MANAGEMENT SCIENCE
Vol. 32, No. 12, December 1986
Printed in U.S.A.

THE USE OF CATEGORICAL VARIABLES IN DATA ENVELOPMENT ANALYSIS*

RAJIV D. BANKER AND RICHARD C. MOREY
School of Urban and Public Affairs, Carnegie-Mellon University, Pittsburgh, Pennsylvania 15213
Fuqua School of Business, Duke University, Durham, North Carolina 27706

Data Envelopment Analysis has now been extensively applied in a range of empirical settings to identify relative inefficiencies, and provide targets for improvements. It accomplishes this by developing peer groups for each unit being operated. The use of categorical variables is an important extension which can improve the peer group construction process and incorporate "on-off" characteristics, e.g., presence of drive-in window or not in a banking network. It relaxes the stringent need for factors to display piecewise constant marginal productivities. In so doing, it substantially strengthens the credibility of the insights obtained. The paper treats the cases when the categorical variable can be controllable or uncontrollable by the manager, for the cases of technical and scale inefficiency. The approach is illustrated using real data.

1. Introduction

Data Envelopment Analysis (DEA) has now been extensively applied in a range of empirical settings to identify technical inefficiencies of decision making units (DMU's), and provide targets for improvement for inefficient DMU'S. Technical inefficiency of a DMU is determined relative to other similar units and can focus on either resource conservation or output augmentation. Consider first the definition of a technically inefficient unit in the context of resource conservation where the objective is to minimize the consumption of resources, given a particular output level. Thus, a given operating unit (DMU) is considered to be technically inefficient in the context of resource conservation if some other unit, or some convex combination of other units, can: (i) produce at least the same amounts of all outputs; (ii) use less of at least one controllable resource input and not more of any other controllable resource; and (iii) accomplish the above with at least the same difficulty in terms of environmental factors. Conversely, a unit is said to be technically efficient on a relative basis in the context of resource conservation if the above is not possible. If the focus is on output augmentation, a unit is considered to be technically inefficient if some other unit or some convex combination of other units can: (i) use no more of any of the controllable inputs; (ii) produce at least the same amounts of all outputs and more of at least one output; and (iii) accomplish the above with at least the same difficulty in terms of environmental factors.

The initial relative efficiency analysis of Farrell (1957) was cast in terms of a ratio formulation by Charnes, Cooper and Rhodes (1978); they also provided equivalent linear programming formulations to evaluate efficiency. Banker (1980b, 1984) and Banker, Charnes and Cooper (1984) provided an axiomatic production economics framework for relative efficiency evaluation in multiple output production settings. Separate linear programming formulations, depending on whether the focus was on input conservation or on output augmentation, were developed to assess technical and scale efficiencies, returns to scale and most productive scale size. The analysis was further extended by Banker and Morey (1984) to distinguish between controllable and non-controllable or fixed resources utilized in the production or conversion process.

* Accepted by Arie Y. Lewin; received February 9, 1984. This paper has been with the authors 7 months for 2 revisions.

0025-1909/86/3212/1613$01.25

DEA has particular appeal in that it deals with multiple outputs and multiple inputs and does not require *a priori* or subjective tradeoffs between various types of outputs or the use of prices for aggregating the resources. Further, the method uses standard LP codes to identify peer groups for each unit being evaluated. Using as a reference these peer group members, DEA provides quantitative insights as to the aspects and sizes of adjustments needed to render an inefficient unit efficient. Applications of this efficiency analysis technique to criminal superior courts (Lewin et al. 1982), Armed Forces recruiting districts (Lewin et al. 1981), school districts (Bessent et al. 1983), pharmacies (Capettini et al. to appear), hospitals (Banker, Conrad and Strauss 1986), electric power generation plants (Banker 1984), manufacturing productivity analysis (Banker 1985), etc., also demonstrate the flexibility of DEA.

The DEA approach for determining relative technical inefficiency is based on building a composite DMU which is a convex combination of other DMUs' inputs and outputs. This assumption of convexity is equivalent to assuming that if two production possibilities are observed in practice, then any production plan which is a convex weighted combination of the two production possibilities is also achievable. This convexity assumption, together with minimum extrapolation (see Appendix B), manifests itself in that DEA estimates the efficient production frontier in a piecewise linear fashion. DEA is then an extremal prediction method which estimates (in the resource conservation formulation) the minimum level of resources needed for a DMU, faced with a given environment, to produce a set of required outputs.

However, suppose we are attempting to estimate the resources (such as labor and capital) that a branch of a bank needs to obtain a given level of deposits, given a population base of say 100,000, with a specific income, age, and other demographic characteristics. Then in DEA the branch in question might well be compared to a composite branch built from a branch with a population of 80,000 and another with a population of 120,000, both weighted equally. While this may seem like a very reasonable approximation, it is clear that the branches employed for this comparison would be less controversial if we were to insure that the peer group consisted *only* of branches with a population of 100,000 or less. We may allow a branch operating in an even more difficult situation, such as a branch with a population base of say only 80,000, to be included in the peer group, but not one with a larger population, since we might be unwilling to assume that the marginal productivity associated with this factor were the same for a branch with a population of 120,000 as that for the branch with a population of 100,000.

The above considerations become even more important when one realizes that some factors in such relative efficiency analysis are 0-1 variables: some branches of a bank may have a drive-in capability and some others may not; some branches may have automatic tellers and some may not. What is desired for the above situations is a method for insuring that the composite reference members be constructed from DMUs which are in the same category or possibly from those in a category which is deemed to be operating in an even more difficult or unfavorable situation. For example, suppose branch banks are rated in terms of their competitive environment as "mild," "medium," or "difficult" competition. Then a particular branch of a bank, with a "difficult" competition rating, should only be compared to other branches with the same rating. To do otherwise, say based on some sort of scale, would imply that a constant marginal productivity applies. At the same time, a branch with a "mild" rating could possibly have its peer group composed of branches with "mild," "medium," or "difficult" competition ratings if these branches outperformed the branch bank in question, in spite of the more difficult competitive situation faced by them. In this paper, we shall develop modified models to accomplish the above goals, i.e., for selected factors and given scalings we will relax the requirement in DEA models that a constant marginal productivity situation applies.

