Journal of Econometrics 46 (1990) 73–91. North-Holland

POLYHEDRAL CONE-RATIO DEA MODELS WITH AN ILLUSTRATIVE APPLICATION TO LARGE COMMERCIAL BANKS*

A. CHARNES, W.W. COOPER, and Z.M. HUANG

University of Texas, Austin, TX 78712, USA

D.B. SUN

University of Maine, Orono, ME 04473, USA

Polyhedral Cone-Ratio Data Envelopment Analysis Models generalizing the CCR Ratio Model are developed for situations with a finite number of DMUs and employing polyhedral cones of virtual multipliers. They provide improved definitions of efficiency over CCR models whose input–output data and/or numbers of DMUs are inadequate to capture aspects or restrictions which should be involved. The focus here is on the sum form for cones which easily provides for capturing exogenous expert opinion as well as mathematical reduction to the old form with its very powerful software. Transformation from the usual intersection form to it and vice versa is explicitly given. Thereby the advantages of either or both are available. The theory is illustrated with two-dimensional examples and by real banking examples for motivation.

1. Introduction

Data Envelopment Analysis (DEA) as brought forward in this journal by Charnes, Cooper, Golany, Seiford, and Stutz (1985) encompasses development of empirical economic production functions on the one hand, and managerial performance evaluation (efficiency analysis) on the other hand from selected or available input and output observed (or sampled) data. As shown in this 1985 'Foundation' paper, *all* the different DEA models correspond mathematically to the Charnes–Cooper test in the Charnes–Cooper monograph (1961) for vector (here Pareto–Koopmans) optimality (or 'dominance' in current mathematical terminology) in goal (or multi-objective) programming. In DEA the selections of inputs and outputs, in contrast to usual microeconomic or operations research normative individual firm developments, are in terms of more aggregative quantities such as, for example, may be available from summarized accounting. Thus, the construction of a good approximate empirical production function may be flawed by the absence of important input or output variables or by the presence of unsuitable ones. Also, depending on the DEA model chosen, the production

*This research was partly supported by National Science Foundation Grant SES-8520806 with the Center for Cybernetic Studies, University of Texas at Austin. It was also partly supported by the IC2 Institute at The University of Texas. Reproduction in whole or in part is permitted for any purpose of the United States Government.

0304-4076/90/$3.50

possibility set may bê a bad description of what is actually possible [see Seiford and Thrall (this issue)]. Further, there may be sufficient observations to construct an adequate production possibility set with *any* model or to robustly determine a good empirical economic production function. Re classical econometrics, see Charnes, Cooper, Seiford, and Stutz (1983).

Because of the mathematical duality relationships between the production function side and the managerial performance or efficiency 'valuational' side, additional relevant information may be brought to bear on construction of a more adequate analysis, e.g., DEA model, either by direct modification of the production possibility set or by restricting the ranges of relative valuation of inputs or outputs on the performance valuational side. Of course, combinations of both can be done. A DEA analysis is *not* achieved until the final selection of inputs, outputs and *model* and its application, via perhaps including window analysis, envelopment maps, etc., are performed.

In this paper, we restrict ourselves to valuational side modifications. In Charnes, Cooper, Huang, and Wei (C^2HW) (1986, 1989) the CCR model was extended to include more general cones than the nonnegative orthants for the virtual multipliers and also obtained were their corresponding dual 'envelopment' problems. Here we consider only the situations of a finite number of DMUs and polyhedral cones, i.e., cones which may be represented *equivalently* either as intersection of a finite number of half spaces or as cones spanned by a finite number of vectors.

The Supercollider location case of Thompson and Thrall (1986) which involved six DMUs (possible locations) rated five of the six as efficient with the CCR model. Additional conditions on relative valuation of inputs and outputs which required ratios of pairs of multipliers to lie within certain ranges called 'assurance regions' were developed. These conditions were equivalent to special cases of intersections of additional half-spaces restricting the virtual multipliers to cones smaller than the nonnegative orthants. Desired here was *not* an empirical production function but a choice of one of the DMUs. With these additional conditions on the multipliers, i.e., with this special cone-ratio model, as desired, only one DMU, Waxahachie, was efficient. The corresponding empirical economic production function, which was *not* a desired objective, consists simply of all positive multiples of the Waxahachie input–output vector.

In contrast, as developed in the Ph.D. thesis of D.B. Sun (1987), a DEA model to objectively assess managerial performance of bank managers was desired, responding to Professor Joseph Ewers' unhappiness with the Federal Reserve's firing of certain bank managers whose managerial performance Ewers assessed as being excellent. (Ewers is a nationally recognized expert on Savings and Loan Institutions who was for twenty years President of the Federal Home Loan Board.) Inputs and outputs were chosen from the FDIC Call Reports. Yet with four inputs, four outputs, and 47 DMUs (banks) the CCR model recognized a few notoriously inefficient banks as efficient.

Evidently the model with Call Reports data inputs and outputs was not adequate to represent the valuational processes of bank experts. Further, no obvious translations of the implicit expert to explicit relative valuations of inputs and outputs was apparent. A new method of bringing such expert overall valuational knowledge to bear through developing a better DEA model was generated.

This method generated the cone of a cone-ratio model as the cone spanned by the optimal virtual multipliers in the CCR model of three banks unanimously recognized by bank experts as being preeminently efficient. Thus, here, the sum form for the cones in the cone-ratio model, which has other important computational advantages, was natural, was immediately available, and was used as shown in a later section of this paper.

Why only three expert efficient choices? The DMU sample was that of the largest U.S. banks whose operations were not geographically restricted nor were there substantial differences in the range of banking operations (loans, etc.) performed. Three, unanimously agreed to be efficient, were sufficient to provide for a reasonable range of flexibility in relative valuations of inputs and outputs.

We next develop the models or relations of the two types in restricted mathematical generality but suitable enough for many applications. More general theory and models are to be found in C^2WH (1986, 1989). We illustrate the results by means of small examples.

