A STRUCTURE FOR CLASSIFYING AND CHARACTERIZING EFFICIENCIES AND INEFFICIENCIES IN DATA ENVELOPMENT ANALYSIS

by

A. Charnes  W.W. Cooper
R.M. Thrall*
A STRUCTURE FOR CLASSIFYING AND CHARACTERIZING EFFICIENCIES AND INEFFICIENCIES IN DATA ENVIRONMENT ANALYSIS

by

A. Charnes        W.W. Cooper
R.M. Thrall*

Original Version: May 1985
First Revision: April 1986
Second Revision: September 1986

*Rice University, Houston, Texas

This research was partly supported by National Science Foundation Grants SES-8408134 and SES-8520806 and Office of Naval Research Grant N00014-86-C-0398 with the Center for Cybernetic Studies, The University of Texas at Austin. Reproduction in whole or in part is permitted for any purpose of the United States Government.
DEA (Data Envelopment Analysis) attempts to identify sources and estimate amounts of inefficiencies contained in the outputs and inputs generated by managed entities called DMUs (= Decision Making Units). Explicit formulation of underlying functional relations with specified parametric forms relating inputs to outputs is not required. An overall (scalar) measure of efficiency is also obtained for each DMU from the observed magnitudes of its multiple inputs and outputs without requiring use of a priori weights or relative value assumptions. A partition of DMUs into six classes is established via primal and dual representation theorems and three classification theorems which do not depend on non-archimedian analysis. Earlier theory is extended to explain the consequences of zero inputs and outputs and to utilize zero virtual multipliers (shadow prices). Three of the six classes are scale inefficient and two of the three scale efficient classes are also technically (= zero waste) efficient.

**KEYWORDS:** dual linear programs; multicriterion efficiency analysis, scale efficiency, strong complementary slackness, technical efficiency, virtual multipliers.
1. BACKGROUND AND INTRODUCTION.

Data Envelopment Analysis (DEA), as reported in [8], is an approach to measuring the efficiency of entities, e.g., not-for-profit entities, which have multiple outputs and multiple inputs that do not readily submit to treatment via unit prices or other a priori weights that might be used (e.g., in index numbers to secure comparability. In addition to providing meaningful scalar efficiency values, DEA is designed to identify sources and estimate the amounts of inefficiencies that might be present in the various output and input vectors. In these respects DEA may be likened to standard cost systems used in manufacturing enterprises although, unlike the latter, it does not require prior undertakings of extensive engineering studies. Also, unlike the latter systems, it seeks to identify and estimate inefficiencies in outputs as well as inputs. Finally DEA seeks to accomplish all of this without requiring a priori specification of the underlying functional forms -- which can be a forbidding task in statistical regression analyses and like approaches to these same problems. See [6].

The basic theory for DEA as it exists at this point in time is set forth in [9] -- see also [8]. The purpose of the present paper is to relax some of the present conditions which restrict DEA uses to situations in which (a) all inputs and outputs are positive and (b) all variables are confined to positive ranges in the solution sets (see the discussion of D-proper virtual multipliers in Section 3). The principal results are: the primal and dual representations in Theorems 4 and 5, the

---

1 See [13] and [14] for difficulties even when so-called flexible functional forms are used.

2 We shall also abbreviate other references such as CCR [p. 41]
characterization of decision making units (DMU's) in Theorems 6, 7, and 8, and the analysis of the role of zero inputs and outputs in Theorem 18 and its corollaries.

Expanding the uses of DEA for evaluating possible additions or deletions of individual product lines is illustrative of possibilities for managerial uses at the level of individual Decision Making Units (DMUs). See [5] for a discussion of the possible interacting effects on efficiency evaluations that can occur with decisions to add or delete activities in the programs of a community college. The ability to bring DEA to bear on issues such as the evaluation of natural monopoly provides another class of examples in the arena of public policy analysis. See the treatment of "economies of scope" and "economies of scale" and their relation to issues of natural monopoly in the recent breakup of AT&T as discussed in [14] and [15].

To make this more precise assume that we have \( j = 1, \ldots, n \) DMUs, for each of which an observed vector of outputs \( y \), and an observed vector of inputs \( x \) is available. Thus for DMU \( j \) we have

\[
y_j = \begin{bmatrix} y_{1j} \\ \vdots \\ y_{rj} \\ \vdots \\ y_{sj} \end{bmatrix}, \quad x_j = \begin{bmatrix} x_{1j} \\ \vdots \\ x_{ij} \\ \vdots \\ x_{mj} \end{bmatrix}
\]

with \( y_{rj} > 0, \ r = 1, \ldots, s \) and \( x_{ij} > 0, \ i = 1, \ldots, m \) for every \( j = 1, \ldots, n \).

Using the CCR ratio form from [11] we may then evaluate the efficiency of each (or all) of these DMUs relative to the others via
\[(2) \quad \max h_0(u,v) = \frac{u^T y_0}{v^T x_0}\]

subject to

\[1 \geq \frac{u^T y_j}{v^T x_j}, \quad j = 1, \ldots, n\]

and

\[\rho < e^T \leq \frac{u^T}{v^T x_0}, \quad \frac{u^T}{v^T x_0}\]

where T indicates transposition. The vectors u and v are called "virtual multipliers", with s and m components, respectively, while the vector \(e^T\) has all of its elements to unity with the number of its elements determined by context. E.g., \(e^T\), above is \(1 \times s\) and is multiplied by a "small" non-Archimedean constant, \(\epsilon\), by which we mean that each of the components of \(\epsilon e^T\) is greater than zero but smaller than any positive real number. In particular, given any positive real number \(z\) there is no real scalar \(k\) with the property

\[(3) \quad k \epsilon \geq z.\]

In the above formulation \(y_0\) and \(x_0\) are the output and input vectors for one of the \(j = 1, \ldots, n\) DMUs incorporated in the constraints. Hence we have

\[\max h_0(u,v) = h^*_0 \leq 1, \text{ and via the following}^{1} \text{ we have:}\]

\(^1\text{First given in [10] and reproduced in [8].}\]
Non-Archimedean Efficiency Theorem:

Part 1: \( h^*_o = 1 \) if and only if DMU \( o \) is efficient in all inputs and outputs.

Part 2: \( h^*_o < 1 \) if and only if DMU \( o \) is inefficient in some inputs or outputs.

To obtain this result and also to place the DEA problem in readily computable form we used the Charnes-Cooper transformation of fractional programming, viz.,

\[
\begin{align*}
\nu &= tu, \quad \omega = tv, \quad t \geq 0
\end{align*}
\]

to obtain the following dual pair of linear programming problems:

\[
\begin{align*}
\text{max} & \quad \nu^T y_o \\
\text{subject to} & \quad \frac{\phi}{\nu} = \omega^T x_o = tv^T x_o = \omega^T x_o \\
& \quad \phi \geq 0
\end{align*}
\]

\[
\begin{align*}
\text{min} & \quad 0 \\
\text{subject to} & \quad -\phi \geq -\omega^T X - s^- \\
& \quad y_o = \nu^T y - \omega^T X \\
& \quad s^+ \geq s, \quad s^-, s^+
\end{align*}
\]

where the condition \( \omega^T x_o = 1 \) insures that the scalar \( t \) used in (4) will be positive so the reverse transformation may be used to secure (2) from (5).

