The Journal of Productivity Analysis, 4, 419–432 (1993)

Some Extensions of Techniques to Handle Non-Discretionary Factors in Data Envelopment Analysis*

BOAZ GOLANY
Faculty of Industrial Engineering and Management, Technion—Israel Institute of Technology, Haifa 32000, Israel
IC^2 Institute, The University of Texas at Austin, 2815 San Gabriel, Austin, TX 78705-3596

YAAKOV ROLL‡
Faculty of Industrial Engineering and Management, Technion—Israel Institute of Technology, Haifa 32000, Israel

Abstract

Data Envelopment Analysis (DEA) assumes, in most cases, that all inputs and outputs are controlled by the Decision Making Unit (DMU). Inputs and/or outputs that do not conform to this assumption are denoted in DEA as *non-discretionary* (ND) factors. Banker and Morey [1986] formulated several variants of DEA models which incorporated ND with ordinary factors. This article extends the Banker and Morey approach for treating non-discretionary factors in two ways. First, the model is extended to allow for the *simultaneous* presence of ND factors in both the input and the output sets. Second, a generalization is offered which, for the first time, enables a quantitative evaluation of *partially* controlled factors. A numerical example is given to illustrate the different models.

1. Introduction

Charnes, Cooper and Rhodes-CCR [1978] published a breakthrough paper which gave birth to the Data Envelopment Analysis (DEA) approach, a new methodology to evaluate the relative efficiencies of Decision Making Units (DMUs). The CCR model is solved separately for each DMU. Its data are input-output observations on a number of DMUs, all using the same types of inputs (but not necessarily in the same amounts) to produce the same types of outputs. The problem that the original CCR model solves can be stated as follows: find a set of virtual prices of outputs and inputs (weights) so as to maximize the (efficiency) ratio of the weighted sum of outputs to the weighted sum of inputs for the DMU being evaluated, while ensuring that using the same weights, no other DMU will have an efficiency ratio that is larger than one.

In determining the relative efficiency of DMUs, DEA provides an estimate for the projection of inefficient DMUs onto an "efficiency frontier." These projections involve input reductions or output enhancements, or both. DEA assumes that the evidence provided by the efficient DMUs (those that constitute the frontier), makes such projections feasible. The developments proposed in this article deviate from this assumption by accepting that some of the factors (inputs and/or outputs) cannot be changed at the discretion of the respective DMUs. The meaning of a non-discretionary (ND) factor, in the DEA context, is that

*The editor for this paper was Wade D. Cook.

DMUs have no control over it, even though this factor is still of importance in assessing relative efficiencies among DMUs. On the input side, it may be instrumental in the DMU's ability to generate outputs and on the output side, it may limit possible reduction in the inputs. Otherwise, if it did not play a significant role in the process of converting inputs to outputs, it would have been excluded from the analysis.

As the DEA methodology has developed, several alternative efficiency definitions have been suggested. However, while competing variants of the basic DEA model (see, e.g., Seiford and Thrall [1990]) have proliferated in the last decade, problems associated with ND inputs and outputs have received far less attention. Charnes et al. [1980] coined the term *non-discretionary* inputs. This issue was further discussed in Charnes and Cooper [1985] and Charnes et al. [1985] where directions for possible formulations of the efficiency evaluation problem in the presence of ND factors were laid out. In particular, the topics of simultaneous treatment of ND inputs and outputs and partially controlled ND factors were presented there as open questions. Banker and Morey-BM [1986] addressed the ND question by proposing various alterations to the original DEA formulations. Their approach, which has essentially become the standard way to handle the ND problem in DEA, is the basis of the extensions and generalizations offered in this article. Ray [1988], offered a different interpretation to the same phenomenon. He argues that the production function can be viewed as multiplicatively separable in the discretionary and ND factors. He attributes any technical inefficiency to the existence of ND inputs.

In a previous publication, Golany and Roll [1989], the authors proposed an application procedure for implementing DEA. In this article we extend that procedure by focusing on ND input and output factors in DEA. Various DEA applications, see Seiford [1990], provide ample motivation to the necessity of such an extension. In real-world applications, whether in production or service situations, where a measure of relative efficiency of different DMUs is sought, instances in which some inputs or outputs are exogenously fixed, albeit still relevant to the analysis, are abundant. For example, in applications relating to the service industry (e.g., schools, courts, hospitals, etc.) the size of the population serviced, as well as its socio-economic characteristics, are crucial factors in the evaluation, but they are not at the discretion of the DMUs. Likewise, in private sector applications (e.g., fast food restaurants, commercial banks, etc.) the number of competitors and the intensity of the competition are important ND factors. In some applications (e.g., Air-Force Wings, Farming, etc.) the weather is an important ND input. Occasionally, certain outputs are set as targets and the DMU cannot "produce" more or less than the specified amounts (e.g., Army recruiting). These ND outputs, although not controlled by the DMU, are critical to the evaluation of its overall efficiency.