2. The cone-ratio DEA model

In order to generalize to infinitely many DMUs and more general conditions that may impose restrictions on the dual evaluators of outputs and inputs, the CCR model as in Charnes, Cooper, and Rhodes (1978), was generalized by Charnes, Cooper, Wei, and Huang (1986) to the 'Cone-Ratio CCR Model', here presented only for the finite number of DMUs case. These problems are typically:

$$V_p = \text{Max } \mu^T \overline{Y}_0, \tag{1}$$

$$\text{s.t.} \quad \omega^T \overline{X} + \mu^T \overline{Y} \leq 0,$$

$$\omega^T \overline{X}_0 = 1,$$

$$\omega \in V, \qquad \mu \in U,$$

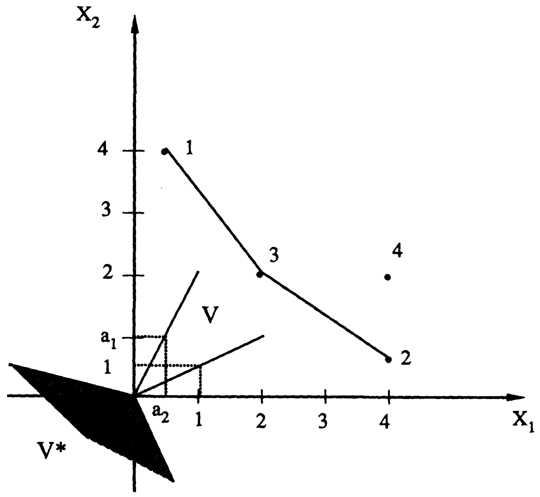

Fig. 1

and its dual (in the DEA form) [see Ben-Israel, Charnes, and Kortanek (1969)]

$$V_D = \text{Min}\,\theta, \tag{2}$$

$$\text{s.t.}\quad \bar{X}\lambda + \theta\bar{X}_0 \in -V^*,$$

$$\bar{Y}\lambda - \bar{Y}_0 \in -U^*,$$

$$\lambda \geq 0,$$

where $\bar{X}$ is the $m \times n$ input matrix for the n DMUs to be considered and $\bar{Y}$ is the $s \times n$ matrix of their outputs. $V \subseteq E^m_+$, $U \subseteq E^s_+$ are closed convex cones and V^* and U^* are the negative polar cones of U and V, respectively. As shown in fig. 1, the spanning vectors of a polar V^* are the directions normal to the hyperplanes bounding V when V is polyhedral.

We use $\bar{X}_j$ and $\bar{Y}_j$, respectively, to represent the input vector and output vector of the jth DMU and assume that $\bar{X}_j \in \text{Int}(-V^*)$, $\bar{Y}_j \in \text{Int}(-U^*)$ for any j. $\text{Int}(-V^*) = \{v\text{: } v'v > 0, \text{ for all } v' \in V \text{ and } v' \neq 0\}$. $\text{Int}(-U^*) = \{u\text{: } u'u > 0, \text{ for all } u' \in U \text{ and } u' \neq 0\}$. $\text{Int}(V^*)$ and $\text{Int}(U^*)$ are not empty since V and U are 'acute' cones as defined via $V \subseteq E^m_+$, $U \subseteq E^s_+$. See Yu (1974). Acute cones are those properly contained in a half-space.

Both (1) and (2) have optimal solutions. With suitable regularity conditions, the optimal solution values are equal, $V_D = V_P = \widehat{\mu^T}\bar{Y}_0 \leq \widehat{\omega^T}\bar{X}_0 = 1$.

Definition 1. DMU_o is said to be efficient if there exists an optimal solution $(\hat{\mu}, \hat{\omega})$ of (1) such that

$$\hat{\mu}^T \bar{Y}_0 = 1 \quad \text{and} \quad \hat{\mu} \in \operatorname{Int} U, \quad \hat{\omega} \in \operatorname{Int} V.$$

The cone-ratio CCR model thus extends the CCR model by employing closed convex cones U and V which need not be nonnegative orthants. If we set $V = E_+^m$, $U = E_+^s$, then the two models coincide. See Charnes, Cooper, Wei, and Huang (1986, 1989) for further discussion.

Polyhedral cones V and U and the cone-ratio CCR model

As mentioned, with only a finite number of inputs, outputs, and DMUs, it may suffice to employ only polyhedral cones V and U to achieve desirable variants of past DEA efficiency evaluations. Polyhedral convex cones V and U may be expressed in the sum form as

$$\begin{aligned} V &= \{A^T\alpha : \alpha \geq 0\}, \quad \alpha \in E_+^l, \quad A^T = (a^1, a^2, \ldots, a^l), \\ &\qquad a^i \in E_+^m, \quad i = 1, \ldots, l, \\ U &= \{B^T\gamma : \gamma \geq 0\}, \quad \gamma \in E_+^k, \quad B^T = (b^1, b^2, \ldots, b^k), \\ &\qquad b^r \in E_+^s, \quad r = 1, \ldots, k, \end{aligned} \tag{3}$$

and

$$V^* = \{v : Av \leq 0\} \quad \text{and} \quad U^* = \{u : Bu \leq 0\}.$$

Construction of a polyhedral convex cone V may be illustrated by the following example. Suppose the DMUs have two inputs. In the CCR model, the ratio of their marginal substitution rate is $0 < \widehat{\omega_1}/\widehat{\omega_2} < \infty$, where $\widehat{\ }$ means optimal. Now suppose market information sets the range of this ratio as $c_1 \leq \widehat{\omega_2}/\widehat{\omega_1} \leq c_2$, with $c_2 \geq c_1 > 0$. This can be rewritten as

$$-\widehat{\omega_2} + c_2\widehat{\omega_1} \geq 0, \qquad \widehat{\omega_2} - c_1\widehat{\omega_1} \geq 0. \tag{4}$$

Thereby, $\hat{\omega} \in V = \{\omega : C\omega \geq 0,\ \omega \geq 0\}$, where

$$C = \begin{bmatrix} c_2 & -1 \\ -c_1 & 1 \end{bmatrix}, \qquad \omega = \begin{bmatrix} \omega_1 \\ \omega_2 \end{bmatrix}.$$

V may also be defined equivalently as $\omega \in V = \{A^T\alpha : \alpha \geq 0\}$, where

$$A = \begin{bmatrix} 1 & c_1 \\ 1 & c_2 \end{bmatrix}, \qquad \alpha = \begin{bmatrix} \alpha_1 \\ \alpha_2 \end{bmatrix}.$$

Then $-V^* = \{v : Av \geq 0\}$.