All of the theorems of ordinary linear programming continue to hold for the non-Archimedean extensions included in (5). Both problems have feasible solutions. Hence the dual theorem of linear programming gives the equality in

\[\text{See [8]}\]
where the asterisks indicate optimum values and to establish the inequality
one need only observe that \( w^T X \geq v^T Y \) is required for the problem on the left
in (5) and this requires \( x^T X_o \geq y^T Y_o \) which, together with the condition
\( w^T X_o = 1 \), implies \( y^T Y_o \) for feasibility.

The expression \( v^T y_o \) on the left of (6) can be a real number only if
the components of \( s^- \) and \( s^+ \) are all zero on the right. With \( s^{**} = 0 \) it is
not possible to improve the output performance by augmenting the correspond-
ing \( y_o \) components by the amount of these non-zero slacks. But even if \( s^{**} = 0 \) it may still be possible to improve the inputs, if the optimal value of
the scalar \( \theta^* < 1 \). If, however, \( \theta^* = 1 \) and all slacks are zero then also
\( v^T y_o = 1 \) and all components of \( v^T \) are real valued. On the other hand if
any components of \( s^- \) or \( s^+ \) are not zero then we must have \( v^T y_o < 1 \) and
\( v^T \) must contain some non-Archimedean components. Indeed transposing back
to (2) via (4) the sources (and amounts) of inefficiencies can be identified
that produce \( h_o^* < 1 \) while, conversely, if \( h_o^* = 1 \) then also \( v^T y_o = 1 \) and no
inefficiencies are present.

Examples.

As these developments show, these non-Archimedean elements play an
important role in the theory and possible uses of DEA. To explore possible
consequences of eliminating these non-Archimedean elements we utilize the
following example which involves \( j = 1, 2 \) DMUs each of which produces 1 unit
of output (indicated in the first row) and each of which utilizes two inputs in the amounts indicated in rows 2 and 3.

\[
\begin{array}{cc}
\text{DMU} & \\
1 & 2 \\
\hline
y_{1j} & 1 & 1 \\
x_{1j} & 1 & 1 \\
x_{2j} & 5 & 6 \\
\end{array}
\]

The DEA efficiency evaluation problems are, respectively, for DMU_1

\[
\begin{align*}
\max h(w) &= \frac{u_1}{v_1 + 5v_2} \\
\text{subject to} & \\
1 & \geq \frac{u_1}{v_1 + 5v_2} \\
1 & \geq \frac{u_1}{v_1 + 6v_2}
\end{align*}
\]

and for DMU_2

\[
\begin{align*}
\max h(w) &= \frac{u_1}{v_1 + 6v_2} \\
\text{subject to} & \\
1 & \geq \frac{u_1}{v_1 + 5v_2} \\
1 & \geq \frac{u_1}{v_1 + 6v_2}
\end{align*}
\]

All variables are merely constrained to be non-negative and not required to be positive as before. Here

\[w^T = (u_1, v_1, v_2)\]

is a member of the set of all "virtual multipliers" \(W\) when the elements of \(w\) are all non-zero and non-negative. If \(W_1\) and \(W_2\) represent the set of optimizing multipliers for, respectively, DMU_1 and DMU_2 in the above example, we have \(W_1 = W\) but \(W_2 \neq W\) since an optimum is achieved for DMU_2 only when \(v_2 = \phi\).
With \( v_2 = \phi \) an optimum is achieved for \( \text{DMU}_2 \) with \( h(w) = 1 \) so that its behavior is incorrectly characterized as efficient in all inputs and outputs. This is a consequence of eliminating the non-Archimedean condition on the variables in (2), and, as already noted, the appearance of these non-Archimedean elements in an optimum solution make it possible to identify the sources of inefficiency and estimate their amounts by reference to the variables with which they are associated in (5).

In the section that follows we shall give an entirely new route which will associate with \( \text{DMU}_2 \) a multiplier set \( W_2 \) related to both sets of LP problems. In the above example we will find that \( \dim W_1 = 3 \), which is the maximum number of positive components for variables in the above problem, whereas \( \dim W_2 = 2 \) because \( v_2^* = \phi \) in every optimum solution for \( \text{DMU}_2 \).

This suggests there may be a way of classifying DMUs by reference to the dimensions of their multiplier sets and that proceeding in this manner will enable us to retrieve what was lost by eliminating the non-Archimedean elements, \( \epsilon \). Thus, as was done in [16] we partition the set of all DMUs into the following six classes

\[ E, E', F, \text{NE}, \text{NE}', \text{NF} \]

where the first three classes are on a frontier and the second three are not. The following figure adapted from p.46 represents the situation for two inputs and one output. The solid lines connecting the points \( E \) and \( E' \) represent frontier points that are efficient. Points like \( F \) which are also on the frontier are not efficient since, as in the case shown, one of the

\[ \text{That is, the DMUs associated with these points are at MPSS (Most Productive Scale Size) in the sense of Banker [2].} \]
inputs may be reduced and this can continue until E is achieved without reducing output or increasing the other input.

All of the points E, E', F have optimal O values of unity in (5) and hence are scale efficient in the sense of Banker [2]. As already noted, however, F is not technically efficient. Finally the points NF, NE, and NE' are not on the frontier and hence are neither technically nor scale efficient.

In the above example DMU₁ is associated with E and DMU₂ with F since optimality is achieved with \( h^* = 1 \) in both cases. The sets E' and I are empty but, of course, this need not be true in other examples -- to which we shall turn after introducing suitable notations and definitions to cover all possible cases in the section that follows.

This paper is a technical supplement to [16] which will also be referenced as CCT. The other two principal references are [11], also referenced as CCR, which was the first of a long and continuous series of papers on DEA, and [8] which is a recent review and codification of DEA theory and contains an extensive bibliography for DEA.

This paper contains proofs of the new results announced in CCT [16]. Although we will rely for the most part on the notations and terminology of the three principal references the proofs require introduction of a number of new concepts and notations.

2. DATA DOMAINS.

Data Envelopment Analysis (DEA) is a framework for investigating the efficiency of a collection of Decision Making Units (DMUs) with common inputs and outputs see CCR [11], pp. 429-430 and [8], pp. 1-6.
We consider \( n \) DMUs, each with \( m \) inputs and \( s \) outputs. We define a data domain \( D \) as a (ordered) set
\[
J = (DMU_1, ..., DMU_n)
\]
of DMUs and a matrix
\[
P = [P_1, ..., P_m]
\]
with \( s + m \) rows and \( n \) columns. The \( j \)th column
\[
P_j = \begin{bmatrix} y_j \\ -x_j \\ \vdots \end{bmatrix}
\]
is made up from an input vector \( x_j \) whose \( i \)th component \( x_{ij} \) is the amount of input \( i \) used by \( DMU_j \) and an output vector \( y_j \) whose \( r \)th component \( y_{rj} \) is the amount of output \( r \) produced by \( DMU_j \). We write
\[
D = (J, P).
\]

We assume that the components \( x_{ij} \) and \( y_{rj} \) are positive or zero and require that for no \( j \) is either \( x_j = 0 \) or \( y_j = 0 \).

We say that a data domain \( D \) is in reduced form if for no pair \((j,k)\) with \( j \neq k \) and scalar \( \alpha \) is
\[
P_k = \alpha P_j.
\]

Let \( J^* \) be a subset of \( J \), and let \( P^* \) be the submatrix of \( P \) consisting of those columns \( P_j \) for which \( DMU_j \in J^* \). Then we call the data domain
\[
D^* = (J^*, P^*)
\]
a subdomain of \( D \).