The article is organized as follows. It starts by introducing a general formulation for DEA that encompasses many of the DEA variants including the approach suggested by BM. An extension of the BM formulation which allows for the simultaneous existence of ND inputs *and* outputs is presented and discussed. The model is then generalized to the case of partial ND factors, namely, when DMUs have only a limited control over certain factors. A numerical example is provided to illustrate the different treatments, and the discussion section summarizes the article.

2. Preliminaries

Different variants of DEA models can be presented by means of the following mathematical programming formulation:

$$\begin{aligned} &Max\ e_{0,0} \\ &\text{s.t.} \\ &\qquad e_{j,0} \leq 1,\ \forall\, j \\ &\qquad \mu_{r0} \geq \epsilon,\ \forall\, r \\ &\qquad \nu_{i0} \geq \epsilon,\ \forall\, i \end{aligned} \tag{1}$$

where the notation is:

$j = 1, \ldots, n$	the index for the DMUs, 0 used as the index for the analyzed DMU,
$r = 1, \ldots, R$	the index for the outputs, ($y_{rj} \geq 0$ is output r of DMU_j)
$i = 1, \ldots, I$	the index for the inputs, ($x_{ij} \geq 0$ is input i of DMU_j),
$e_{j,0}$	the relative efficiency of DMU_j when DMU_0 is evaluated,
μ_{r0}, ν_{i0}	output and input weights, respectively, associated with the evaluation of DMU_0,
ϵ	a non-Archimedean infinitesimal.

In formulation (1), the input (x_{ij}) and output (y_{rj}) factors are known quantities observed from the activities of the n DMUs and the factor weights (μ_{r0} and ν_{i0}) are the decision variables.

Defining

$$e_{j,0} = \frac{\sum_{r=1}^{R} \mu_{r0} \cdot y_{rj}}{\sum_{i=1}^{I} \nu_{i0} \cdot x_{ij}} \tag{2}$$

we retrieve from (1) the original DEA model developed by Charnes, Cooper and Rhodes [1978] and known as the CCR model. Similarly, defining the relative efficiency measure as,

$$e_{j,0} = \frac{\sum_{r=1}^{R} \mu_{r0} \cdot y_{rj} - u_0}{\sum_{i=1}^{I} \nu_{i0} \cdot x_{ij}} \tag{3}$$

(where u_0 is an unrestricted decision variable) yields the model developed by Banker, Charnes and Cooper [1984] and known as the BCC model.

Banker and Morey [1986] developed a series of models which extend the BCC and CCR models to include ND factors. Starting with ND factors only on the input side, they continued to develop models for ND outputs ("output targets") as well. Letting D represent the set of discretionary inputs and F the set of ND inputs, they defined the efficiency ratio which extends the CCR[1] model as (ibid. p. 517):

$$e_{j,0} = \frac{\sum_{r=1}^{R} \mu_{r0} \cdot y_{rj} - \sum_{i \in F} \nu_{i0} \cdot (x_{ij} - x_{i0})}{\sum_{i \in D} \nu_{i0} \cdot x_{ij}} \tag{4}$$

Applying the transformation of fractional programming problems into linear programs which was used in previous DEA models, and looking at the dual form of the linear program, (4) yields the following formulation (given as (30)–(34) in the BM paper):

$$Min\ \theta_0 - \epsilon \cdot \left[\sum_{r=1}^{R} s_r + \sum_{i \in D} \sigma_i \right] \tag{5}$$

s.t.

$$\sum_{j=1}^{n} \lambda_j \cdot y_{rj} - s_r = y_{r0}, \ \forall\ r$$

$$\sum_{j=1}^{n} \lambda_j \cdot x_{ij} + \sigma_i - \theta_0 \cdot x_{i0} = 0,\ i \in D$$

$$\sum_{j=1}^{n} \lambda_j \cdot x_{ij} + \sigma_i = \sum_{j=1}^{n} \lambda_j \cdot x_{i0},\ i \in F$$

$$\lambda_j,\ \sigma_i,\ s_r \geq 0,\ \forall\ j,\ i,\ r$$

where s_r and σ_i are slack variables on outputs and inputs,[2] respectively, λ_j are coefficients associated with the selection of an efficient frontier point for the evaluation of DMU_0, and θ is an "intensity" variable. In an optimal solution, the model accords θ the value of the largest possible proportional reduction in all inputs.