More generally, an 'assurance' region given by (without loss of generality)

$$a_{ij} \leq w_i/w_j \leq b_{ij}, \qquad i<j, \quad i,j=1,\dots,n,$$

satisfying $b_{ij} \geq a_{ij} \geq 0$ with at least one $b_{ik} > a_{ik}$ and $b_{ij} > 0$ can be written as

$$-w_i + b_{ij}w_j \geq 0, \qquad w_i - a_{ij}w_j \geq 0,$$

with $i<j$ and $i,j=1,\dots,n$.

The corresponding constraint cone with matrix C from this can be written in intersection form as $\{w\colon Cw \geq 0, w \geq 0\}$. Here C is full column rank. This can be rewritten in sum form as $\{w\colon w = A^T\alpha, \alpha \geq 0\}$ where $A^T = (C^TC)^{-1}C^T$.[1] Conversely, if we have the constraint cone in sum form with $A \geq 0$, $\{w\colon w = A^T\alpha, \alpha \geq 0\}$ where A^T is of full column rank, an intersection form of $\{w\colon Cw \geq 0, w \geq 0\}$ can be had from $C = A^T(AA^T)^{-1}$.

In general, going from intersection to sum form, the columns of an A^T can be the extreme points of $\{w\colon Cw \geq 0, \varepsilon^T w = 1, w \geq 0\}$, which may involve large numbers of extreme point determinations for simple C matrices. But if C is of full row rank, $0 \leq A = C^T(CC^T)^{-1}$ is available.

Without intersection to sum transfer, the cone defining inequalities beyond nonnegativity corresponds to modifying (extending) the production possibility set of the dual linear programming problem (the DEA side).

As will later be seen, polyhedral cones can tighten efficiency criteria in DEA tests. Before giving examples, we further expose some theoretical underpinnings as follows.

With U and V as polyhedral convex cones represented in the sum form, using eq. (3), problems (1) and (2) can be transformed into

$$\begin{aligned} V_{P_1} = {} & \operatorname{Max} \gamma^T(B\bar{Y}_0), \\ \text{s.t.} \quad & \alpha^T(A\bar{X}) + \gamma^T(B\bar{Y}) \leq 0, \\ & \alpha^T(A\bar{X}_0) = 1, \\ & \alpha \geq 0, \quad \gamma \geq 0, \qquad \alpha \in E^l, \quad \gamma \in E^k, \end{aligned} \tag{5}$$

$$\begin{aligned} V_{D_1} = {} & \operatorname{Min} \theta, \\ \text{s.t.} \quad & (A\bar{X})\lambda + \theta(A\bar{X}_0) \geq 0, \\ & (B\bar{Y})\lambda - (B\bar{Y}_0) \geq 0, \\ & \lambda \geq 0. \end{aligned} \tag{6}$$

[1]See the publicly available 1989 Charnes, Cooper, Huang, and Sun CCS 636 report 'Relations Between Half-Space and Finitely Generated Cones in Polyhedral Cone-Ratio DEA Models'.

Letting $\bar{X}' = A\bar{X}$, $\bar{Y}' = B\bar{Y}$, the cone-ratio CCR model then coincides with a CCR model evaluating the same DMUs but with the transformed data $\bar{X}'$ and $\bar{Y}'$. Note that $\bar{X}'$ and $\bar{Y}'$ are strictly positive, since $a^{iT} \in V$, $b^{rT} \in U$, and $\bar{X}_j \in \text{Int}(-V^*)$, $\bar{Y}_j \in \text{Int}(-U^*)$, $j = 1, \ldots, n$, $i = 1, \ldots, l$; $r = 1, \ldots, k$.

The following theorem establishes the existence of efficient DMUs for the cone-ratio CCR DEA model [see C^2WH (1986) for proof].

Theorem 1. There exists at least one efficient DMU with the cone-ratio CCR model provided that U and V are polyhedral cones.

Since problem (6) and problem (2) are equivalent, an optimal solution (λ^*, θ^*) to (6) is also an optimal solution to problem (2). Furthermore, since $U \subseteq E^s_+$, $V \subseteq E^m_+$, then $E^s_+ \subseteq -U^*$ and $E^m_+ \subseteq -V^*$. Although the conditions for optimal solutions of problem (6) are more restrictive than those of the corresponding CCR model, if DMU$_o$ is efficient according to problem (6), it must be efficient for the corresponding CCR model.

Now let $T = \{(X, -Y): (X, -Y) \in (\bar{X}\lambda, -\bar{Y}\lambda) + (-V^*, -U^*), \lambda \geq 0\}$ be the production possibility set. Then:

Definition 2. $(\bar{X}_o, -\bar{Y}_o) \in T$ is said to be a nondominated point of T associated with $V^* \times U^*$, if there exists no $(X, -Y) \in T$ such that

$$(X, -Y) \in (\bar{X}_o, -\bar{Y}_o) + (V^*, U^*), \qquad (X, -Y) \neq (\bar{X}_o, -\bar{Y}_o).$$

Given this definition Theorem 2 follows:

Theorem 2. Let $(\bar{X}_o, -\bar{Y}_o)$ be a nondominated point of T associated with $V^* \times U^*$. Then DMU$_o$ is efficient [see proof in C^2WH (1986)].