Associated with a data domain \( D \) is a set of vectors
\[
K(P) = \{ p | p \leq P \lambda = P_1 \lambda_1 + ... + P_n \lambda_n, \lambda \geq 0 \}
\]
called the **conical hull** of \( P \).

Let \( D = (J, P) \) be a data domain and let \( D^* \) be the subdomain of \( D \) defined as follows. Let

\[
J^* = (DMU_j \text{ in } N \text{ for which there is no } k < j \text{ with } P_k \text{ proportional to } P_j).
\]

Then, clearly, \( D^* \) is reduced.

3. **VIRTUAL MULTIPLIERS AND DEA-SCALE-EFFICIENCY.**

Let \( u \) and \( v \) be non-negative, non-zero vectors with \( s \) and \( m \) components, respectively. We call the vector

\[
w = \begin{bmatrix} u \\ v \end{bmatrix}
\]

a **virtual multiplier** and denote by \( W \) the set of all virtual multipliers.

Next, for \( w \in W \) and for any \( DMU \), let \( g_j(v) = v^T x_j \), \( f_j(u) = u^T y_j \). Then, if \( g_j(v) > 0 \) let

\[
h_j(w) = f_j(u)/g_j(v).
\]

We say that \( w \) is **D-improper** if either

\[
(a) \quad f_j(u) = 0 \text{ for every } DMU
\]

or

\[
(b) \quad \text{for some } DMU_j, f_j(u) > 0, g_j(u) = 0.
\]

Otherwise we call \( w \) **D-proper.**

If \( w \) is D-proper and if

\[
h_j(w) \geq h_k(w) \text{ for all } DMU_k \text{ for which } g_k(v) > 0
\]
then we call $w$ an optimizing multiplier for $DMU_j$ and denote by $W_j$ the set of all optimizing multipliers for $DMU_j$. We call $W_j$ the multiplier set for $DMU_j$.

Notes: (a) If $w$ is D-improper it is not assigned to any $W_k$. In Section 12 we will consider the significance of the D-improper virtual multipliers. (b) If $f_k(u) = g_k(v) = 0$ then $h_k(w)$ has the form $0/0$ and is not only not defined but is not considered in or needed in applying (17).

(c) If $w = \begin{bmatrix} u \\ v \end{bmatrix} \in W_j$, then so is $w' = \begin{bmatrix} au \\ \beta v \end{bmatrix}$ for all $\alpha > 0, \beta > 0$.

$W_j$ may be empty; if so we say that $DMU_j$ is DEA-inefficient and denote by $N$ (or $N(D)$) the set of all DEA-inefficient DMUs.

If $W_j$ is not empty we say that $DMU_j$ is DEA-scale-efficient and denote by $RE$ (RE(D)) the set of all DEA-scale-efficient DMUs.

It follows readily from (17) that if $D^*$ is a subdomain of $D$ then

(18) $N(D^*) \subset N(D)$

and

(19) $N^* \cap RE(D) \subset RE(D^*)$.

Lemma 1. Suppose $k \neq h$, $P_k = \alpha P_h$ and let $J^*$ be $J$ with $DMU_k$ deleted.

Then:

(20) $K(P^*) = K(P)$

(21) $N(D^*) = N(D) \cap N^*$

and

(22) $RE(D^*) = RE(D) \cap J^*$. 

11
Proof: to establish (20) we observe that
\[ P\lambda = P^*\lambda^* \]
where, for \( j \neq h, k \), \( \lambda^*_j = \lambda_j \) and \( \lambda^*_h = \lambda_h + a\lambda_k \). By the non-negativity of inputs and outputs, \( a > 0 \) so that \( \lambda^*_h \geq 0 \). For (21) and (22) we note from (16)
\[ h_k(w) = h_h(w) \] for all \( w \), if \( P_k = aP_h \).

Corollary 2. Let \( D^* \) be the reduced subdomain obtained from \( D \) according to (14). Then (20), (21), and (22) hold for \( D^* \).

The conclusions of Corollary 2 hold as well for any reduced subdomain \( D^* \) obtained from \( D \) by successive deletion of DMUs proportional to some remaining one.

It follows from Corollary 2 that for questions relating to scale efficiency or to the extended domain we may without loss of generality limit ourselves to reduced data domains.

Reduction assumption: Henceforth, unless otherwise specified, we assume that all data domains \( D \) considered are reduced.

4. CLASSIFICATION OF DEA EFFICIENCIES.

For our purposes, we make no distinction between two DMUs whose input-output vectors are equal. Hence we may identify a DMU with its input-output vector \( p = [y \ -x] \).
Conceptually, any vector \( p \) with \( x \) and \( y \) non-negative and non-zero may be treated as a DMU. For any \( p \) in \( K(P) \) we construct a new data domain \( D^+ = D^p \) with

\[
J^+ = (p_1, \ldots, p_n, p) \text{ and } P^+ = [Pp].
\]

(\( D^+ \) may not be reduced.)

We next define the frontier \( FR \) of \( D \) as the set of all \( p \) in \( K(P) \) for which \( p \) is scale efficient in \( D^+ = D^p \). Clearly,

\[
RE \subset FR.
\]

In what follows we will give characterizations of \( FR \) in terms of \( RE \).

We partition the set \( RE \) of scale efficient DMUs into three classes.

\[
E = \{ \text{DMU} k \in RE | \dim W_k = s + m \},
\]

\[
E' = \{ \text{DMU} k \in RE | \dim W_k < s + m \text{ and there exists } w \in W_k \text{ such that } w > 0 \},
\]

\[
F = \{ \text{DMU} k \in RE | \text{every } w \in W_k \text{ has at least one zero component} \}.
\]

We use the terms DEA-extreme-efficient, and DEA-non-extreme-efficient to describe, respectively, DMUs in \( E \) and \( E' \). We do not consider elements of \( F \) to be DEA-efficient because (as will be seen below) the presence of zeros is related to the existence of slacks, which in turn indicate what has been called technical inefficiency. Thus, \( E \cup E' \) is the set of all DEA-efficient DMUs, and \( F \) consists of the DMUs that are scale but not technically efficient.
5. SOMEEXAMPLES.

Three related simple examples are displayed. Each column gives the components \( y_{rij} \) and \( x_{ij} \) of the output and input vectors \( y_j \) and \( x_j \) of DMU \( j \). We recall that \( W \) is the set of all virtual multipliers

\[
W = \begin{bmatrix} u \v v \end{bmatrix},
\]

where \( u \) is a non-negative, non-zero \( s \) vector and \( v \) is a non-negative, non-zero \( m \) vector.

In Example 1:

\[ W_1 = W, \quad W_2 = \{ \text{all } w \text{ with } v_2 = 0 \} \]
\[ \dim W_1 = j = m + s \text{ so that } \text{DMU}_1 \in E \]
\[ \dim W_2 = 2 < m + s \text{ and has no positive vector so that } \text{DMU}_2 \in F. \]

In Example 2:

\[ W_1 = \{ \text{all } w \text{ with } v_3 \leq v_2 \} \]
\[ W_2 = \{ \text{all } w \text{ with } v_2 \leq v_3 \} \]

Both \( W_1 \) and \( W_2 \) have dimension \( 4 = m + s \) and hence \( E = \{ \text{DMU}_1, \text{DMU}_2 \} \). \( E', \ F, \) and \( N \) are empty. Note also that
\[
W_1 \cap W_2 = \{ \text{all } w \text{ with } v_2 = v_3 \}
\]
has dimension 3 which is less than \( 4 = m + s \).