To better understand formulation (5) above, we offer the following explanation. While it is true that any DMU under evaluation has no control over its ND inputs, other DMUs may have an advantage (or disadvantage) from having at their disposal more (or less) of the same factor. These *differences* in the availability of resources (even though they are uncontrollable) should be reflected in the relative efficiency rating.[3]

Another way to present the same argument is by considering the case when a ND factor is constant across all DMUs. In all probability, such a factor would not enter the analysis. Now, partitioning a ND factor into a constant level—that prevailing in the DMU under evaluation, and differences from that level, only the differences should be considered. When the level is constant, all differences are zero, and this factor will have no effect on the analysis.

The extension of the BCC model to include ND inputs includes an additional convexity constraint on the λ_j values—requiring that their sum will be equal to one (constraint (23) in the BM article).

Finally, we note that (4) is not a unique way of getting at program (5). This non-uniqueness is a direct consequence of the transformation procedure leading from fractional to linear programs. Thus, assigning the following efficiency ratio definition into (1), and executing the same transformation will also yield program (5):

$$e_{j,0} = \frac{\sum_{r=1}^{R} \mu_{r0} \cdot y_{rj}}{\sum_{i \in D} \nu_{i0} \cdot x_{ij} + \sum_{i \in F} \nu_{i0} \cdot (x_{ij} - x_{i0})} \tag{6}$$

We shall find it more convenient to refer to (6) rather than (4) in presenting the extensions to program (5).

3. Extensions

Charnes and Cooper [1985], ibid p. 88, pointed out two directions for further research in the area of ND factors:

"We shall not pursue these topics further except to note that extensions to partially controllable inputs and outputs require the introduction of additional constraints while the question of simultaneous treatment of non-discretionary inputs and outputs remains open."

3.1. Simultaneity

Most of the ND cases treated in the literature so far, both in theory and in application, considered ND factors only on the input side. In the few places where the existence of ND outputs was recognized (e.g., Banker and Morey [1986] and Dyson and Thanassoulis [1991]), separate models were formulated to handle the ND inputs and the ND outputs. A straightforward extension of the BM approach, given in (5), into simultaneous ND inputs and outputs can rely on the efficiency definition in (6) where the inputs and outputs appear separately in the two sides of the efficiency ratio.

Accordingly, a generalized efficiency ratio definition for cases where ND factors are found on both the input and the output sides is:

$$e_{j,0} = \frac{\sum_{r \in D_y} \mu_{r0} \cdot y_{rj} + \sum_{r \in F_y} \mu_{r0} \cdot (y_{rj} - y_{r0})}{\sum_{i \in D_x} \nu_{i0} \cdot x_{ij} + \sum_{i \in F_x} \nu_{i0} \cdot (x_{ij} - x_{i0})} \tag{7}$$

Where D_x, F_x and D_y, F_y are the sets of discretionary and ND factors, respectively, for the inputs and outputs.

Obviously, only discretionary factors affect the efficiency of the analyzed DMU (where $j = 0$). Efficiency ratios for other DMUs ($j \neq 0$) take into account, in the dimensions of the ND factors, only deviations from the levels of the same factors in DMU_0.

The linear programming formulation of the DEA model corresponding to this case is:

$$Min\ \theta_0 - \epsilon \cdot \left[\sum_{r \in D_y} s_r + \sum_{i \in D_x} \sigma_i \right] \tag{8}$$

s.t.