This will also serve to extend the nonparametric regularity tests suggested by Diewert and Parkan (1984) for testing the consistency of the hypothesis of productive efficiency of all DMU's with observed input-output data and a monotonic nondecreasing and concave (or quasi-concave) production function. In the multiple output case, their tests for quasi-concavity of the production function require the asymmetrical singling out of an output and input in turn. Our approach permits the simultaneous consideration of quasi-concavity for a subset of the inputs [i.e., the function $-X_s = f(-Y; X_{-s})$ is quasi-concave, where X_s is the vector of all inputs $i \in S \subseteq \{1, \cdots m\}$ and X_{-s} is the vector of all inputs other than inputs $i \in S$] and concavity for the remaining inputs. Although our focus is on the estimation of efficient production frontiers, the consistency of the productive efficiency hypothesis with the specified regularity conditions for the production function is tested easily by checking whether *each* DMU has an efficiency rating of one.

The remainder of the paper is structured as follows: §2 contains modifications to the mathematical formulation for evaluating technical inefficiency when a noncontrollable variable is considered categorical. A noncontrollable variable is one over which the manager of the DMU exercises essentially no control. It may be a key demographic (such as median income or population), a cost of living measure, competition (such as dollars of advertising by the competition) or even a resource used by the DMU when the level of that resource is not managed locally at the DMU level; an example of the latter may be the level of national advertising impacting a given branch where headquarters (and not the DMU) controls that particular facet. §3 compares for a real data set the results of treating a particular noncontrollable variable in the usual continuous fashion versus the categorical approach of §2. This section illustrates, at least for this example, that very little discriminating power is lost, while greatly increasing the credibility of the findings. §4 generalizes this categorical variable approach to the case where both technical and scale inefficiencies are of interest. The empirical illustration in the preceding section is continued to deal with this additional facet. Finally, §5 provides a mixed integer LP formulation for evaluating technical inefficiency when a controllable output variable is treated as categorical. An example of this situation is the level of customer service that a branch bank chooses to provide.

2. Mathematical Development for the Case of Noncontrollable Categorical Variables

We consider N DMU's indexed by $j = 1, 2, \ldots, j_0, \ldots, N$, where it is desired to assess the j_0th DMU's relative performance. The rth type of output for the jth DMU is denoted by $\{y_{rj}; r = 1, 2, \ldots, R; j = 1, 2, \ldots, N\}$ and the ith input (to be conserved) for the jth DMU is denoted $\{x_{ij}; i = 1, 2, \ldots, m; j = 1, \ldots, N\}$. Further, we index the inputs so that the first m' inputs (denoted $i = 1, 2, \ldots, m'$) are controllable according to above discussion (where controllable is to be understood as capable of being varied by the manager of the DMU) and the other $(m - m')$ inputs (denoted $i = m' + 1, \ldots, m$) are non-controllable. Banker and Morey (1984) provide the following LP formulation for assessing the relative technical efficiency of the j_0th DMU, when none of the inputs or outputs are categorical. The following formulation is in the context of resource conservation:

$$\text{Minimize } h_0 = Z_0 - \epsilon\Big(\sum_{i=1}^{m'} s_i^- + \sum_{r=1}^{R} s_r^+\Big) \tag{1}$$

subject to:

$$\sum_{j=1}^{N} \lambda_j X_{ij} + s_i^- = Z_0 X_{ij_0} \qquad (i = 1, 2, \ldots, m'), \tag{2}$$

$$\sum_{j=1}^{N} \lambda_j X_{ij} + s_i^- = X_{ij_0}, \qquad i = (m'+1, m'+2, \ldots, m), \tag{3}$$

$$\sum_{j=1}^{N} \lambda_j Y_{rj} - s_r^+ = Y_{rj_0}, \qquad (r = 1, 2, \ldots, R), \tag{4}$$

$$\sum_{j=1}^{N} \lambda_j = 1, \tag{5}$$

$$\lambda_j \geq 0 \qquad (j = 1, 2, \ldots, j_0, \ldots, N), \tag{6}$$

$$s_i^- \geq 0; \qquad s_r^+ \geq 0 \qquad (i = 1, 2, \ldots, m; r = 1, 2, \ldots, R),$$

where $\epsilon > 0$ is small "non-Archimedean" number. (In practice ϵ is chosen to be a very small number, such as 10^{-6}. Alternatively, this prioritized minimization problem can be solved in two stages, as described by Banker and Morey 1984.)

The above formulation is consistent with the axiomatic framework for DEA provided by Banker, Charnes and Cooper (1984). This is written in terms of four postulates: convexity, monotonicity, inclusion of observations and minimum extrapolation (see Appendix B), which in turn determine a unique production possibility set based on the observed input-output data. Further this unique set can be employed to provide points of comparison for the DMU being evaluated. Note also that the optimal level of the variable Z_0, denoted Z_0^* from the first stage LP, (1)–(6) provides the so-called Farrell (1957) score which is a radial measure of inefficiency in that each controllable resource can be simultaneously contracted by the factor Z_0^*. The maximum resource conservation potential for the ith controllable resource is given in terms of the following target for a lower level of resource consumption:

$$X'_{ij_0} = Z_0^* X_{ij_0} - s_i^{-*} \qquad (i = 1, 2, \ldots, m'). \tag{7}$$

For completeness the corresponding formulation, in the context of output augmentation and in the presence of exogenously fixed outputs (see Banker and Morey 1984), is presented in the Appendix B.