In Example 3:

\[ a \text{ and } b \text{ are (positive) parameters and we distinguish three cases.} \]

Case 3.1: \( a + b < 1 \). Then

\[ W_1 = \{ \text{all } w \text{ with } v_3 \leq \frac{a}{1-b} v_2 \}, \]
\[ W_2 = \{ \text{all } w \text{ with } v_3 \geq \frac{(1-a)/b} v_2 \}, \]
\[ W_3 = \{ \text{all } w \text{ with } (a/(1-b))v_2 \leq v_3 \leq ((1-a)/b)v_2 \}. \]

Each of these has dimension \(4 - m + s\) and hence \(E = \{ \text{DMU}_1, \text{DMU}_2, \text{DMU}_3 \}\).

**Figure 1: Projection on \(v_2, v_3\) plane**

Figure 1 is the projection of \(W\) onto the \(v_2, v_3\) plane. Note that from \(a + b < 1\) we conclude that \((1-a)/b > 1 > a/(1-b)\) and hence the multipliers \(w\) with \(v_3 = v_2\) belong to \(W_3\).

Case 3.2: \(a + b = 1\).

The definitions for the \(W_i\) from case 3.1 are still valid but now \(a/(1-b) = (1-a)/b = 1\). Hence, \(\dim W_1 = \dim W_2 = 4, \dim W_3 = 3\) so that \(E = \{ \text{DMU}_1, \text{DMU}_2 \}, E' = \{ \text{DMU}_3 \}\), and \(F\) and \(N\) are empty.

Note that as \(a + b\) increases toward \(a + b = 1\) in Case 3.1, the bounding lines in Figure 1 approach as a limiting position the line \(v_2 = v_3\).

In Case 3.2

(30) \[ W_3 = W_1 \cap W_2 \]

and does not depend on the value of \(a\) (so long as \(a + b = 1\)). In this case we have
Figure 2 illustrates $P_3$ for several values of $a$.

The independence of $W_3$ from $a$ illustrates a property of the multipliers which requires special attention in some of the proofs which follow.

Case 3.3: $a + b > 1$.

In this case $W_1$ and $W_2$ are the same as in Case 3.2 but $W_3 = \{\text{all w with } v_2 = v_3 = 0\}$ has dim 2. Here $E = \{DMU_1, DMU_2\}$, $F = \{DMU_3\}$, and $E'$ and $N$ are empty. We can write

$$W_3 = W_1 \cap W_2 \cap \{\text{all w with } v_2 = v_3 = 0\}$$

and for $1-b \leq c \leq a$ we have

$$P_3 = (1-c)P_1 + cP_2 + R(c)$$

where

$$R(c) = \begin{bmatrix} 0 \\ 0 \\ a-c \\ c+b-1 \end{bmatrix}$$
The non-zero components of \( R(c) \) are called "slacks" and measure the technical inefficiency of DMU. The lack of uniqueness of the expression (33) for \( P \) and the consequential lack of uniqueness of the slack vector is significant in the theory of DEA-efficiency. [The fact that the sum \( a + b - 1 \) of the slacks in (34) is unique is not generally the case. For example, if we replace the value \( x_{33} = 6 \) with \( x_{33} = 5.9 \) the sum of the slacks \( a + b - 1 + (0.1)c \) is not independent of \( c \).]

Table 4 gives an example which shows that DEA-efficiency, by itself, may not always guarantee desirability. In this example with \( m = 4, s = 3, n = ms = 12 \) (clearly generalizable for any \( m \) and \( s \)) every DMU is extremely efficient. To see this write \( j \) in the form

\[
j = (\alpha - 1)s + \beta
\]

where \( 1 \leq \alpha \leq m \) and \( 1 \leq \beta \leq s \). Then, if we define \( w_j \) by

\[
(35) \quad u_{rj} = \delta_{r\beta} \quad (r = 1, \ldots, s),
\]

\[
v_{ij} = \delta_{i\alpha} \quad (i = 1, \ldots, m)
\]

where \( \delta_{pq} \) is the kronecker delta, we have

\[
(36) \quad h_j(w_j) = 3 > h_k(w_j) \text{ for all } k \neq j.
\]

From (36) it follows at once that DMU \( e \in RE_j \); however, from Theorem 6(E2) proved later, the "greater than" in (36) implies also that DMU \( e \in E \), i.e., that \( \dim W_j = s + m - 7 \).
\[
\begin{array}{ccc}
\text{Example 1: } m = 2, & \text{Example 2: } m = 3, \\
\text{s = 1, } n = 2 & \text{s = 1, } n = 2 \\
\end{array}
\]
DMU

<table>
<thead>
<tr>
<th></th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>7</th>
<th>8</th>
<th>9</th>
<th>10</th>
<th>11</th>
<th>12</th>
</tr>
</thead>
<tbody>
<tr>
<td>output</td>
<td>1</td>
<td>3</td>
<td>2</td>
<td>2</td>
<td>3</td>
<td>2</td>
<td>2</td>
<td>3</td>
<td>2</td>
<td>2</td>
<td>2</td>
<td>2</td>
</tr>
<tr>
<td></td>
<td>2</td>
<td>2</td>
<td>3</td>
<td>2</td>
<td>2</td>
<td>3</td>
<td>2</td>
<td>2</td>
<td>3</td>
<td>2</td>
<td>2</td>
<td>2</td>
</tr>
<tr>
<td></td>
<td>3</td>
<td>2</td>
<td>2</td>
<td>3</td>
<td>2</td>
<td>2</td>
<td>3</td>
<td>2</td>
<td>3</td>
<td>2</td>
<td>2</td>
<td>3</td>
</tr>
<tr>
<td>input</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>2</td>
<td>2</td>
<td>2</td>
<td>2</td>
<td>2</td>
<td>2</td>
<td>2</td>
<td>2</td>
</tr>
<tr>
<td></td>
<td>2</td>
<td>2</td>
<td>2</td>
<td>2</td>
<td>2</td>
<td>2</td>
<td>2</td>
<td>1</td>
<td>1</td>
<td>2</td>
<td>2</td>
<td>2</td>
</tr>
<tr>
<td></td>
<td>3</td>
<td>2</td>
<td>2</td>
<td>2</td>
<td>2</td>
<td>2</td>
<td>2</td>
<td>1</td>
<td>1</td>
<td>2</td>
<td>2</td>
<td>2</td>
</tr>
<tr>
<td></td>
<td>4</td>
<td>2</td>
<td>2</td>
<td>2</td>
<td>2</td>
<td>2</td>
<td>2</td>
<td>2</td>
<td>2</td>
<td>1</td>
<td>1</td>
<td>1</td>
</tr>
</tbody>
</table>

Example 4: an example with every DMU extremely efficient

5. A LINEAR PROGRAMMING FORMULATION.

Following CCT [16] we next turn to a linear programming formulation of DEA-efficiency. Let \( \alpha \) and \( \beta \) be positive scalars and let

\[
\mathbf{W}' = \begin{bmatrix} \alpha \mathbf{u} \\ \beta \mathbf{v} \end{bmatrix}
\]

Lemma 3. For any \( k \), \( \mathbf{w} \in \mathbf{W}_k \) iff \( \mathbf{w}' \in \mathbf{W}_k \).