$$\sum_{j=1}^{n} \lambda_j \cdot y_{rj} - s_r = y_{r0},\ r \in D_y$$

$$\sum_{j=1}^{n} \lambda_j \cdot y_{rj} - s_r = \sum_{j=1}^{n} \lambda_j \cdot y_{r0},\ r \in F_y$$

$$\sum_{j=1}^{n} \lambda_j \cdot x_{ij} + \sigma_i - \theta_0 \cdot x_{i0} = 0,\ i \in D_x$$

$$\sum_{j=1}^{n} \lambda_j \cdot x_{ij} + \sigma_i = \sum_{j=1}^{n} \lambda_j \cdot x_{i0},\ i \in F_x$$

$$\lambda_j,\ \sigma_i,\ s_r \geq 0,\ \forall\ j,\ i,\ r$$

3.2. *Partial Discretion—Relative Approach*

In all the DEA literature dealing with ND factors, the topic is always presented in a "0–1" manner, i.e., a factor is either fully controllable or totally uncontrollable. However, in many real-life applications, one finds that a different concept prevails among managers of DMUs. Often, they prefer to state that they can exercise only a limited control over some variables. For example, managers can make marginal alterations in personnel scheduling (they might, for example, have the right to authorize a limited amount of overtime) but, they have to follow general guidelines of their organization in many other respects involving the use of their human resources. Looking at many of the reported applications of DEA, one will easily find cases where either inputs or outputs were treated as discretionary factors while, in fact, they should have been treated as partially controlled factors.

To formulate these cases in a way which will generalize the previous developments, we let $0 \leq \delta_r \leq 1$, and $0 \leq \delta_i \leq 1$ represent the degree of discretion that the DMU has on output r and input i, respectively. $\delta = 1$ (for any index) means that the factor is fully discretionary and $\delta = 0$ means that the factor is totally ND. Any value between 0 and 1 implies partial control over the factor.

One can now view the amount of a given factor for any DMU as consisting of two parts—the discretionary and ND portions. For example, if the factor in question is an input i^*, such as total labor hours available, and there is, say, the possibility of altering the base level by 15% (i.e., $1 - \delta_{i^*} = 0.85$). Then, 85% of the input is ND while 15% is discretionary. Thus, for example, the expression in (6) would appear as

$$e_{j,0} = \frac{\sum_r \mu_{r0} \cdot y_{rj}}{\sum_{i \in D_x} \nu_{i0} \cdot x_{ij} + \nu_{i^*0} \cdot \delta_{i^*} \cdot x_{i^*j} + \nu_{i^*0} \cdot [(1 - \delta_{i^*})x_{i^*j} - (1 - \delta_{i^*})x_{i^*0}]} \tag{9}$$

$$= \frac{\sum_r \mu_{r0} \cdot y_{rj}}{\sum_{i \in D_x} \nu_{i0} \cdot x_{ij} + \nu_{i^*0} \cdot [x_{i^*j} - (1 - \delta_{i^*})x_{i^*0}]}$$

This definition yields the following extension to the CCR fractional programming model for the efficiency evaluation of DMU_0:

$$Max \frac{\sum_{r=1}^{R} \mu_{r0} \cdot \delta_r \cdot y_{r0}}{\sum_{i=1}^{I} \nu_{i0} \cdot \delta_i \cdot x_{i0}} \tag{10}$$

s.t.

$$\frac{\sum_{r=1}^{R} \mu_{r0} \cdot (y_{rj} - (1 - \delta_r) \cdot y_{ro})}{\sum_{i=1}^{I} \nu_{i0} \cdot (x_{ij} - (1 - \delta_i) \cdot x_{io})} \leq 1, \; \forall j$$

$$\mu_{r0}, \nu_{i0} \geq \epsilon, \; \forall \, i, r$$

Clearly, when all δ_r, $\delta_i = 1$, formulation (10) is equivalent to the original CCR ratio model. Also, when the δ_r, δ_i parameters assume only zero or one values, (10) corresponds to the efficiency definition in (7). Finally, note that the ratios in the objective and the constraints are identical (since in the objective we have $1-(1-\delta_i) = \delta_i$ and $1-(1-\delta_r) = \delta_r$).

Applying once again the transformation of fractional to linear programs we obtain the (dual) linear representation of (10):

$$Min\ \theta_0 - \epsilon \cdot \left(\sum_{r=1}^{R} s_r + \sum_{i=1}^{I} \sigma_i \right) \tag{11}$$

s.t.

$$\sum_{j=1}^{n} \lambda_j \cdot y_{rj} - s_r = y_{r0} \cdot \left[\delta_r + (1-\delta_r) \cdot \sum_{j=1}^{n} \lambda_j \right], \ \forall\ r$$

$$\sum_{j=1}^{n} \lambda_j \cdot x_{ij} + \sigma_i = x_{i0} \cdot \left[\theta_0 \cdot \delta_i + (1-\delta_i) \cdot \sum_{j=1}^{n} \lambda_j \right], \ \forall\ i$$

$$\lambda_j,\ \sigma_i,\ s_r \geq 0,\ \forall\ j,\ i,\ r$$

It is interesting to note that when the convexity constraint associated with the BCC model (sum of λs equal one) is added to (11), the δ_r parameters disappear from the output constraints. Thus, the BCC version of (11) will not distinguish between outputs according to their discretion levels.