Facets, i.e., portions of the frontier which correspond to linear parts of the DEA derived efficient production function, are usually developed from the optimal basic LP solutions information for *in*efficient DMUs. Here the LP problem solved is that on the DEA side, e.g., that corresponding to (2) in the CCR case. We obtain a dual optimal solution from the basic optimal solution (e.g., that stemming from the determined optimal basis). Those corresponding to solution for *efficient* CCR DMUs we shall call 'ebd' (efficient basic duals) and we shall characterize DMUs as efficient on their being or not being in the new constraint cone. To repeat, an 'ebd' is the optimal dual vector pointing out from an optimal extreme point of the constraint set and provided in the optimal basic solution data of the LP problem solved with an extreme point method such as the simplex method.

Assume that we have chosen M ebd's from efficiency evaluations using the CCR model. Let these be $a^i = (\widehat{\mu^i}, \widehat{\omega^i})^T \in E_+^{m+s}$, $i = 1, \dots, M$. They have corresponding halfspaces $A_i = \{Z\colon\ a^{iT}Z \geq 0\}$, with boundary hyperplanes $B_i = \{Z\colon\ a^{iT}Z = 0\}$, $i = 1, \dots, M$. Clearly every observed vector of inputs and negative outputs $(x, -y)^T$ is a Z which is contained in at least one A_i.

Let

$$A = \bigcap_{i=1}^{M} A_i, \qquad A^j = \bigcap_{\substack{i=1 \\ i \neq j}}^{M} A_i.$$

Suppose we use a *subset* of the ebd's as spanning vectors for a constraint cone W. Then by renumbering if necessary,

$$W = \left\{ \sum_{i=1}^{k} \lambda_i a^i;\ \lambda_i \geq 0,\ i = 1, \dots, k \right\}, \qquad 1 \leq k \leq M.$$

Note that such a W is an acute cone, i.e., is properly contained in a half-space, and thereby $\mathrm{Int}(W^*)$ is not empty.

Lemma 1 below (to be proven in the appendix) shows that if an ebd is not in W, then its associated efficient DMU is no longer efficient under the cone-ratio model associated with W. In other words, these DMUs are dominated in the negative polar W^*.

Lemma 1. If a^j is not in W, and $Z_o \in B_j \cap \mathrm{Int}(A^j)$, then Z_o is not a nondominated point of A associated with W^*. I.e., there exists $Z \in A$, such that $Z \in Z_o + W^*/\{0\}$, where $W^*/\{0\}$ is W^* omitting $\{0\}$.

From $a^{jT}(Z_o - \beta \bar{Z}) = 0$ (see appendix) we know Z_o is in fact dominated by another DMU that is an extreme point located on the same facet. Note that Z_o is not an extreme point of A^j since $Z_o \in \mathrm{Int}(A^j)$. Hence, from Lemma 1, we immediately conclude:

Theorem 3. A DMU which is evaluated as efficient by the CCR model is inefficient within the cone-ratio CCR model if its ebd is not in the constraint cone employed.

Theorem 4. A DMU evaluated as efficient by the CCR model is still efficient under a cone-ratio CCR model with W constructed from ebd's if $(\hat{\mu}, \hat{\omega})$, an optimal dual solution ebd to the CCR model, is in the constraint cone (U, V) of the cone-ratio CCR model.

These theoretical conclusions are of practical importance. They make it possible for us to employ expert knowledge for improving a DEA analysis and to do so without unduly straining that knowledge. For example, we can use the ebd's of economically viable efficient DMUs as spanning vectors for the constraint cone and thereby determine which other efficient DMUs are economically viable according to our relevant information which is not in the input and output measures. This approach is illustrated in the evaluation of managerial performance in two bank examples here.

Which and how many ebd's, i.e., DMUs, should be employed in a cone-ratio sharpening of the CCR evaluations? Clearly the evaluations will depend on the efficient DMUs selected and also they can vary with the number employed. Generally, outside experts will agree unanimously on a *few* DMUs as being 'certainly' efficient. Thus for the 'sum' generation of a constraint cone, the top three DMU ebd's were employed. See Sun (1987).

3. Selection of cones for various purposes

We now illustrate use of polyhedral cones U and V in the sum form to tighten the criteria for efficiency evaluations of DMUs. These cones may be classified further into (a) those which emphasize individual inputs and/or outputs and (b) those which favor individual DMUs.

Classification (a): Cones emphasizing inputs and outputs

Example 1

We are to evaluate four DMUs which use two inputs to produce one output. The observed data are:

		DMU			
		1	2	3	4
$\bar{X}$	$\bar{x}_1$	1	4	1.5	4
	$\bar{x}_2$	5	1	1.5	2
$\bar{Y}$	$\bar{y}$	1	1	1	1

The CCR model will identify DMU1, DMU2, and DMU3 as efficient. Let us examine their efficiency again with a polyhedral cone-ratio model. For simplicity, we constrain only cone V and set $U = E^1_+$, the nonnegative real numbers.

Let $V = \{A^T\alpha: \alpha \geq 0\}$, then $-V^* = \{\omega: A\omega \geq 0\}$. Set

$$A^T = \begin{bmatrix} 1 & a^2 \\ a^1 & 1 \end{bmatrix},$$

so we have $-V^* = \{\omega: \omega_1 + a^1 w_2 \geq 0,\ a^2\omega_1 + w_2 \geq 0\}$. See fig. 1.

This is equivalent to using the CCR model to determine efficiency with the transformed inputs $\bar{X}' = A\bar{X}$ and the original output $\bar{Y}$ as represented in the following arrangement:

		DMU			
		1	2	3	4
$\bar{X}'$	$\bar{x}_1'$	$1 + 5a^1$	$4 + a^1$	$1.5 + 1.5a^1$	$4 + 2a^1$
	$\bar{x}_2'$	$a^2 + 5$	$4a^2 + 1$	$1.5a^2 + 1.5$	$4a^2 + 2$
$\bar{Y}$	$\bar{y}$	1	1	1	1

To simplify the formulae for geometric interpretation, if a^1 is sufficiently small and a^2 sufficiently large, and transformed data are equivalent for efficiency evaluation to the approximations:

		DMU			
		1	2	3	4
$\bar{X}'$	$\bar{x}_1'$	1	4	1.5	4
	$\bar{x}_2'$	a^2	$4a^2$	$1.5a^2$	$4a^2$
$\bar{Y}$	$\bar{y}$	1	1	1	1

since a^1 is dominated by the observed value of $\bar{x}_1'$ and a^2 dominates the observed value of $\bar{x}_2'$. Only DMU1, which originally used the least $\bar{x}_1'$, can survive the efficiency test under the constrained cone.