This lemma follows from (16) and (17).

In view of Lemma 3 the determination as to whether or not \( \text{DMU}_j \) is DEA-scale-efficient is equivalent to the existence or non-existence of a virtual multiplier \( \mathbf{w} \) for which

\[
1 - h_j(\mathbf{w}) + \lambda_k(\mathbf{w}) \text{ for all } \text{DMU}_k \text{ for which } g_k(\mathbf{v}) > 0
\]

or even, indeed, for which in addition to (38)

\[
g_j(\mathbf{v}) = 1.
\]

Of course (38) and (39) also imply

\[
f_j(\mathbf{u}) = 1.
\]

For any \( j \) we denote by \( \mathbf{W}_j^m \) the set of all \( \mathbf{w} \) (if any) which satisfy (38) and (39). We call \( \mathbf{W}_j^m \) the normalized multiplier set for \( \text{DMU}_j \). Clearly, if \( \mathbf{w}' \) and \( \mathbf{w} \) satisfy (37) then
(41) \[ W_j = \{w| \text{there exist } a \text{ and } b \text{ for which } w' \in W_j^m\}. \]

Let "o" stand for some integer from \(\{1, \ldots, m\}\). We associate with DMU\(_o\) the scalar \(z^*\) defined by the non-linear optimization problem:

\[
(42) \quad z^* = \max \{h_o(w)|\text{D-proper } w \text{ for which } h_j(w) \leq 1 \text{ for all } j \text{ with } g_j(v) > 0\}. 
\]

Clearly \(z^* = 1\) iff DMU\(_o\) is scale efficient. Moreover, if \(z^* = 1\) there is a \(w \in W^m_o\) which solves (42), i.e., which satisfies (39) for DMU\(_o\).

Since, when \(g_j(v) > 0\)

\[
(43) \quad h_j(w) \leq 1 \quad \text{iff} \quad f_j(u) - g_j(v) \leq 0
\]

and since \(f_j(u) - g_j(v) = 0 - 0 \leq 0\) for any D-proper \(w\) with \(g_j(v) = 0\) we can replace the non-linear system (42) and (39) by the linear program

\[
(44) \quad \max z = \sum_r y_r u_r \\
\text{subject to} \\
\sum_r y_r u_r - \sum_j x_{ij} v_i \leq 0 \quad j=1, \ldots, n \\
\sum_i x_{io} v_i = 1 \\
u_r \geq 0 \quad r=1, \ldots, s \\
v_i \geq 0 \quad i=1, \ldots, m
\]

This program may be regarded as the dual of a related linear program which we will reference as the primal problem. The primal and dual problems are shown in Table 1.
In Table 1 the vector \( \lambda \) defined by \( \lambda_0 = 1, \lambda_j = 0 \) for \( j \neq 0 \), and corresponding \( \theta^0 = 1 \) is always feasible. On the other hand, by hypothesis, there is at least one \( i \) for which \( x_{i0} > 0 \). Then the primal relation

\[
x_{i0}\theta - \sum_j x_{ij}\lambda_j \geq 0
\]

requires the \( \theta \) be non-negative for any feasible \( (\lambda, \theta) \). Hence, the primal problem will have an optimal feasible solution \( (\lambda^*, \theta^*) \) with \( 0 \leq \theta^* \leq 1 \).

We can strengthen this result to

\[0 < \theta^* \leq 1\]

For, let \( (\lambda, \theta) \) be feasible for the primal. Then \( \lambda \neq 0 \) else

\[0 = \sum_j y_{rj}\lambda_j \geq y_{ro}\]

contrary to the hypothesis that \( y_{ro} \) is positive for at least one \( r \).

Suppose, say, that \( \lambda_k > 0 \), then for every \( i \)

\[\theta x_{i0} \geq \sum_j x_{ij}\lambda_j \geq x_{ik}\lambda_k \]

Since, by hypothesis for at least one \( i \) we have \( x_{ik} > 0 \) we conclude \( \theta x_{i0} > 0 \) and hence \( \theta > 0 \). In particular, \( \theta^* > 0 \).
By duality, the dual problem will also have one or more optimal solutions all with \( 0 < z^* = z \leq 1 \). This fact and the condition \( g_o(v) = 1 \) together show that every optimal solution \( w^* \) of the dual is D-proper.

We associate a slack with each inequality in the primal or the dual:

\[
(45) \quad s_i = x_{10} - \sum_j x_{ij} \geq 0 \quad (i = 1, \ldots, m)
\]

\[
(46) \quad s_r^* = -y_{ro} + \sum_j y_{rj} \geq 0 \quad (r = 1, \ldots, s)
\]

\[
(47) \quad t_j = -\sum_r y_{rj}u_r + \sum_i x_{ij}v_i \geq 0 \quad (j = 1, \ldots, n)
\]

By the complementary slackness theorem for dual programs we have for any optimal solutions \( \lambda \) and \( w \) of the primal and dual, respectively,

\[
(48) \quad s_i v_i = s_r^* u_r = t_j \lambda_j = 0 \quad (i = 1, \ldots, m; r = 1, \ldots, s; j = 1, \ldots, n).
\]

If the optimal values \( \theta^* = z^* \) are less than 1 then by (47) the dual slack \( t_o \) is positive so that \( \lambda_o = 0 \) for all optimal \( \lambda \). However, if \( \theta^* = z^* = 1 \) then \( \lambda^o \) is one optimal solution for the primal, and for it the primal slacks are all zero.

A less well-known feature of duality known as the strong complementary slackness theorem states that there exists an optimal solution pair \((\lambda^o, w^o)\) for which, in addition to (48), we have

\[
(49) \quad s_i^o v_i > 0 \quad (i = 1, \ldots, m)
\]

\[
(50) \quad \lambda_j^o + \sum_j \lambda_j x_j + s^o -
\]

See Section 13 below for discussion and references.
\[ y_o = \sum_j \lambda^O_j y_j - s^{O+} \]
\[ p_o = \sum_j \lambda^O_j p_j - s^{+o} = \sum_j \lambda^O_j \left[ y_j \right] - \left[ s^{O+} \right] = p_o^0 - s^{+o} \]

where \( s^* = \frac{s^+}{s^-} \) is the non-negative vector of all primal slacks, and \( p_o^0 \) defines a DMU in the conical hull \( K(P) \) of \( D \).

If \( s^{+o} \) is non-zero then (52) states that there is a DMU in \( K(P) \) with the same or less input and more output. Similarly, if \( s^{-o} \) is non-zero, then (52) states that there is a DMU in \( K(P) \) with the same or more output and less input.

In either case it is not logical to regard DMU \( o \) as being fully efficient. Hence, if \( s^{+o} \) is non-zero we do not call DMU \( o \) DEA-efficient even though it is DEA-scale-efficient. We use the term technical efficiency to indicate \( s^{+o} = 0 \), and use the term DEA-efficient to describe a DMU for which \( z^o = 1 \) and \( s^{+o} = 0 \), i.e., one that is both scale efficient and technically efficient.

