business administration from Oklahoma State University. He has published papers in *Decision Sciences*, *Journal of Management Accounting Research*, *Journal of Business Finance and Accounting*, and *The British Accounting Review*. His research interests include the distributional properties of financial ratios and the modeling of various types of cost data.

Murali S. Shanker is an associate professor in the Department of Management & Information Systems at Kent State University. He received his PhD from the Department of Operations and Management Science, University of Minnesota. His research interests are in knowledge management and data mining, simulation, and distributed processing. His papers have appeared in journals such as *INFORMS Journal on Computing, IIE Transactions*, *Omega*, and *Annals of Operations Research*. His current research includes developing tools to optimize web-server performance, and in conducting statistical analyses of the behavior of neural networks as applied to classification and prediction problems.

Aimao Zhang is a candidate for the doctorate in Business Administration at Southern Illinois University at Carbondale, Illinois. Her work has previously appeared in *Annals of Operations Research*. Her research interests include organizational transformation, industrial economy, and computing related to operations research and production management.