The generalized DEA model given in (11) benefits from an important advantage over previous models designed specifically to handle ND factors. Here, by applying a simple transformation,

$$x_{ij}^o = x_{ij} - (1-\delta_i) \cdot x_{i0}, \quad y_{rj}^o = y_{rj} - (1-\delta_r) \cdot y_{r0} \tag{12}$$

and assigning the (transformed) x_{ij}^o and y_{rj}^o in (10) and (11) we get back the original CCR model. This implies that as long as we maintain some simple conditions (e.g., not all $x_{ij}^o = 0$), the entire theoretical structure erected on the foundations of the CCR model holds. This structure includes theorems on the existence of feasible solutions (and consequently, the assurance that in any solution $\theta \leq 1$ always holds), the characteristics of efficient units (all zero slacks and $\theta = 1$), etc. The transformation also carries with it an important computational advantage, namely, by a simple data preprocessing technique one is able to continue using existing DEA software packages. None of these advantages are readily available with the other ND models (including (8)) since the formulations there include additional constraints which do not appear in the original model.

Transformation (12) paves the way to obtaining a "DEA projection formula" for model (11), that is, the changes required in the input-output values of DMU_0 in order for it to be located on the DEA "efficiency frontier."[4] First, the projected point in the transformed (x^o, y^o) space is:

$$x_{i0}^{o*} = \sum_{j=1}^{n} \lambda_j^* \cdot x_{ij}^o; \qquad y_{r0}^{o*} = \sum_{j=1}^{n} \lambda_j^* \cdot y_{rj}^o \tag{13a}$$

where the asterisks represent optimal solution values in (11). Now, applying transformation (12) backwards by adding to both sides of the equations in (13a) the terms $(1 - \delta_i) \cdot x_{i0}$ and $(1 - \delta_r) \cdot y_{r0}$, respectively, we get back from the (x^o, y^o) space to the (x, y). Hence:

$$x_{i0}^* = \sum_{j=1}^{n} \lambda_j^* \cdot x_{ij} + \left(1 - \sum_{j=1}^{n} \lambda_j^*\right) \cdot (1 - \delta_i) \cdot x_{i0}$$

$$y_{i0}^* = \sum_{j=1}^{n} \lambda_j^* \cdot y_{rj} + \left(1 - \sum_{j=1}^{n} \lambda_j^*\right) \cdot (1 - \delta_r) \cdot y_{r0} \tag{13b}$$

3.3. *Partial Discretion—Absolute Approach*

An alternative formulation to describe partial control over input and/or output factors can be effected by imposing constraints on the input-output slacks. This approach may become useful when decision makers prefer to express their degree of discretion by *absolute* terms, i.e., by providing ranges for possible changes in either the inputs or the outputs, rather than by *relative* terms as demonstrated through the δ parameters. Because of obvious difficulties that may arise from introducing the bounds directly on input slacks (since θ will not be effected by these bounds), we prefer to place restrictions directly on the potential changes in input or output values. A somewhat similar approach was adopted by Ray [1991] who imposed upper bounds on projected output values. In general, a CCR model with bounds on potential changes in input or output values will be given as:

$$Min\ \theta_0 - \epsilon \cdot \left(\sum_{r=1}^{R} s_r + \sum_{i=1}^{I} \sigma_i\right) \tag{14}$$

s.t.

$$U_{r0} \geq \sum_{j=1}^{n} \lambda_j \cdot y_{rj} = y_{r0} + s_r, \ \forall\ r$$

$$L_{i0} \leq \sum_{j=1}^{n} \lambda_j \cdot x_{ij} = \theta_0 \cdot x_{i0} - \sigma_i, \ \forall\ i$$

$$\lambda_j,\ s_r,\ \sigma_i \geq 0, \ \forall\ j, r, i$$

where U_{r0} and L_{i0} are upper bounds on output values and lower bounds on input values, respectively.