Now focusing on the resource conservation formulation of (1)–(6), consider the case where one of the *noncontrollable* inputs, say $i = m$, is a *categorical* variable, i.e., a variable which has a finite number of values that the variable can take on. The convexity postulate (see Appendix B.2(i)) provides that all convex combinations of the N referent points may be used as points to evaluate the relative performance of the j_0th DMU. But for a categorical variable, a convex combination: $\sum_{j=1}^{N} \lambda_j x_{mj}$, $\lambda_j \geq 0$, $\sum_{j=1}^{N} \lambda_j = 1$, need not belong to the categorical scale, and it may not even have a meaningful interpretation.

Further, the convexity axiom, along with the minimum extrapolation axiom (see Appendix B.2(iv)), implies a piecewise linear surface for the production possibility set and, therefore, that the marginal productivity is constant on each such linear segment. However, for a categorical scale, this assumption of piecewise constant marginal productivity may not be valid. Hence, we modify the Banker-Charnes-Cooper postulates by deleting the operability of the convexity postulate for the categorical variable. In this situation, the production possibility set T will then be given by:

$T = \{(X, Y)$ where $X = (X_1, X_2, \ldots, X_m)$ and $Y = (Y_1, Y_2, \ldots, Y_R)$

where there exists a set $\{\lambda_j; j = 1, 2, \ldots, N\}$ characterized by $\lambda_j \geq 0$

$(j = 1, 2, \ldots, N)$ and: $\sum_{j=1}^{N} \lambda_j = 1$ such that: $Y_r \leq \sum_{j=1}^{N} \lambda_j Y_{rj}$ $(r = 1, 2, \ldots, R)$,

$$X_i \geq \sum_{j=1}^{N} \lambda_j X_{ij} \; (i = 1, 2, \ldots, m-1) \text{ and: } X_m \geq X_{ml} \text{ for all those indices } l \text{ for which } \lambda_l > 0\}. \quad (8)$$

To elaborate on (8), we observe that for the case of the noncontrollable category variable, the linear constraint of (3) namely,

$$\sum_{j=1}^{N} \lambda_j X_{mj} + s_m^- = X_{mj_0}$$

is to be replaced by *constraints* of the form:

$$X_{ml} \leq X_{mj_0} \quad \text{if} \quad \lambda_l > 0, \text{ i.e.,} \quad (9)$$

for all DMU's l which are used in the composite convex combination (i.e., for which λ_l is strictly positive), the level[1] of their category variable must be at the same or lower level than is the case for the j_0th DMU.

To convert (9) into a form suitable for our LP model, we shall define k new variables $d_{m,j}^{(\alpha)}$, where $k + 1$ is the number of values the categorical variable x_m can take on. For concreteness, suppose there are 4 distinct levels for this categorical variable, which are denoted as "none", "low", "average", and "high", where "high" denotes the most favorable situation. Then we shall define three descriptor binary variables $\{d_{m,j}^{(1)}, d_{m,j}^{(2)}, d_{m,j}^{(3)}\}$ for each of the DMU's to replace the continuous variables $x_{m,j}$ $(j = 1, 2, \ldots, j_0, \ldots, N)$. If the jth DMU is actually in the "none" level for the mth input, we capture this by setting $d_{m,j}^{(1)} = d_{m,j}^{(2)} = d_{m,j}^{(3)} = 0$; if the jth DMU is in the "low" level, its descriptor variables are given by:

$$d_{m,j}^{(1)} = 1, \qquad d_{m,j}^{(2)} = d_{m,j}^{(3)} = 0; \quad (10)$$

if the jth DMU is in the "average" category, then:

$$d_{m,j}^{(1)} = 1 = d_{m,j}^{(2)}, \qquad d_{m,j}^{(3)} = 0; \quad (11)$$

and finally, if the jth DMU is in the "high" category, we have:

$$d_{m,j}^{(1)} = d_{m,j}^{(2)} = d_{m,j}^{(3)} = 1. \quad (12)$$

The modifications to the LP modeling of (1)–(6) are accomplished by deleting from the constraint set (3), namely,

$$\sum_{j=1}^{N} \lambda_j x_{m,j} + s_m^- = x_{m,j_0}$$

and replacing it by 3 new constraints, namely:

$$\sum_{j=1}^{N} \lambda_j d_{m,j}^{(1)} \leq d_{m,j_0}^{(1)}, \quad (13)$$

$$\sum_{j=1}^{N} \lambda_j d_{m,j}^{(2)} \leq d_{m,j_0}^{(2)}, \quad (14)$$

[1] Diewert and Parkan (1984) employed a similar idea when they operationalized the notion of technical progress for their nonparametric tests for consistency of observed production data with postulated regularity conditions for the production function. However, they considered only one categorical variable, viz. time. Therefore, they did not consider the identification and the invalidation of the observations that are inappropriate as referent points for each DMU, in situations when only a partial ordering of the data is possible because of the presence of more than one categorical variables. Furthermore, they focused strictly on consistency tests. They did not address the question of efficiency evaluation when the observed data are not consistent with their efficiency postulate.

$$\sum_{j=1}^{N} \lambda_j d_{m,j} \le d^{(3)}_{m,j_0}. \tag{15}$$

Observe that if the j_0th DMU is the lowest state, i.e., $d^{(1)}_{m,j_0} = d^{(2)}_{m,j_0} = d^{(3)}_{m,j_0} = 0$, then constraints (13), (14), and (15) imply that:

$$\sum_{j=1}^{N} \lambda_j d^{(k)}_{m,j_0} \le 0 \qquad \text{for} \qquad k = 1, 2, 3. \tag{16}$$

Hence the composite group (i.e., those l for which $\lambda_l^* > 0$) must consist only of DMU's for which all the descriptor values are also 0, i.e., the reference set of DMU's must be composed of all "none" rated DMU's. Similarly, if the j_0th DMU is in the "low" or second category, then:

$$\sum_{j=1}^{N} \lambda_j d^{(1)}_{m,j} \le 1, \qquad \sum_{j=1}^{N} \lambda_j d^{(2)}_{m,j} \le 0, \qquad \sum_{j=1}^{N} \lambda_j d^{(3)}_{m,j} \le 0, \tag{17}$$

which implies that λ_j^* can only be positive for those DMU's which are also in the "none" or "low" categories.