Note that we do not preclude the consideration of stronger forms of efficiency than DEA-efficiency. Thus, whereas to be DEA-inefficient is a mark against any DMU, to be DEA-efficient need not be enough to settle efficiency in some contexts. We do not treat the question of general efficiency in this paper.

Suppose that DMU \( o \) is DEA-efficient. Then from (49) we conclude that every component of \( w^o \) is positive. Conversely, if \( w_o \) contains any positive vector \( w \) then from (48) applied to \( \lambda^o \) and \( w \) we conclude that \( s^{+o} = 0 \) and hence DMU \( o \) is DEA-efficient.
7. THE PRINCIPAL CLASSIFICATION THEOREMS.

We begin with a few more definitions. For any \( j \) we define a set \( D(j) \) called the dominant of \( \text{DMU}_j \) by

\[
D(j) = \{ \text{DMU}_k | W_j \subset W_k \}
\]

and then also introduce

\[
E(j) = D(j) \cap E
\]

Note, in particular, that if \( \text{DMU}_j \in N \) then \( D(j) = J \) and \( E(j) = E \).

Next, for any virtual multiplier vector \( w \in W \) we define another subset of \( J \) by

\[
J(w) = \{ \text{DMU}_j | w \in W_j \}
\]

Referring again to the dual optimal vectors \( \lambda^0, \nu^0 \), of the previous section we define partitions \( J^+, J^0; I^+, I^0; R^+, R^0 \) of \( J, I = \{1, \ldots, m\}, R = \{1, \ldots, s\} \), respectively, by

\[
J^+ = \{ \text{DMU}_j | \lambda^0_j > 0 \}, \quad J^0 = \{ \text{DMU}_j | \lambda^0_j = 0 \}
\]

\[
I^+ = \{ i | v^0_i > 0 \}, \quad I^0 = \{ i | v^0_i = 0 \}
\]

\[
R^+ = \{ r | u^0_r > 0 \}, \quad R^0 = \{ r | u^0_r = 0 \}.
\]

Then from (51) through (55) we get

\[
t^0_j = - \sum_r y_{jr} u^0_r + \sum_i x^0_{ij} v^0_i \quad \begin{cases} > 0 & \text{for } \text{DMU}_j \in J^0 \\ = 0 & \text{for } \text{DMU}_j \in J^+ \end{cases}
\]

\[
s^0_i = x^0_{io} - \sum_j x^0_{ij} \lambda^0_j \quad \begin{cases} > 0 & \text{for } i \in I^0 \\ = 0 & \text{for } i \in I^+ \end{cases}
\]

\[
s^{0+}_r = -y_{ro} + \sum_j y_{jr} \lambda^0_j \quad \begin{cases} > 0 & \text{for } r \in R^+ \\ = 0 & \text{for } r \in R^0 \end{cases}
\]

For any subsets \( J', I', R' \), of \( J, I, R \), respectively, we set

\[
\lambda(J') = \{ \lambda | \lambda_j = 0 \text{ if } \text{DMU}_j \in J' \}
\]

\[
W(R', I') = \{ w = (u, v) \in W \left| u_r = 0, \quad r \in R' \right. \}
\]

\[
P(R', I') = \{ p = \begin{bmatrix} y \\ -x \end{bmatrix} \left| y_r = 0, r \in R' \right. \}
\]

\[
x_i = 0, \quad i \in J'
\]
We will say that a vector \( w \) is maximally positive in \( W(R', I') \) if 
\( w \in W(R', I') \) and 
\[(63) \ u_r > 0, \ v_i > 0 \text{ for all } r \in R', \ i \in I',\]
i.e., all of the components of \( w \) which are not required to be zero are 
at actually positive. Note that in view of (49) the slack vector \( s^+0 \) in (52) 
is maximally positive in \( W(R^+, I^+) \).

Theorem 4. Primal Representation Theorem.

Suppose that \( DMU_o \in RE \). Then we can write 
\[(64) \ p = \sum_{DMU_j \in E(o)} \alpha_j p_j - s^+0 \]
where \( \alpha_j \geq 0 \) for all \( j \) and \( s^+0 \) is a slack vector which is maximally positive 
in \( W(R^+, I^+) \).

Neither the coefficients \( \alpha_j \) nor the slack vector \( s^+0 \) are necessarily unique.

Theorem 5. Dual Representation Theorem.

Suppose that \( DMU_o \in RE \). Then 
\[(65) \ \widehat{W}_o = \bigcap_{DMU_j \in E(o)} W_j \cap W(R^0, I^0)\]

Theorem 6. \( DMU_o \) belongs to \( E \) if and only if any one of the following
properties holds:

- (E1) \( \dim W_o = s + m \),
- (E2) for some \( w \in W_o \), \( J(w) = \{DMU_o\} \),
- (E3) \( \dim W_o \cap W_j = s + m \) iff \( j = 0 \),
- (E4) the primal program (of Table 1) has the unique optimal solution
  \( \lambda_o = 1, \lambda_j = 0 \) for all \( j \neq 0 \).
- (E5) \( D(o) = \{DMU_o\} \)
- (E6) \( J^+(o) = \{DMU_o\} \)

Theorem 7. \( DMU_o \) belongs to \( E' \) if and only if any one of the following
properties holds:
Theorem 8. DMU\(_o\) belongs to F if and only if any one of the following conditions holds:

(F1) W\(_o\) is non empty but contains no positive vector.

(F2) W\(_o\) is non empty and w\(^o\) has at least one zero component.

(F3) W\(_o\) is non empty and s\(^o\)(o) \(\neq 0\) (in Theorem 4)

(F4) W\(_o\) is non empty and s\(^o\) \(\neq 0\) (in (52)).

We will prove these theorems in the next three sections.

8. PROOFS OF THE PRINCIPAL CLASSIFICATION THEOREMS (PART 1).

In this section we proceed under the hypothesis that DMU\(_o\) is scale efficient. Hence from (52) and (56) we have

\[ P_o = \sum_{DMU_j \in J^+} \lambda_j^o p_j - s_j^o \]

where \( \lambda_j^o \) > 0 and s\(^o\) is maximally positive in W(R\(^+\),I\(^+\)).

Lemma 9. J\(^+\) = D(o) = J(w\(^o\)).

Proof: suppose that \( \lambda_k^o = 0 \), then t\(_k^o\) > 0 and from (57) we conclude that \( \lambda^o \in W_k \) and hence DMU\(_k\) \( \in D(o) \). Hence J\(^+\) \( \supseteq \) D(o).

On the other hand, if DMU\(_k\) \( \in J^+ \) then \( \lambda_k^o > 0 \). Now for w \( \in w^m\) we get from (66) that

\[ 0 - w^T_p = \sum_{DMU_j \in J^+} \lambda_j^o w^T p_j - w^T s_j^o. \]
From (60) with \( \lambda = \lambda_0 \) we get \( w_T s^{+0} = 0 \). Next, since in each product \( \lambda_j^0 w_T p_j \) the first factor is positive and the second is non-negative, for the sum to be zero each term \( w_T p_j = 0 \) and therefore \( D(o) \supset J^+ \).

Finally, \( DMU_j \in J(w^0) \) if and only if \( t_j = 0 \) and hence, by (49), if and only if \( \lambda_j^0 > 0 \). Thus (cf (56)) \( J^+ = J(w^0) \).