Conversely, if a^2 is sufficiently small and a^1 sufficiently large, only DMU2 which originally used the least $\bar{x}_2'$ will remain efficient. For the same reason, the transformed data are equivalent for efficiency evaluation to:

		DMU			
		1	2	3	4
$\bar{X}'$	$\bar{x}_1'$	$5a^1$	a^1	$1.5a^1$	$2a^1$
	$\bar{x}_2'$	5	1	1.5	2
$\bar{Y}$	$\bar{y}$	1	1	1	1

In the first case, V is heavily tilted toward input $\bar{x}_1'$. This shows that more emphasis is now put upon $\bar{x}_1'$. As a result, conserving $\bar{x}_1'$ becomes of key concern. It is not strange that only DMU1, which consumed the least $\bar{x}_1'$, can survive this condition seriously favoring $\bar{x}_1'$. On the other hand, in the second case, emphasis is directed to $\bar{x}_2'$ and it makes conservation of $\bar{x}_2'$ much more desirable. Hence only DMU2 remains efficient. If graphed, cone V is tilted toward axis $\bar{x}_2'$ in case two.

A convenient way to interpret the implication of these cones is to link them to the nondominated solution in the multi-objective programming problem. DMU2 and DMU3 are dominated in the polar cone $-V^*$ by DMU1 in case one; DMU1 and DMU3 are dominated by DMU2 in $-V^*$ in case two.

We see from the above examples that a constraint cone tilted toward any objective (input and/or output) emphasizes that objective. This provides us with the possibility of taking account of different concerns for objectives which may not be explicitly rendered in the observed quantities themselves.

Classification (b): Cones favoring DMUs

First, let us look at a special case that excludes 'weak efficient' DMUs, i.e., a case that ensures strict positivity of $(\hat{\mu}, \hat{\omega})$ in problem (1). Then, DMU_o is efficient if $\hat{\mu}^T \overline{Y}_0 = 1$. We need to construct a constraint cone to exclude the hyperplanes $(0, a_2, \dots, a_n)$, $(a_1, 0, \dots, a_n), \dots, (a_1, a_2, \dots, 0)$, but to include $(0, 0, \dots, 0)$. So we may set

$$A_{m \times m} = \begin{bmatrix} 1 & 1 & \cdot & \cdot & \cdot & \cdot & 1 & \varepsilon \\ 1 & 1 & \cdot & \cdot & \cdot & \cdot & \varepsilon & 1 \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ \varepsilon & 1 & \cdot & \cdot & \cdot & \cdot & 1 & 1 \end{bmatrix},$$

where ε is a positive number. $B_{s \times s}$ for output transformation is of the same matrix form. Example 2 illustrates how to detect the weak efficient DMUs using this approach.

Example 2

Consider the following seventeen DMUs. Each DMU uses two inputs to produce one output. The observed data are in the table 1.

Consider DMU3 and DMU10. They are scale- but not technically-efficient, i.e., $\theta^* = 1$ but the slacks of inputs are not all zero. Specifically, the slack of input 1 is 5 for DMU3 and the slack of input 2 is 10 for DMU10. While they may be termed 'weak efficient', they are not really efficient at all. But DMU3 and DMU10 seem to be fully efficient. (The θ^* of both are 1.0000.) We can use the polyhedral constraint cone described above with $\varepsilon = 0.01$ to uncover

Table 1

DMU	$\bar{y}$	$\bar{x}_1$	$\bar{x}_2$	DMU	$\bar{y}$	$\bar{x}_1$	$\bar{x}_2$
DMU1	2	10	10	DMU10	2	4	30
DMU2	2	20	5	DMU11	2	6	15
DMU3	2	30	4	DMU12	2	25	4
DMU4	2	27	9	DMU13	2	7	13
DMU5	2	14	8	DMU14	2	40	5
DMU6	2	5	20	DMU15	2	20.5	4.9
DMU7	2	4	20	DMU16	2	4.1	19.5
DMU8	2	12	18	DMU17	2	5	15
DMU9	2	8	12				

the 'true' *inefficiency of these DMUs* as follows. Take

$$\begin{bmatrix} B & 0 \\ 0 & A \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0.01 \\ 0 & 0.01 & 1 \end{bmatrix}.$$

With the transformed data (BY, AX), we get θ^* of DMU3 and DMU10 as 0.9884 and 0.9767, respectively.

The next example exhibits use of a polyhedral cone to identify the economically efficient DMUs.

Example 3

We use the same data as in Example 2. From the optimal solutions to the CCR model, we obtain ebd's. Now assume that market information indicates that the price ratio of inputs $\bar{x}_1$ and $\bar{x}_2$ are in the range k_1 to k_2, so that DMU managers should want to adjust their input consumption accordingly. If $k_1 = \frac{1}{5}$ and $k_2 = 1$, only the efficient DMUs whose $\widehat{\omega_2}/\widehat{\omega_1}$ are in the range $(\frac{1}{5}, 1)$ are economically efficient.

We see that we can use the $\hat{\omega}$ of DMU16 and DMU9 to establish the cone. Thus we take

$$\begin{bmatrix} B & 0 \\ 0 & A \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 \\ 1 & 0 & 0 \\ 0 & 0.125 & 0.025 \\ 0 & 0.05 & 0.05 \end{bmatrix}.$$

Evaluating with the transformed data $(B\overline{Y}, A\overline{X})$, we obtain the new efficiency scores as in table 2.