3. Empirical Illustration

In this section we illustrate the technique using real data and assess the sensitivity of the results from treating a noncontrollable variable as a continuous variable versus treating it as a categorical variable. It is easy to see that the latter categorical approach limits more stringently the DMUs which can be a member of each DMU's peer group. Therefore, the technical efficiency scores will tend to be somewhat higher, but for the DMU's still identified as technically inefficient, we have improved the credibility of the peer group.

The data set consists of 69 pharmacies in the State of Iowa, with data obtained for 2 outputs, and 4 input factors, 3 of which were controllable. The outputs were the annual number of prescriptions and dollar level of prescriptions dispensed over the year. The inputs were: the total labor cost for the year; other operating costs related to the prescription department; average size of the inventory carried; and market size (as measured by the population of the city served by the given pharmacy).

For the 69 pharmacies, the population of the city ranged from a few hundred to over 200,000, with a mean of 41,500. In one treatment, the market size, a non-controllable factor, was treated as a continuous variable, using the actual population sizes. Under this treatment we evaluated the relative technical efficiency of each of the 69 pharmacies. In the other treatment, we categorized the market size variable into 11 categories, based on the classification used by the U.S. Census of Population and Housing, 1980. The relevant categories were:

Category	Population
1	Under 200
2	200–499
3	500–999
4	1,000–1,499
5	1,500–1,999
6	2,000–9,999
7	10,000–19,999
8	20,000–24,999
9	25,000–49,999
10	50,000–99,999
11	100,000–249,999

TABLE 1

Output and Input Data for Two DMU's

	Pharmacy #15 (one being evaluated)	Pharmacy #9 (dominates DMU #15)
(1) Number of Prescriptions	26,830	33,800
(2) Dollars of Sales	$178,513	$183,000
(3) Total Labor Cost in Prescription Department	$53,848	$31,760
(4) Other Operating Costs	$14,208	$5,559
(5) Average Prescription Inventory Value	$59,750	$27,000
(6) Population Being Served	2,500 (Category 6)	1,694 (Category 5)

The levels of the 2 outputs and 4 input factors for the 69 pharmacies are presented in the Appendix A, together with their averages and ranges. The technical efficiency evaluations were repeated with the only change being the treatment of market size, now considered a categorical variable.

As mentioned earlier, the number of DMUs identified as technically inefficient will not increase under the categorical treatment, since the reference groups are more restrictive, i.e., the reference group consists only of DMU's with the same or lower categories of market size. The empirical analysis confirms that indeed this is the case, but by using the more defensible categorical approach, the DEA, at least in this case, loses relatively little discriminating power. Similarly the adjustments needed under the two treatments, for the inefficient operations to become relatively efficient, are quite close; however this may not always be the case.

To illustrate concretely, we present the differences in results for one DMU, say for $j_0 = 15$. We first observe that the 15th pharmacy is indeed relatively technically ineffi-

TABLE 2

Comparisons of Technical Efficiencies and Efficient Targets for DMU #15: Continuous versus Categorical Treatment of Market Size Variable

Characteristics of 15th Pharmacy	Actual	Level Needed to Eliminate Technical Inefficiency When Population Treated As Continuous	Level Needed to Eliminate Technical Inefficiency When Population Treated As Categorical
TECHNICAL EFFICIENCY SCORE		0.534	0.561
OUTPUTS			
(1) Number of Prescriptions Dispensed	26,830	34,333	32,894
(2) Dollar Value of Prescriptions Dispensed	$178,513	$178,513	$178,513
INPUTS			
(1) Labor Costs	$59,823	$31,953	$33,536
(2) Other Operating Costs	$14,208	$4,908	$5,517
(3) Average Value of Drug Inventory Carried	$59,750	$21,946	$26,216
(4) Population Served (Market Size)	2,500 (Category 6)	—	—
PEER GROUP	—	Composite Pharmacy Consists of DMU's 9, 12, and 44 with Respective Weights of 0.386, 0.312, and 0.302.	Composite Pharmacy Consists of DMU's 9 and 37 with Respective Weights of 0.956 and 0.044.

TABLE 3

Characteristics of Members of Composite Reference Groups Used for the Technical Efficiency Rating of DMU #52 (population of 23,166, category 8)

Treatment of Market Size

	Continuous	Categorical
Relative Technical Efficiency Score	0.654	0.856
Reference Group and its Characteristics	DMU #1 (category #4) DMU #30 (category #6) DMU #41 (category #9) DMU #44 (category #6) DMU #66 (category #11)	DMU #9 (category #5) DMU #12 (category #4) DMU #33 (category #6)

cient since Pharmacy 9 (see Table 1) dominates it on every dimension, i.e., #9 produces more of each of the outputs, using less of each resource in a more difficult environment. Indeed the 9th pharmacy turns out to be in the peer group of the 15th pharmacy, regardless of which treatment is used. We also observe, as expected, that the technical efficiency score is slightly smaller (0.534 vs. 0.561) when the continuous treatment is used since the continuous treatment is less constraining. Note also that the adjustments (see Table 2) needed for the 15th pharmacy to be deemed efficient are quite close, e.g., a reduction to $33,536 in labor cost under the categorical treatment versus $31,953 from the continuous approach. However, to concretely show that the categorical approach

TABLE 4

Comparison of Technical Efficiency Scores under Two Modeling Treatments of Market Size: Continuous and Categorical