Clearly, by setting all output bounds to infinity and input lower bounds to zero we retrieve from (14) the original CCR model. As U_{r0} (L_{i0}) is set closer to $y_{r0}(x_{i0})$ the less discretion the DMU has over these factors. There are three possible disadvantages of (14) when compared to (11). First, it requires much more (subjective) data to be entered into the analysis (one set of upper and lower bounds for each DMU). In some situations it may be difficult to specify values for these bounds. Second, it again makes changes in the model itself, thus inhibiting the use of existing DEA theory and software. Third, setting the bounds too close to the observed (x_0, y_0) may limit the reference set to a greater extent than is desirable.

4. Numerical Example

This section presents a simple numerical example which illustrates the way relative efficiency ratings and projections are affected by the generalized ND model (11). An artificial set of data, containing 15 DMUs with two outputs and three inputs, was generated and each of the input and output factors was associated with a given degree of discretion. Raw data are presented on the left of Table 1 and the δ (discretion) parameters are given at the bottom row of that table.

The projections of inefficient DMUs to the frontier, as affected via formula (13), are presented in Table 2. Notice, for example, DMU_8 whose efficiency score is extremely low. It is evaluated with respect to a combination of DMUs 5 and 7. Both of these efficient units have a value $x_1 = 2$ compared with $x_{18} = 9$. As $(1 - 0.8) = 0.2$ (or %20) of the latter is deducted from the former value, the corresponding input constraint allows for a very small θ value. Thus, rather than draw the interpretations from the values of θ, we recommend observing the projected values. There we find, for the same DMU, that its projected inputs are indeed much smaller than their current values, but, this reduction does not go below %22 (in the first input).

Table 1. Efficiency results using model (11).

DMU	y1	y2	x1	x2	x3	Efficiency	Facet
1	5	15	2	15	17	0.937	$\lambda_7 = 0.938$
2	5	12	4	14	24	0.280	$\lambda_7 = 0.75$
3	8	11	5	9	16	0.268	$\lambda_5 = 0.019$, $\lambda_7 = 0.981$
4	6	11	2	14	24	0.705	$\lambda_5 = 0.274$, $\lambda_7 = 0.430$
5	9	15	2	10	27	1.00	
6	8	16	6	12	6	1.00	
7	8	16	2	3	15	1.00	
8	9	15	9	13	20	0.026	$\lambda_5 = 0.091$, $\lambda_7 = 1.089$
9	9	11	7	9	8	1.00	
10	3	14	7	1	12	1.00	
11	6	10	3	2	25	1.00	
12	7	13	8	4	18	0.435	$\lambda_7 = 0.482$, $\lambda_{10} = 0.391$, $\lambda_{11} = 0.341$
13	4	14	6	6	9	1.00	
14	10	15	4	11	13	1.00	
15	4	11	9	5	21	0.199	$\lambda_7 = 0.462$, $\lambda_{10} = 0.257$
δ	0.95	1	0.8	0.75	0.1		

Table 2. Projections onto the efficiency frontier using formula (13).

DMU	Proj(y1)	Proj(y2)	Proj(x1)	Proj(x2)	Proj(x3)
1	7.52	15.01	1.90	3.05	15.02
2	6.06	12.00	1.70	3.12	16.65
3	8.00	16.00	2.08	3.17	14.83
4	6.00	11.00	1.53	5.07	20.25
5	9.00	15.00	2.00	10.00	27.00
6	8.00	16.00	6.00	12.00	6.00
7	8.00	16.00	2.00	3.00	15.00
8	9.45	18.79	2.04	3.60	15.55
9	9.00	11.00	7.00	9.00	8.00
10	3.00	14.00	7.00	1.00	12.00
11	6.00	10.00	3.00	2.00	25.00
12	7.00	16.60	4.38	2.31	16.98
13	4.00	14.00	6.00	6.00	9.00
14	10.00	15.00	4.00	11.00	13.00
15	4.52	11.00	3.23	1.99	15.32

The third input is characterized by a small degree of discretion ($\delta = 0.1$). Such a factor would have traditionally been considered as a ND factor and treated through model (5) (or model (8) if it coincided with another ND output). Table 3 facilitates comparison between the models as it provides the efficiency scores and resulting projections of inefficient DMUs onto the efficiency frontier constructed by model (5).

Table 3. Optimal slacks and efficient projections using model (5).