The proof of the next lemma utilizes a well-known property of analytic functions which when applied to two quotients \( h_o \) and \( h_j \) of homogeneous linear functions, of \( (s + m) \) variables states that equality of \( h_o \) and \( h_j \) over an \( (s + m) \) sphere implies that there is a positive scalar \( \alpha \) for which the numerator and denominator of \( h_j \) are respectively just \( \alpha \) times the numerator and denominator of \( h_o \).

**Lemma 10.** Let \( DMU_0 \in E \). Then \( \dim W_0 \cap W_j = s + m \) iff \( j = o \).

**Proof:** The "if" follows immediately from the definition of \( E \). For the "only if" we have \( h_o(w) = u_T y_o / v_T x_o \), \( h_j(w) = u_T y_j / v_T x_j \). Now, if \( \dim W_0 \cap W_j = s + m \) then \( W_0 \cap W_j \) contains an \( (s + m) \) sphere on which \( h_o = h_j \). It follows from the cited property of analytic functions that for some positive scalar \( \alpha \) we have \( y_j = \alpha y_o \) and \( x_j = \alpha x_o \) and hence \( p_j = \alpha p_o \). Since our data domain \( D \) is reduced this requires \( \alpha = 1 \) and \( j = o \).

**Lemma 11.** \( D(o) = \{ DMU_o \} \) iff \( DMU_0 \in E \)

**Proof:** If \( \{ DMU_o \} = D(o) = J(w^0) \) then \( W_0 \) contains all elements of \( W \) sufficiently close to \( w^0 \) and hence has dimension \( s + m \) so that \( DMU_o \in E \).

Conversely, if \( DMU_o \in E \) then for any \( w \in W_0 \) we have from (52) and the definition of \( W_0 \) that

\[
0 = w_T p_o = \sum_j \lambda_j^0 (w_T p_j) = w_T s^{+0}.
\]

From (48) we have \( w_T s^{+0} = 0 \) (whether or not \( DMU_o \in E \)). From (44) we have \( w_T p_j \leq 0 \) and hence \( \lambda_j^0 > 0 \) only if \( w_T p_j = 0 \). Thus if \( \lambda_j^0 > 0 \) we have
\( W_o \subseteq W_j \) and hence \( \dim W_o \cap W_j = \dim W_o = s + m \). Now by Lemma 10, we must have \( j = o \).

**Lemma 12.** If \( \text{DMU}_o \) does not belong to \( E \), then \( \lambda_o^o < 1 \) and, hence,

\[
(68) \quad P_o \ifrac{\sum_{\text{DMU}_j \in J^+ - \{\text{DMU}_o\}} \rho_j P_j - s^+}{J^+ - \{\text{DMU}_o\}}
\]

where \( \rho_j = \lambda_j / (1 - \lambda_o)^{-1} \) and where

\[
(69) \quad s^+ = (1 - \lambda_o)^{-1} s^o
\]

is maximally positive in \( W(R^+, I^+) \)

Proof: from (66) we have

\[
(70) \quad (1 - \lambda_o^o) x_o = \sum_{\text{DMU}_j \in J^+ - \{\text{DMU}_o\}} \lambda_j^o x_j - s^o.
\]

Since \( \text{DMU}_o \notin E \), \( J^+ - \{\text{DMU}_o\} \) is non-empty (by Lemma 11). Now, since each \( \lambda_j^o \) in the summation is positive, since each \( x_j \) has no negative and at least one positive component, and since \( R^o \geq 0 \) the right hand side of (70) is a non-negative non-zero vector. Since \( x_o \geq 0 \) this requires \( 1 - \lambda_o^o > 0 \).

**Lemma 13.** Let \( \lambda^* \) be defined by \( \lambda^*_o = 1, \lambda^*_j = 0 \) for all \( j \neq o \).

If \( \text{DMU}_o \in E \) then \( \lambda^* = \lambda^o \) and is the unique optimal solution of the primal problem. Moreover, \( s^+ = 0 \) and \( R^+ = R, I^+ = I \).

In the proof of Lemma 11 we established that \( \lambda^*_j = 0 \) for all \( j \neq o \). Thus (48) reduces to

\[
(71) \quad P_o = i_o P_o - s^o
\]

so that \( \lambda_o^o = 1 \) and \( s^+ = 0 \). Next, from (58) and (59) we conclude that \( R^+ = R \) and \( I^+ = I \) (and, of course, that both \( R^o \) and \( I^o \) are empty).

At this stage we have established Theorems 4 and 5 for \( \text{DMU}_o \in E \).

For, \( E(o) = D(o) = \{\text{DMU}_o\} \) and \( W(R^o, I^o) = W, \) (71) reduces to (64), and (65) reduces to \( W_o = W_o \cap W \).

**Lemma 14.** If the primal system has a unique optimal solution then that solution is the \( \lambda^* \) described in Lemma 13 and \( \text{DMU}_o \in E \).
Proof: Since $\lambda^*$ is an optimal solution whenever $\text{DMU}_o \in \text{RE}$ the unique optimal solution must be $\lambda^*$. Then $\lambda^o = \lambda^*$ and $D(o) = J^+ = \{\text{DMU}_o\}$. Now by Lemma 11 we conclude that $\text{DMU}_o \in E$.

The general framework of the proof of Theorem 4 is to begin with (68) and set up a step-by-step reduction of the summation range from $J^+ - \text{DMU}_o$ to $E(o)$. This would be trivial if $W_j$ properly contained $W_o$ for every $\text{DMU}_j$ in $J^+ - \{\text{DMU}_o\}$. The following Example 3.4 (which is derived from the earlier Example 3) shows that we can have $W_j = W_o$ for $j \neq 0$ and $\text{DMU}_j \in J^+$ and thus illustrates the need for the constructions introduced below when we pick up the proof of Theorem 4.

\[
\begin{array}{cccccc}
\text{DMU} & 1 & 2 & 3 & 4 & 5 & 6 \\
\hline
y_{1j} & 1 & 1 & 1 & 1 & 1 & 1 \\
x_{1j} & 1 & 1 & 1 & 1 & 1 & 1 \\
x_{2j} & 5 & 6 & 5.2 & 5.4 & 5.6 & 5.8 \\
x_{3j} & 6 & 5 & 5.8 & 5.6 & 5.4 & 5.2 \\
\end{array}
\]

Example 3.4. Expansion of Example 3. $s = 1$, $m = 3$, $n = 6$.

We have

\[
W_1 = \{\text{all } w \text{ with } v_3 \leq v_2\}
\]

\[
W_2 = \{\text{all } w \text{ with } v_2 \leq v_3\}
\]

\[
W_3 = W_4 = W_5 = W_6 = \{\text{all } w \text{ with } v_2 = v_3\} = W_1 \cap W_2
\]

\[
\dim W_1 = \dim W_2 = 4 = s + m
\]

\[
\dim W_3 = \dim W_4, \dim W_5 = \dim W_6 = 3 < s + m
\]

Therefore, $E = \{\text{DMU}_1, \text{DMU}_2\}, E' = \{\text{DMU}_3, \text{DMU}_4, \text{DMU}_5, \text{DMU}_6\}$. Now let $o = 4$; we
can verify that $\lambda^0 = \frac{1}{60}(13, 7, 5, 5, 5, 5)$ and $\omega^0 = \begin{bmatrix} u^0 \\ v^0 \end{bmatrix} = \frac{1}{12} \begin{bmatrix} 12 \\ 1 \\ 1 \end{bmatrix}$ are optimal primal and dual solutions that satisfy the strong complementary slackness property. Next, the equation

\[(64') \quad P_4 = (0.6)P_1 + (0.4)P_2 - 0\]

meets the conditions of Theorem 4 since $E(o) = \{DMU_1, DMU_2\} = E$ and $R(o) = 0$ is maximally positive in $W(R^+, I^+)$ since $R^+ = R$ and $I^+ = I$.