	Continuous Treatment	Category Treatment
(1) Number of Inefficient Units	41 inefficient	36 inefficient
(2) Lowest Technical Efficiency	0.403 (unit #23)	0.451 (unit #47)
(3) Percent change in terms of resource conservation or output augmentation for controllable factors		
(i) Input #1—Labor costs	18.34% reduction over all 69 units	16.85% reduction over all 69 units
(ii) Input #2—Other Operating Costs	25.77% reduction	22.45% reduction
(iii) Input #3—Average value of inventory	20.25% reduction	17.60% reduction
(iv) Output #1—Number of prescriptions	4.43% increase	3.54% increase
(v) Output #2—Dollar levels of sales for prescriptions	1.84% increase	1.21% increase
	Average Change	Number of Units
(4) Average change in technical efficiency score		
(i) For those with a continuous treatment score of 1.0	no change	28
(ii) For those with a continuous treatment score of between 0.800 and 0.999	+0.025	13
(iii) For those with a continuous treatment score of between 0.700 and 0.799	+0.068	14
(iv) For those with a continuous treatment score of less than 0.700	+0.034	14

can produce quite different results, consider the comparison of the results of the relative technical efficiency for DMU #52 under the two treatments for market size (see Table 3). DMU #52 with a population of 23,166 had a rating of 0.654 under the continuous approach compared to .856 under the categorical approach. The composite reference group under the continuous treatment consisted of DMU's 1, 30, 41, 44, and 66 with respective λ weights of 0.438, 0.133, 0.333, 0.040, 0.056 versus the composite group of DMU's 9, 12, and 33 and λ weights of 0.293, 0.407, and 0.300 for the categorical approach.

We observe that the reference group for DMU #52 (with a population of 23,166) under the continuous treatment used DMU #41 (with a population of 30,213) and DMU #66 with a population of 201,404, nearly ninefold that of DMU #52. The use of this questionable peer group under the continuous treatment allowed the technical score to be pulled down substantially over that obtained under the categorical approach, where only DMU's with a category of 8 or less could be included in the composite group.

Next consider the results for all 69 pharmacies. When market size was treated as a continuous variable, 41 (of the 69) pharmacies were technically inefficient, compared to 36 when the more stringent category approach was used. When market size was treated as a continuous variable, the DMU with the lowest technical efficiency score was DMU #23 with an index of 0.403; when population was treated through the categorical approach, the least efficient DMU was #47 with a score of 0.451. Table 4 also shows that the changes in the scores for various groupings of the inefficient DMU's are relatively small; but more importantly, the overall reductions in resources necessary for the pharmacies to be rated relatively technically efficient are close, e.g., 18.34% in labor costs for the continuous versus 16.85% for the categorical approach.

4. Generalization of Categorical Treatment for Identification of Technical/Scale Inefficiencies

Banker (1980b, 1984), Banker, Charnes and Cooper (1984) and Banker and Morey (1984) (for the case of noncontrollable inputs) have shown that the combined technical and scale inefficiency of the j_0th DMU can be evaluated by modifying the LP's in (1)–(6) by removing the convexity constraint of (5) and replacing the constraints in (3) by the following constraints:

$$\sum_{j=1}^{N} \lambda_j x_{ij} + s_i^- - x_{ij_0} \cdot \sum_{j=1}^{N} \lambda_j = 0 \qquad (i = m' + 1, \ldots, m) \tag{18}$$

where $\lambda_j \geq 0$ for $j = 1, 2, \cdots N$. It can be shown easily, following the arguments in Banker (1984) and Banker, Charnes and Cooper (1984) that if $\sum_{j=1}^{N} \lambda_j^* < 1$ in the linear program defined by the objective function in (1) and the constraints in (2), (15), (4) and (6) then the j_0th DMU is in the increasing returns to scale region relative to the controllable inputs and outputs. Similarly $\sum_{j=1}^{N} \lambda_j^* > 1$ indicates that the j_0th DMU is in the decreasing returns to scale region relative to the controllable inputs and outputs. If the sum of the weights equals unity, then Banker (1984) has shown that the j_0th DMU is at the most productive scale of operation. It is further shown that the combined technical scale inefficiency is the product of the technical inefficiency and scale inefficiency.

The categorical modeling performed earlier for the case of technical inefficiency is similar to that used when technical and scale inefficiency is of interest. We shall first consider the results for the 15th DMU again. Table 5 displays the comparison of the results for DMU #15 when the market size variable is treated as continuous versus when it is treated as categorical. We observe in Table 5 that the scores are all very

TABLE 5

Comparison of Combined Technical and Scale Efficiencies for DMU #15 Continuous vs. Categorical Treatment of Market Size

	Continuous Treatment	Categorical Treatment
(1) Combined Technical and Scale Efficiency Score	0.521	0.532
(2) Technical Efficiency Score	0.532	0.561
(3) Scale Efficiency Score	0.975	0.948
(4) Returns to Scale	Decreasing Returns $(\sum_{j=1}^{69} \lambda_j^* = 1.899)$	Decreasing Returns $(\sum_{j=1}^{69} \lambda_j^* = 1.529)$
(5) Adjusted Levels for 15th Pharmacy to be Technically and Scale Efficient		
(a) No. of Prescriptions Dispensed (Actual of 26,830)	20,981*	20,414*
(b) Dollars of Prescriptions Dispensed (Actual of $178,513)	$94,011	$116,713
(c) Labor Costs (Actual of $59,823)	$16,404	$ 20,816
(d) Other Operating Costs (Actual of $14,208)	$ 9,968	$ 15,416
	Composite Pharmacy Consists of DMU's 12, 41, 44 with respective weights of 1.750, 0.066, and 0.083.	Composite Pharmacy Consists of DMU's 9 and 37 with respective weights 0.545 and 0.984.

* We observe that to be both technically and scale efficient, the 15th DMU should operate at a reduced scale of operation, consuming less resources but also producing lower levels of outputs.

similar, and the decreasing returns are detected by both approaches. It is also interesting to observe the change in the peer groups.