For a two-input case, as in this example, we may relate the constraint cone to the nondominated solutions associated with the negative polar of the

Table 2

DMU	Score	DMU	Score
DMU1	1.0000	DMU10	0.3333
DMU2	0.8000	DMU11	0.9524
DMU3	0.1923	DMU12	0.6897
DMU4	0.5556	DMU13	1.0000
DMU5	0.9091	DMU14	0.4444
DMU6	0.8889	DMU15	0.7874
DMU7	1.0000	DMU16	1.0000
DMU8	0.6667	DMU17	1.0000
DMU9	1.0000		

constraint cone. Here, we see that DMU2, DMU5, and DMU15 are no longer efficient, since their ebd's fall outside the range $(\frac{1}{5}, 1)$. But DMU1 is still efficient. As a matter of fact, its $\hat{\omega}$ could be any value between the ebd's of DMU9 and DMU5. With the constraint condition, DMU1 has the ebd of DMU9 ebd. To see this, note that DMU1 is an extreme point. Recall that from Lemma 1 contrarily, $Z_o = (\bar{x}_1, -\bar{y}_1)^T \in B_1$ but it is not in $\text{Int}(A^1)$.

We have thus illustrated how a constraint cone can be selected to favor desired patterns of input usage and output production in efficiency evaluation. Further, as shown in Charnes, Cooper, Wei, and Huang (1986) as well as in Sun (1987), these cone ratio approaches can be adapted for use with other models, such as the 'additive' model, which embody the DEA concepts and methods of computation and analysis.

4. Applications to commercial banks

We turn finally to a realistic application to large commercial banks. As reported in Sun (1987), the data involved were drawn from the call reports (1980–1985) to the FDIC[2] for 48 U.S. commercial banks drawn from the top 300 banks headquartered in America which are also members of FDIC.

Using expert advice from a banking specialist the following outputs and inputs were used in this study:

Outputs	Inputs
1. Total operating income	1. Total operating expense
2. Total interest income	2. Total noninterest expense
3. Total noninterest income	3. Provision for loan losses
4. Total net loans	4. Actual loan losses

[2]The supplemental data and expert opinions used are described in Sun (1987).

To be noted is that the provision for loan losses and actual loan losses treated as inputs are indicators of risks in banking operations. Total net loans is a measure of the size of services that a bank produces while the other inputs and outputs are mainly profit-related measures.

The results obtained from the CCR ratio model as applied to the data for these inputs and outputs were not satisfactory so recourse was made to a polyhedral cone-ratio DEA model with results that passed muster in subsequent reviews with wide experience in banking.

We use (6) to make the re-evaluation. The associated tranformational matrix $\begin{bmatrix} B & 0 \\ 0 & A \end{bmatrix}$ in 1983 is illustrated in table 3. It consists of optimal virtual multiplier vectors of three model banks: Morgan Guaranty, First Wachovia National Bank, and First Interstate of Nevada. It is a 6×8 matrix, where B and A are 3×4 matrices (A stands for the virtual multipliers of inputs and B the virtual multipliers of outputs), since we use three model banks and performance is measured in four inputs and four outputs.

We provide only a pair of examples to show what occurred and how the CCR model and its cone ratio extensions were used. For the first example, we use Citibank which, for 1983, showed the results listed under the column headed Value Observed. The column headed CCR model in table 4 shows the values for efficient performance as estimated by this model. The values exhibited under the column designated as cone-ratio CCR show the values which efficient performance would have exhibited as estimated with the cone-ratio CCR model.

As can be seen, the values in the latter two columns differ. The CCR model rated Citibank performance as efficient but the cone-ratio CCR model did not. The value of $\theta^* = 0.9693$ obtained from the latter model applied to all of the observed input values produces the values shown for these same inputs in the last column with the result that these inputs are all reduced by about 3%.

Turning to the output values, we obtain the adjustments needed for efficiency attainment by means of the formula

$$\bar{Y}_o^* = \sum_{j=1}^{n} \lambda_j^* \bar{Y}_j,$$

where the $\bar{Y}_j$ are the vectors of observed values which correspond to the efficient DMUs used in the evaluation of DMU_o and the λ_j^* are the optimal solution values. $\bar{Y}_o^*$ is the value corresponding to the point on the efficient facet from which the outputs observed in Y_o are evaluated.

In the case of Citibank's 1983 performance, the banks appearing in the optimal basis – and thus the banks used in evaluating the efficiency of Citibank's performance – are the Republic National Bank of New York and

Table 3

0.1221E − 6	0.5440E − 7	0.1727E − 9	0.3625E − 8	0.0	0.0	0.0	0.0
0.1275E − 5	0.9091E − 11	0.1727E − 9	0.1253E − 7	0.0	0.0	0.0	0.0
0.3699E − 5	0.9091E − 11	0.1727E − 9	0.3635E − 7	0.0	0.0	0.0	0.0
0.0	0.0	0.0	0.0	0.1582E − 9	0.1542E − 6	0.1402E − 6	0.1006E − 5
0.0	0.0	0.0	0.0	0.1582E − 9	0.1473E − 5	0.2393E − 6	0.3441E − 6
0.0	0.0	0.0	0.0	0.1582E − 9	0.4273E − 5	0.6940E − 6	0.9979E − 6

Table 4
Citibank (1983).

	Value observed	Value if efficient	
		CCR model	Cone-ratio CCR
	Output		
Total income	13572000	13572000	13443860
Interest income	10615000	10615000	10020451
Noninterest income	553000	553000	271151
Net loans	69286000	69286000	82397984
	Input		
Provisions	320000	320000	310176
Total expense	12171000	12171000	11797350
Noninterest expense	3061000	3061000	2967027
Loan losses	263000	263000	254926

Texas Commerce Bank, with λ_j^* values of 2.85 and 12.22, respectively. Applying these values to the 1983 data for these two banks produced the results for the output values shown in the upper part of the last column in table 4. Comparison with the observed values for Citibank showed that this would have resulted in a decrease of total income by some 1%, a decrease in interest income of 6%, and a decrease in noninterest income by 50% whereas net loans would have increased by 19%. To be noted, therefore, is the fact that the reduction of inputs (by some 3%) may then be accompanied by a decrease in some outputs and an increase in others.