			Slacks					Projections				
DMU	Eff. (5)	Facet	y1	y2	x1	x2	x3	y1	y2	x1	x2	x3
1	1.000		0	0	0	0	0	5	15	2	15	17
2	0.375	$\lambda_7 = 0.75$	1	0	0	3	6.75	6	12	1.5	2.25	2.25
3	0.396	$\lambda_5 = 0.082$ $\lambda_7 = 0.907$	0	4.75	0	0.02	0	8	15.75	1.98	3.55	6.33
4	0.708	$\lambda_5 = 0.333$ $\lambda_7 = 0.375$	0	0	0	5.46	2.37	6	11	1.42	4.46	14.62
5	1.000		0	0	0	0	0	9	15	2	10	27
6	1.000		0	0	0	0	0	8	16	6	12	6
7	1.000		0	0	0	0	0	8	16	2	3	15
8	0.259	$\lambda_{11} = 0.052$ $\lambda_7 = 1.086$	0	2.9	0	0	5.17	9	17.9	2.33	3.36	0
9	1.000		0	0	0	0	0	9	11	7	9	8
10	1.000		0	0	0	0	0	3	14	7	1	12
11	1.000		0	0	0	0	0	6	10	3	2	25
12	0.614	$\lambda_7 = 0.363$ $\lambda_{10} = 0.388$ $\lambda_{11} = 0.488$	0	3.13	0	0	0	7	16.13	4.91	2.45	11.04
13	1.000		0	0	0	0	0	4	14	6	6	9
14	1.000		0	0	0	0	0	10	15	4	11	13
15	0.320	$\lambda_7 = 0.437$ $\lambda_{10} = 0.286$	0.36	0	0	0	5.2	4.36	11	2.88	1.6	1.51

In Table 3, we point out that efficiency scores in model (5) will always be larger than or equal to the corresponding scores in model (11) for the same data set. This is due to the fact that (11) is a relaxation of (5). Also, the selection of the comparison set (the "facet") was almost identical in the two models (except for some changes in λ values for the DMUs in the facet).

It is interesting to notice again DMU_8, whose projected value in the third (ND) input is reduced to zero. As this factor does not affect the value of θ, and its slack is free to absorb any difference between the evaluated DMU and its comparison set, the model's recommendations in this dimension are to be ignored. This issue of seemingly unlimited positive slacks in the ND dimensions is addressed again in the final section of this article.

5. Discussion

Non-discretionary factors in relative efficiency analyses are such factors that—although uncontrollable by DMUs—still impact relative efficiency ratings. In this article we have extended and generalized the approach taken by Banker and Morey [1986] to treat ND factors in two ways. First, the study of cases where both input and output ND factors appear simultaneously is made possible. Second, we have relaxed the "zero-one" distinction between discretionary and ND factors. Motivated by numerous real-life situations where partial control over certain factors exist, we argue that a model allowing varying degrees of control (0–100%) over factors may be more suitable to many cases.

The extension to simultaneous ND inputs and outputs, following the approach proposed by BM, modifies the DEA *model* so as to account for the fact that a DMU does not control such factors. The generalization leading to partially controlled factors can be viewed as one which introduces appropriate changes in the relevant *data*, rather than the model itself. Thus, important methodological and computational advantages are secured.

Finally, we point the reader's attention to several important issues concerning the ND factors. First, just as there is no accepted quantitative technique to identify and determine the appropriate number of relevant factors in ordinary DEA applications and instead one uses a variety of ad-hoc procedures (see Golany and Roll [1989]), there is not also an accepted method to select the ND factors. In this respect, one should be cautioned against adding a disproportionate number of ND factors into the analysis which may bring about unduly high efficiency ratings (see Ray [1988]). Further, in many situations it may be difficult to specify the values for the δ parameters which describe the degree of discretion over factors. It might be helpful in such cases to perform a sensitivity analysis by allowing the δ values to vary. The efficiency rating, as well as the other measures derived from DEA, will then be contingent upon the values assumed for the δ parameters.

Second, by the approach taken here, as well as by the BM approach, some ND factors, in some DMUs, may assume negative values (as a result of the differences in both the numerator and denominator of the efficiency ratio). This may happen when DMUs are characterized by ND input (output) levels which are lower (larger) than the respective factors at the DMU in the objective. Regular LP models, and DEA codes based on them, can handle these cases without difficulty.