In Table 6, we display the average results for all 69 pharmacies. There were 52 DMU's identified as having some technical or scale inefficiency under the continuous approach versus 43 under the categorical approach. Hence, there was some loss in discriminating power. Note that the average change in the index from the two approaches is small, as are the percent changes needed to eliminate all of the inefficiencies. Finally, under the continuous approach, 31 of the DMU's were estimated to be in the increasing returns portion of the production frontier. Under the categorical approach, 24 were judged to be in the area of increasing returns.

5. Mixed-Integer LP Formulation for Case Where Categorical Variable Is Controllable

We have seen in §2 that an input factor (such as market size) which is a noncontrollable factor can be modeled as a categorical variable in a manner which insures that the composite DMU will consist only of DMU's which are in the same or more difficult categories. Also, we have seen that a simple LP code can still be utilized to evaluate technical efficiencies and adjustments needed to the controllable factors for the unit to be technically, or technically and scale efficient. Now consider the evaluation of relative technical efficiency when the category variable is deemed to be controllable. We shall see that a mixed-integer LP formulation will be needed in this case.

To illustrate the approach, we shall consider a third output of a pharmacy which is controllable by the pharmacy, namely, its service orientation—the extent to which it provides additional services such as 24-hours-open, home delivery, pharmacist avail-

TABLE 6

Comparison of Technical and Scale Efficiency Scores

	Continuous Treatment	Categorical Treatment
(1) Number of Inefficient Units	52 inefficient	43 inefficient
(2) Lowest Efficiency Score	0.292 (unit #23)	0.292 (unit #23)
(3) Percent change in terms of resource conservation or output augmentation for controllable factors.		
(i) Input #1—Labor costs	11.5% reduction over all 69 units	13.8% reduction over total for all 69 units
(ii) Input #2—Other operating costs	35.59% reduction	37.1% reduction
(iii) Input #3—Average value of inventory	13.61% reduction	17.0% reduction
(iv) Output #1—Number of prescriptions	21.39% increase	11.9% increase
(v) Output #2—Dollar levels of sales for prescriptions	18.53% increase	10.4% increase
(4) Percent of inefficient DMU's with increasing returns indicated	31/52 (59.6%)	24/43 (55.8%)
	Average Change	**Number of Units**
(5) Average absolute change in efficiency score		
(i) For those with a continuous treatment score of 1.0	No change	17
(ii) For those with a continuous treatment score of between 0.800 and 0.999	+0.037	13
(iii) For those with a continuous treatment score of between 0.700 and 0.799	+0.037	13
(iv) For those with a continuous population score of less than 0.700	+0.40	21

able for consultation with physicians, maintenance of tax records, etc. Suppose that each pharmacy can be rated as having one of L different categorical service orientations where the categories range from category 1 (representing the lowest service orientation, such as might be present in a discount pharmacy operation) to category L, representing the highest service orientation. Further, we shall illustrate the approach in the context of output augmentation (see Appendix B) where it is assumed that the *sole* intention is to identify the maximum gain possible in the categorical service level of a pharmacy. For simplicity, we shall treat market size as a continuous variable. Then one defines parameters $\{w_{l,j}\colon l = 1, 2, \ldots, L-1\}$ as a service descriptor vector, one for each pharmacy j, where the lowest service category is represented by the vector of 0's, the second lowest service category is represented by the vector $(1, 0, \ldots, 0)$, the third lowest by the vector $(1, 1, 0, 0, 0)$ and the highest service category by a vector of 1's.

The problem is formulated as the following mixed-integer LP model:

$$\text{Maximize} \sum_{l=1}^{L-1} t_l \tag{19}$$

subject to

$$\sum_{j=1}^{N} \lambda_j x_{rj} \le x_{ij_0}, \qquad i = 1, 2, 3, 4, \tag{20}$$

$$\sum_{j=1}^{N} \lambda_j Y_{rj} \geq Y_{rj_0}, \qquad r = 1, 2, \tag{21}$$

$$\sum_{j=1}^{N} \lambda_j w_{l,j} - t_l = w_{l,j_0}, \qquad l = 1, \ldots, L-1, \tag{22}$$

$$t_{l-1} \geq t_l, \qquad l = 2, \ldots, L-1, \tag{23}$$

$$\sum_{j=1}^{N} \lambda_j = 1, \qquad \lambda_j \geq 0, \qquad j = 1, 2, \ldots, N, \tag{24}$$

where t_l's are 0-1 variables.

The constraints in (22) are similar to the corresponding constraints in (13)–(15), and ensure that DMUs included in the composite referent point are necessarily of the same or higher service orientation. However additionally we must insure that the improvements in the service orientation are identified sequentially; improvements in service orientation for the j_0th DMU can only occur by sequentially converting the elements of the service vector to 1's; but with the restriction that if a particular element of the service descriptor vector, say the $(l-1)$th slot is 0, then the element in the lth slot must also have a value of 0. The additional set of constraints in (23), together with the fact that the t_l's are 0-1 variables, insures that if t_{l-1} is 0, then $t_l = 0$.

6. Summary

This paper introduces the use of categorical variables, both for noncontrollable and controllable inputs or outputs, into the DEA approaches for the evaluation of the technical efficiency and the combined technical and scale efficiency of decision making units. It relaxes the need to assume that all factors are measured on a continuous scale and display piecewise constant marginal productivity on this continuous scale. It strengthens the credibility of the insights obtained about the improvements possible because the referent composite DMU's consist only of DMU's which have been matched more carefully with the DMU being evaluated. Straightforward LP techniques can still be used if the categorical variable is noncontrollable; however if the category variable is a controllable factor, then mixed-integer LP approaches are needed to insure that the resource conservation or output augmentation results are well defined and meaningful.