For another example, we turn to Continental Illinois for 1984, which is known to have been a disastrous year for this bank. The data for this case and the corresponding CCR model and cone-ratio CCR model estimated efficiency adjusted values are shown in table 5 which has the same arrangement as table 4.

In this case, the CCR model gave a value of $\theta^* = 0.919$ which was reduced to 0.2351 by our cone-ratio extension of this model. Evidently a needed drastic reorientation of this bank's activities is signaled by the latter value, as

Table 5

Continental Illinois nb & tc (1984).

	Value observed	Value if efficient	
		CCR model	Cone-ratio CCR
	Output		
Total income	3998187	3998187	4209945
Interest income	3334291	3334291	3380729
Noninterest income	70064	96783	172986
Net loans	23693936	24791577	22922308
	Input		
Provisions	1171878	143749	275509
Total expense	3703887	3405380	870784
Noninterest expense	779890	717036	183352
Loan losses	1165487	96907	274006

was subsequently confirmed by the complete overhaul initiated with the FDIC bail-out attempt for Continental Illinois.

Turning from the inputs to the outputs for Continental Illinois in 1984, we observe that Wachovia National Bank and Trust Co. is the only DMU appearing in the basis from which Continental Illinois was evaluated. Thus applying the value of $\lambda_j^* = 4.74$ to the data for Wachovia in 1984, we obtain the new Y_o^* output values for Continental Illinois which are shown in the last column of table 5. Associated with this nearly 77% reduction in its inputs, as shown in table 5, Continental Illinois might also have increased its total income by 5% and its interest income by some 14% and 147%, respectively, while decreasing its total loans by nearly 4%.

Fig. 2 provides a geometric portrayal which can illustrate what is happening in the above cases. As is evident from these examples, output adjustments to attain efficiency in the case of the cone-ratio model need not be limited to movement in the 'northeast' direction, as is true for the CCR model. Thus, in the case shown in fig. 2, the output adjustment for *DMU5* is restricted to projections on *AA*′. In the cone-ratio CCR model, however, the projections can be to *BB*′.

5. Summary and conclusions

This extract from a more extended study should help to show some of the differences that may be expected as opportunities and vistas for research and use are opened by the new cone-ratio DEA models. Evidently a good deal of flexibility is added and ways are opened for the use of either expert opinion without strain or a knowledge of only ranges of values with associated

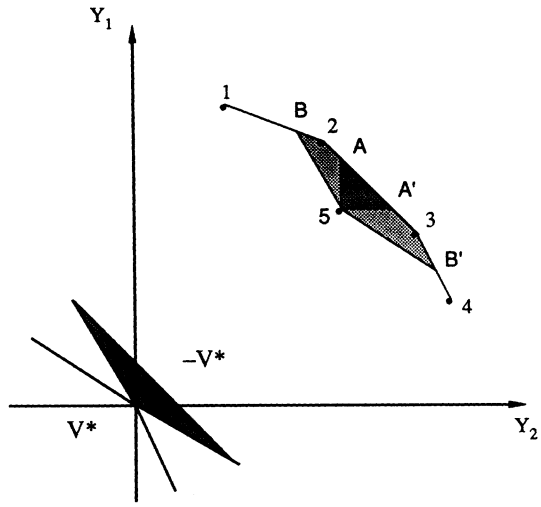

Fig. 2

inequalities can be employed. Needless to say, these uses can also provide guidance and act as a control on such opinions.[3]

These cone-ratio developments open other possibilities as well. For instance, the deficiencies exhibited by the ordinary CCR ratio model may reflect rather the fact that FDIC call report data are insufficient to provide all of the indicators needed to distinguish between efficient and inefficient performance.[4] Indeed, as shown in Charnes, Cooper, Golany, Halek, Schmitz, and Thomas (1986), uses of DEA admit of extensions that include 'goals' which might be specified for attainment as well as laws or regulations, risk factors, and/or economic 'climate'. Finally, cone-ratio extensions can be applied to the elimination of activities and/or merger schemes along the lines of what was done in Bessent, Bessent, Charnes, Cooper, and Thorogood (1983).

These generalize the old ratio models by allowing additional relative valuational considerations on inputs, outputs, and Pareto-optimality to appear in the form of cones which may differ from the nonnegative orthants appearing in the older models.

In this paper attention is focused on the 'sum' form for the cones and a means of eliciting these cones when direct relative valuation of inputs and

[3]See, e.g., the discussion in Thomas (1986) of the way DEA was used to guide and evaluate the performance of the auditors of the Texas Public Utility Commission in their managerial audits.

[4]See the similar comments in Divine (1986) on the use of DEA for effecting bond-rating evaluations for electric utilities which are more comprehensive than the ratings provided in Standard and Poor's or other bond rating services.

outputs is not available. This means uses expert opinion to select a few most efficient DMUs based on their experiential perceptions and generates from their optimal dual evaluators valuational cones in sum form.

The sum form for the cones has the important advantage that by a simple initial matrix calculation the problem is reduced to the old form with its powerful software vital for performing DEA analyses with the usual large numbers of DMUs to be evaluated and/or an empirical production function to be determined. The intersection form which has occurred in some so-called assurance region considerations can be transformed via a related matrix into the sum form and thereby can reap also the sum form advantages. Combinations of both types can also be made.

As noted, the sum form in equivalent intersection form is generally more complicated than the simple trade-off bounds utilized in assurance regions considerations. It would be of great interest to build up a set of typical applications equivalences as a means of better economic perception, and hopefully understanding, of the relevant trade-off ranges of input, output values arising in expert efficiency opinion.

Appendix

Lemma 1. If a^j is not in W, and $Z_o \in B_j \cap \mathrm{Int}(A^j)$, then Z_o is not a nondominated point of A associated with W^*. I.e., there exists $Z \in A$, such that $Z \in Z_o + W^*/\{0\}$, where $W^*/\{0\}$ is W^* omitting $\{0\}$.