Third, in this article (as in the BM article) we continue to assume that ND inputs (outputs) can be underutilized (overproduced). This allows the slack variables in the restriction for the fixed inputs and outputs to assume positive values. Ray [1991] argues that the non-discretionary inputs may be unalterable, and uses that as a motivation to include them in a post-DEA regression analysis. This goes beyond the scope of this article and we intend to address it in further research.

In Memoriam

This article is dedicated by the first author to the memory of his colleague and co-author, Professor Yaakov Roll (1927–1993).

Acknowledgments

The authors would like to thank Prof. W. Cook, the associate editor, for his insightful comments and suggestions. His recommendations contribute substantially to the presentation of our research.

Notes

1. We prefer to focus on BM's extension to the CCR model rather than the BCC model because of the recent criticism by Chang and Guh [1991] of the BCC model with respect to the issue of returns to scale (which is at the heart of the BCC model).
2. Isotonicity in input-output relations (see Charnes and Cooper [1985, p. 102] is assumed here, as in the BM article. This implies that input factors (in particular, ND inputs) are such that more of the factor facilitates more (or equal) outputs. When this is not the case (e.g., level of competition, case mix severity index in hospitals, etc.) we need to reverse the definition of the factor by taking its reciprocal or a complement of its value from an upper bound (see Charnes et al. [1985, p. 99]).
3. For the use of differences between factor values, in a different context of DEA, see Golany and Rousseau [1992].
4. These terms are explained in Charnes and Cooper [1985].

References

Banker, R.D., A. Charnes, and W.W. Cooper. (1984). "Some Models for Estimating Technical and Scale Inefficiencies in Data Envelopment Analysis." *Management Science*, 27, 1370–1382.

Banker, R.D. and R.C. Morey. (1986). "Efficiency Analysis for Exogenously Fixed Inputs and Outputs." *Operations Research* 34(4), 513-521.

Chang, K.P. and Y.Y. Guh. (1991). "Linear Production Functions and The Data Envelopment Analysis." *European Journal of Operations Research* 52, 215-223.

Charnes, A., C.T. Clark, W.W. Cooper and B. Golany. (1985). "A Developmental Study of Data Envelopment Analysis in Measuring the Efficiency of Maintenance Units in the U.S. Air Force." *Annals of Operations Research* 2, 95–112.

Charnes, A. and W.W. Cooper. (1985). "Preface to Topics in Data Envelopment Analysis." *Annals of Operations Research* 2, 59–94.

Charnes, A., W.W. Cooper, B. Golay, L. Seiford and J. Stutz (1985). "Foundations of Data Envelopment Analysis for Pareto-Koopmans Efficient Empirical Production Functions." *Journal of Econometrics* 30, 91–107.

Charnes, A., W.W. Cooper, A. Lewin, R.C. Morey and J. Rousseau. (1980). "Efficiency Analysis with Non-Discretionary Resources." Research Report CCS 379, The University of Texas at Austin, Center for Cybernetic Studies, Austin, Texas, U.S.A.

Charnes, A., W.W. Cooper and E. Rhodes. (1978). "Measuring Efficiency of Decision Making Units." *European Journal of Operations Research* 2(6), 429–444.

Golany, B. and Y. Roll. (1989). "An Application Procedure for DEA," *OMEGA* 17(3), 237–250.

Golany, B. and J.J. Rousseau. (1992). "Efficiency Evaluation Games," In *Systems Management Science by Extreme Methods*. J.J. Rousseau and F.Y. Phillips (eds.). pp. 327–347, Boston: Kluwer Academic Publishers.

Ray, S. (1988). "Data Envelopment Analysis, Nondiscretionary Inputs and Efficiency: An Alternative Interpretation." *Socio-Economic Planning Sciences* 22(4), 167–176.

Ray, S. (1991). "Resource-Use Efficiency in Public Schools: A Study of Connecticut Data." *Management Science* 37, 1620–1628.

Seiford, L. (1990). "A Bibliography of Data Envelopment Analysis." (1978-1990). Working Paper, the University of Massachusetts, Amherst, U.S.A.

Seiford, L. and R. Thrall. (1990). "Recent Developments in DEA: The Mathematical Approach to Frontier Analysis." *Journal of Econometrics* 46, 7–38.

Thanassoulis, E. and R.G. Dyson. (1992). "Estimating Preferred Target Input Output Levels Using Data Envelopment Analysis." *European Journal of Operations Research* 56(1), 80–97.