Appendix A

Data Set for 69 Pharmacies

Store Number	Prescription Volume	Dollar Value of Prescription (in Thousands)	Average Prescription Inventory	Labor Cost	Other Operating Costs	Population	Population Category
1	12,293	61.00	$ 8,000	$ 17,030	$ 1,280	1,410	4
2	18,400	92.00	9,000	25,890	2,779	1,523	5
3	16,789	92.65	13,694	29,076	2,372	1,354	4
4	10,700	45.00	4,250	17,506	1,385	822	3
5	9,800	50.00	6,500	23,208	639	746	3
6	6,500	29.00	7,000	12,946	802	1,281	4
7	8,200	56.00	4,500	18,001	1,130	1,016	4
8	8,680	45.00	5,000	14,473	1,097	1,070	4
9	33,800	183.00	27,000	31,760	5,559	1,694	5
10	23,710	156.00	21,560	50,972	15,010	1,910	5
11	24,000	120.00	15,000	39,523	4,799	1,745	5

Appendix A (*Continued*)

Data Set for 69 Pharmacies

Store Number	Prescription Volume	Dollar Value of Prescription (in Thousands)	Average Prescription Inventory	Labor Cost	Other Operating Costs	Population	Population Category
12	17,500	75.00	8,500	13,076	3,489	1,353	4
13	25,000	130.00	35,000	35,427	1,704	500	3
14	26,000	122.00	18,000	27,554	2,882	1,016	4
15	26,830	178.513	59,750	53,848	14,208	2,500	6
16	16,600	106.00	19,200	38,253	1,480	2,293	6
17	90,000	450.00	40,000	109,404	83,016	2,718	6
18	11,140	73.624	8,466	18,198	1,278	2,877	6
19	25,868	136.00	16,000	40,891	7,599	4,150	6
20	32,700	191.295	10,000	45,444	5,556	4,421	6
21	29,117	152.864	25,000	35,623	2,121	3,883	6
22	18,000	100.00	14,000	20,192	5,515	3,519	6
23	11,100	60.00	12,500	34,973	10,475	32,366	9
24	23,030	137.778	17,260	32,284	14,498	3,393	6
25	10,656	58.00	7,000	17,920	7,585	4,489	6
26	24,682	152.095	14,000	42,094	3,742	2,217	6
27	26,908	120.00	16,400	35,422	14,236	4,641	6
28	16,464	80.00	13,000	19,100	3,529	5,968	6
29	57,000	321.00	30,000	72,167	8,656	8,715	6
30	17,532	94.747	12,530	19,970	1,714	5,968	6
31	30,035	168.00	31,500	39,183	4,919	5,607	6
32	16,000	100.00	10,000	32,048	3,483	7,324	6
33	63,700	277.00	22,000	68,877	12,279	8,685	6
34	18,000	90.00	10,000	29,812	3,332	8,685	6
35	27,339	139.134	16,000	47,686	2,507	5,420	6
36	19,500	116.00	10,000	33,415	4,738	7,703	6
37	13,000	80.00	9,000	12,359	4,603	4,665	6
38	15,370	102.00	16,439	23,614	2,989	6,317	6
39	18,446	90.00	14,500	36,069	1,793	31,839	9
40	56,000	260.00	39,000	76,307	9,539	15,619	7
41	73,845	364.951	24,927	40,706	12,661	30,213	9
42	28,600	145.00	13,858	39,267	4,609	34,719	9
43	27,000	243.00	33,375	29,509	11,323	31,839	9
44	52,423	279.816	29,044	44,482	5,542	34,719	9
45	73,759	363.388	32,257	61,365	20,550	32,366	9
46	20,500	80.00	8,800	49,671	3,306	43,561	9
47	27,100	115.00	47,000	40,425	10,396	31,263	9
48	15,000	110.00	12,000	33,034	4,915	31,263	9
49	50,895	277.852	28,000	69,163	4,688	15,173	7
50	19,707	128.00	13,300	28,931	16,735	73,064	10
51	17,994	78.80	13,500	29,758	4,260	62,309	10
52	36,135	167.222	24,000	40,927	8,285	23,166	8
53	30,000	153.00	16,000	40,403	2,131	99,836	10
54	26,195	125.00	17,000	38,730	2,539	60,348	10
55	28,000	216.00	25,000	35,978	2,502	99,836	10
56	24,658	152.551	16,000	37,509	6,278	99,836	10
57	36,850	190.00	20,000	46,950	10,715	85,925	10
58	29,250	183.69	14,000	35,966	3,144	85,925	10
59	50,000	250.00	22,000	68,318	8,015	108,987	11
60	40,078	265.443	21,879	69,537	7,778	108,987	11
61	20,200	110.00	15,000	25,425	2,812	201,404	11
62	12,500	75.00	10,000	19,508	2,454	201,404	11
63	30,890	195.00	20,000	28,191	3,367	201,404	11
64	31,000	175.00	18,000	37,073	8,624	108,987	11
65	31,277	192.992	19,051	23,763	3,496	201,404	11

Appendix A (*Continued*)

Data Set for 69 Pharmacies

Store Number	Prescription Volume	Dollar Value of Prescription (in Thousands)	Average Prescription Inventory	Labor Cost	Other Operating Costs	Population	Population Category
66	11,500	75.00	15,000	28,642	3,366	201,404	11
67	30,000	175.668	10,000	35,919	3,868	201,404	11
68	38,383	190.00	24,000	54,653	26,494	108,987	11
69	2,075	8.650	1,800	6,276	3,413	60,348	10
Averages	27,133	147.199	17,976	38,486	7,081	41,500	9
Minimum	2,075	8.650	1,800	6,276	639	500	3
Maximum	90,000	450.000	59,750	113,404	83,016	201,404	11

Appendix B

(B.1)

The corresponding formulation to (1)–(6) in the context of output augmentation is given by:

$$\max \Phi_0 + \epsilon\left(\sum_{r\in\nu_D} s_r^- + \sum_{i=1}^{m} s_i^+\right) \tag{B.1}$$

subject to

$$\sum_{j=1}^{N} \lambda_j Y_{rj} - s_r^- = \Phi_0 Y_{rj_0}, \qquad r \in \nu_D, \tag{B.2}$$

$$\sum_{j=1}^{N} \lambda_j Y_{rj} - s_r^- = Y_{rj_0}, \qquad r \in \nu_F, \tag{B.3}$$

$$\sum_{j=1}^{N} \lambda_j X_{ij} + s_i^+ = X_{ij_0} \qquad (i = 1, 2, \ldots, m), \tag{B.4}$$

$$\sum_{j=1}^{N} \lambda_j = 1, \qquad \Phi_0, \lambda_j, s_r^-, s_i^+ \geq 0, \tag{B.5}$$

where ν_D is the set of indices of output under the discretionary control of the DMU manager and ν_F is the set of indices of exogenously fixed outputs. Note that $\nu_D \cap \nu_F = \varnothing$.