Proof. Suppose, on the contrary, that there is no $Z \in A$ such that $Z \in Z_o + W^*/\{0\}$. Let $S = \{s\colon\ s \in W^*/\{0\} - Z + Z_o,$ for some $Z \in A\}$. It is easy to show that S is a convex set and 0 is not in S. By the separating hyperplane theorem for convex sets, there exists nonzero $p \in E^{m+s}$ such that $p^T s \leq 0$ for all $s \in S$.

For any $Z \in A$, $\lambda > 0$, and $w \in W^*/\{0\}$, let $S_{Z,\lambda,w} = -Z + Z_o + \lambda w$. Then $p^T Z_o + \lambda p^T w \leq p^T Z$ for all $Z \in A$, $\lambda > 0$, and $w \in W^*/\{0\}$. Hence

(a) $\quad p^T Z_o \leq p^T Z \quad$ for all $\quad Z \in A$,

(b) $\quad p^T w \leq 0 \quad$ for all $\quad w \in W^*/\{0\}$.

From (b), since W is an acute cone, $p \in (W^*/\{0\})^* = W$.

Now consider the system

$$\text{(c)} \qquad \begin{cases} p^T Z < 0, \\ a^{jT} Z = 0. \end{cases}$$

There must exist a solution $\bar{Z}$ to (c). Otherwise, for all Z satisfying $a^{jT} Z = 0$, we would have $p^T Z = 0$. In that event, there must exist a scalar h such that $a^{jT} = hp$ with $h > 0$ since $a^j \geq 0$, and $p \geq 0$. This leads to $a^j \bar{Z} W$, which contradicts our assumption.

Now let $\bar{Z}$ be a solution to (c) and consider the point $(Z_o - \beta\bar{Z})$. $Z_o \in \mathrm{Int}(A^j)$ and $a^{iT}Z_o > 0$ for $i \neq j$. There exists $\alpha < 0$ such that for $i \neq j$, $a^{iT}(Z_o - \beta\bar{Z}) = a^{iT}Z_o - \beta a^{iT}\bar{Z} \geq 0$ for all $\beta \in (\alpha, 0)$ and $a^{jT}(Z_o - \beta\bar{Z}) = a^{jT}Z_o - \beta a^{jT}\bar{Z} = a^{jT}Z_o = 0$. That means $(Z_o - b\bar{Z}) \in A$. But, $p^T(Z_o - b\bar{Z}) = p^T Z_o - p^T\bar{Z} < p^T Z_o$ for all $\beta \in (\alpha, 0)$, that contradicts (a). Hence, there exists $Z \in A$ such that $Z \in Z_o + W^*/\{0\}$. Q.E.D.

References

Bessent, A., W. Bessent, A. Charnes, W.W. Cooper, and N. Thorogood, 1983, Evaluation of educational proposals by means of DEA, Educational Administration Quarterly 19, 82–92.

Ben-Israel, A., A. Charnes, and K.O. Kortanek, March 1969, Duality and asymptotic solvability over cones, Bulletin of American Mathematical Society 75, 318–324.

Charnes, A. and W.W. Cooper, 1985, Preface to topics in data envelopment analysis, Annals of Operations Research 2, 59–94.

Charnes, A., B. Golany, R. Halek, G. Klopp, E. Schmitz, and D. Thomas, 1986, Data analysis approaches to policy evaluation and management of Army recruiting activities I: Tradeoffs between joint services and Army advertising, Research report no. 532 (Center for Cybernetic Studies, University of Texas, Austin, TX).

Charnes, A., L. Seiford, and J. Stutz, 1985, Foundations of Pareto–Koopmans optimality efficiency analysis and empirical production functions, Journal of Econometrics 30, 91–107.

Charnes, A., Z.M. Huang, and D.B. Sun, 1989, Relations between half-space and finitely generated cones in polyhedral cone-ratio DEA models, Research report no. 636 (Center for Cybernetic Studies, University of Texas, Austin, TX).

Charnes. A. and E. Rhodes, 1978, Measuring the efficiency of decision making units, European Journal of Operations Research 2, 429–444.

Charnes, A., 1981, Data envelopment analysis as an approach for evaluating program and managerial efficiency – with an illustrative application to the program follow through experiment in U.S. public school education, Management Science 27, 668–697.

Charnes, A., J. Rousseau, and J. Semple, 1978, Data envelopment analysis and axiomatic notions of efficiency and reference sets, Research report no. 558 (Center for Cybernetic Studies, University of Texas, Austin, TX); Arabian Journal of Science and Technology, forthcoming.

Charnes, A., Q.L. Wei, and Z.M. Huang, 1986, Cone-ratio data envelopment analysis and multi-objective programming, Research report no. 559 (Center for Cybernetic Studies, University of Texas at Austin, TX); International Journal of Systems Science 20, 1099–1118.

Charnes, A., L. Seiford, and J. Stutz, 1983, Invariant multiplicative efficiency and piecewise Cobb–Douglas envelopments, Operations Research Letters 2, 101–103.

Divine, J.D., 1986, Efficiency analysis and management of not-for-profit and governmentally regulated organizations, Ph.D. thesis (University of Texas at Austin, TX).

Huang, Z.M., 1985, The second order conditions of nondominated solutions for generalized multiobjective mathematical programming, Journal of Systems Science and Mathematical Science 5, 172–184.

Seiford, L. and R. Thrall, 1989, Recent developments in DEA: The mathematical programming approach to frontier analysis, Journal of Econometrics, this issue.

Sun, D.B., 1987, Evaluation of managerial performance of large commercial banks by data envelopment analysis, Unpublished Ph.D. dissertation (Graduate School of Business, University of Texas, Austin, TX).

Thomas, D.L., 1986, Auditing the efficiency of regulated companies: An application of data envelopment analysis to electric cooperatives (IC^2 Institute, University of Texas, Austin, TX).

Thompson, R.G., F.D. Singleton, R.M. Thrall, and B.A. Smith, December 1985, Comparative site evaluations for locating a high energy physics lab in Texas, Interfaces 16, 35–49.

Yu, P.L., 1974, Cone convexity, cone extreme point, and nondominated solutions in decision problems with multi-objectives, Journal of Optimization Theory and Application 14, 319–377.