(B.2) *The Banker-Charnes-Cooper Postulates*

The estimated production possibility set T must satisfy:

(i) *Convexity Postulate.* If (X_1, Y_1) and $(X_2, Y_2) \in T$, and $\lambda \geq 0$ then $(\lambda X_1 + (1-\lambda)X_2, \lambda Y_1 + (1-\lambda)Y_2) \in T$.

(ii) *Monotonicity Postulate.* (a) If $(X_0, Y_0) \in T$ and $Y \leq Y_0$ then $(X_0, Y) \in T$.

(b) If $(X_0, Y_0) \in T$ and $X \geq X_0$ then $(X, Y_0) \in T$.

(iii) *Inclusion of Observations Postulate.* Each of the observed vectors $(X_j, Y_j) \in T$ $(j = 1, \cdots N)$.

(iv) *Minimum Extrapolation Postulate.* T is the intersection set of all $\hat{T}$ satisfying the earlier three postulates.

If T satisfies the above four postulates, then T can be expressed as:

$$T = \{(X, Y) | X \geq \sum_{j=1}^{N} \mu_j X_j,\ Y \leq \sum_{j=1}^{N} \mu_j Y_j,\ \sum_{j=1}^{N} \mu_j = 1, \mu_j \geq 0\}.$$

References

BANKER, R. D., "A Game Theoretic Approach to Measuring Efficiency," *European J. Oper. Res.*, (October 1980a), 262–266.

——, "Studies in Cost Allocation and Efficiency Evaluation," unpublished doctoral thesis, Harvard University, Graduate School of Business Administration, 1980b.

——, "Productivity Accounting: An Application and Conceptual Extensions," submitted, *Accounting Rev.*

——, "Estimating Most Productive Scale Size Using Data Envelopment Analysis," *European J. Oper. Res.*, (September 1984), 35–44.

——, "Productivity Measurement and Management Control," *The Management of Productivity and Technology in Manufacturing*, P. Kleindorfer (Ed.), Plenum, New York, 1985.

——, "Fuel Productivity and Scale Efficiency of Coal-Fired Steam-Electric Plants," Working Paper, Carnegie-Mellon University, 1984.

——, A. CHARNES AND W. W. COOPER, "Models for Estimation of Technical and Scale Inefficiencies in Data Envelopment Analysis," *Management Sci.*, 30, 9 (September 1984), 1078–1092.

——, ——, —— AND A. SCHINNAR, "A BiExtremal Principle for Frontier Estimation and Efficiency Evaluation," *Management Sci.*, (December 1981), 1370–1382.

——, R. F. CONRAD AND R. P. STRAUSS, "An Application of Data Envelopment Analysis to the Empirical Investigation of a Hospital Production Function," *Management Sci.*, (1986).

—— AND A. MAINDIRATTA, "Piecewise Loglinear Estimation of Efficient Production Surfaces," submitted to *Management Sci.*, (1983).

——, "Nonparametric Analysis of Technical Allocative Efficiency in Production," Carnegie-Mellon University, Working Paper (1985).

—— AND R. MOREY, "Efficiency Analysis with Exogeneously Fixed Inputs and Outputs," *Oper. Res.*, (to appear) (1984).

BESSENT, A., W. BESSENT, J. KENNINGTON AND B. REAGAN, "An Application of Mathematical Programming to Assess Productivity in the Houston Independent School District," *Management Sci.*, (1983).

CAPETTINI, R., D. A. DITTMAN AND R. C. MOREY, "Reimbursement Rate Setting for Medicaid Prescription Drugs Based on Relative Efficiencies," *J. Accounting and Public Policy*, (to appear).

CHARNES, A., W. W. COOPER AND E. RHODES, "Measuring the Efficiency of Decision Making Units," *European J. Oper. Res.*, (1978), 429–444.

——, —— AND ——, "Short Communication: Measuring the Efficiency of Decision Making Units," *European J. Oper. Res.*, (1979), 339.

——, —— AND ——, "Evaluating Program and Managerial Efficiency," *Management Sci.*, (1981), 668–697.

DIEWERT, E., AND C. PARKAN, "Linear Programming Tests of Regularity Conditions for Production Functions," in *Quantitative Studies on Production and Prices*, W. Eichorn, R. Henn, K. Newmann, and R. W. Shephard (Eds), Physics-Verlag, 1984.

FARRELL, M. J., "The Measurement of Productive Efficiency," *J. Roy. Statist. Soc., Ser. A*, (1957), 253–290.

—— AND M. FIELDHOUSE, "Estimating Efficient Production Frontiers Under Increasing Returns to Scale," *J. Roy. Statist. Soc., Ser. A*, (1962), 252–267.

LEWIN, A. Y., R. C. MOREY AND T. J. COOK, "Evaluating the Administrative Efficiency of Courts," *Omega* (1982), 401–411.

—— AND ——, "Measuring the Relative Efficiency of Output Potential of Public Sector Organizations: An Application of Data Envelopment Analysis," *J. Policy Analysis and Information Systems*, (December 1981), 267–285.

SHEPHARD, R. W., *The Theory of Cost and Production Functions*, Princeton University Press, Princeton, N.J., (1970).