On the other hand, $J^+ = J$ and equation (66) becomes

\[(66') \quad P_o = \frac{1}{60}(13P_1 + 7P_2 + 5P_3 + 5P_4 + 5P_5 + 5P_6) = 0\]

so that (67) becomes

\[(67') \quad P_o = \frac{1}{55}(13P_1 + 7P_2 + 5P_3 + 5P_4 + 5P_5 + 5P_6) = 0.\]

What then is our logical path from (66') to (64')? One possible approach is to show that $P_o$ can be written in a form such as (67') but where for each $P_j$ on the right hand side $W_j$ properly contains $W_{o j}$, i.e., in the present case in terms of $P_1$ and $P_2$. We turn to this task in the following section.

9. PROOFS OF THE PRINCIPAL CLASSIFICATION THEOREMS (PART 2).

We begin by recalling the expression

\[(68) \quad P_o = \sum_{DMU_j \in J^- \cdot DMU_o} p_j P_j = -s^+\]

where $s^+$ is maximally positive in $W(R^+, I^+)$ (Recall that $DMU_o \in RE - E$.) Let $T = \{DMU_j | j \in J - \{DMU_o\}$ and $W_j = W_{o j}$.

**Lemma 15.** Let $J^* = J - T$, let $P^*$ be obtained from $P$ by deleting all columns $P_j$ for which $DMU_j \in T$, and let $o^* = o$. Then $W_{o^*} = W_{o}$.  

Proof: If $T$ is empty there is nothing to prove. Suppose that $T$ is non-empty (it follows from Lemma 10 that this cannot happen if $DMU_o \in E$). Since every condition for membership in $W_{o^*}$ holds also for $W_{o}$ (cf. (17)), $W_{o} \subset W_{o^*}$.  

30
Suppose then that \( w' \in W_o^* \) and \( w' \notin W_o \). By the definition of \( w' \) we have

\[
(72) \quad h_j(w') \leq h_o(w') \quad \text{for all } DMU_j \in J^*.
\]

Since \( w' \notin W_o \) there is some \( DMU_k \in T \) for which \( h_o(w') < h_k(w') \), and if there are several such \( DMU_k \) we choose one for which

\[
(73) \quad h_j(w') \leq h_k(w') \quad \text{for all } DMU_j \in T.
\]

Now, since \( J = J^* \cup T \), we conclude from (72) and (73) that \( w' \in W_k \) contrary to our hypothesis that

\[ W_k = W_o. \]

**Lemma 16.** For any \( DMU_o \in R^E - E \) there is an expression of the form

\[
(74) \quad P_o = \sum_{j \in J^* - T - \{DMU_o\}} p_j P_j - s^{+}\]

where \( p_j \geq 0 \) and \( s^{+}\) is maximally positive in \( W(R^*, I^*) \), and for each \( p_j \) appearing on the right hand side \( W_j \) properly contains \( P_o \).

**Proof:** By the definition of \( T \) for each \( DMU_j \in J^* - T - \{DMU_o\} \) we do have \( W_j \) properly containing \( W_o \) and hence also \( W_j \) properly contains \( W_o^* = W_o \). Next applying our general theory to \( J^* \) we get an expression of the form (74) but cannot assume that the slack vector \( s^{+}\) is maximally positive in \( W(R^*, I^*) \).

However, since \( w_o \in W_o^* \) we do have that \( s^{+}\) is maximally positive in \( W(R^*, I^*) \). Moreover such an expression holds for each \( DMU_k \in T \), i.e.,

\[
(75) \quad P_k = \sum_{DMU_j \in J^* - T - \{DMU_o\}} p_{jk} P_j - s^{+}, \quad DMU_k \in T.
\]

Next substitute (75) in (67) for each \( k \in T \) and we get

\[
(76) \quad P_o = \sum_{DMU_j \in J^* - T - \{DMU_o\}} p_{j}^* P_j - s^{+}\]

where

\[
\rho_j^* = \rho_j + \sum_{DMU_k \in T} \rho_{jk} \geq 0 \quad \text{for all } DMU_j \in J^* - T - \{DMU_o\}
\]

and where
belongs to \( W(R^+,I^+) \) since each summand does and is maximally positive in \( W(R^+,I^+) \) since \( s_k^+ \) is.

**Lemma 17.** For any scale efficient \( \text{DMU}_o \) we have

\[
N^+ = \{ \text{DMU}_k \mid W_k \supseteq W_o \}
\]

and for and \( \text{DMU}_k \subseteq J^+ \) we have

\[
J(k)^+ \subseteq J^+, R(k)^+ \supseteq R^+, I(k)^+ \supseteq I^+
\]

and

\[
W(R(k)^+, I(k)^+) \subseteq W(R^+, I^+).
\]

**Proof:** If \( \text{DMU}_k \in J^0 \) then we have \( \sum y_{rk}u_r^0 - \sum x_{ik}v_i^0 < 0 \)
so that \( W_o \) is not a subset of \( W_k \). On the other hand, if \( w \in W_o \) and \( \text{DMU}_k \in J^+ \), then from \( \lambda_k^0 > 0 \) we conclude (via complementary slackness of \( \lambda_k^0 \) and \( w \)) that

\[
\sum y_{rk}u_r - \sum x_{ik}v_i = 0;
\]

hence \( w \in W_k \) and consequently \( W_o \subseteq W_k \).

Next, suppose that \( W_k \supseteq W_o \). Then, if \( W_j \supseteq W_k \), we conclude that \( W_j \supseteq W_o \); hence \( J(k)^+ \subseteq J^+ \). Moreover, if \( r \in R(k)^0, i \in I(k)^0 \) then for any \( w \in W_k \) we have \( u_r = v_i = 0 \). In particular, \( w^0 \in W_k \) so that \( u_r^0 = v_i^0 = 0 \) and hence \( R(k)^0 \supseteq R^0, I(k)^0 \supseteq I^0 \). It follows from (56) that \( R(k)^+ \subseteq R^+, I(k)^+ \subseteq I^+ \) and hence also that \( W(R(k)^+, I(k)^+) \subseteq W(R^+, I^+) \).

We are now ready to complete the proof of Theorem 4.

Suppose that \( \text{DMU}_o \in R^+ - E \) and we have an expression for \( P_o \) of the form

\[
(77) \quad P_o = \sum_{\text{DMU}_j \in K} \lambda_{kj} p_j - s_k^+ \text{ where }
\]

\[
(78a) \quad E(o) \subseteq K \subseteq J^+ - T - \{ \text{DMU}_o \}
\]

\[
(78b) \quad \lambda_{kj} > 0 \text{ for all } \text{DMU}_j \in K \text{ and }
\]

\[
(78c) \quad s_k^+ \text{ is maximally positive in } W(R^+, I^+).
\]
Next, suppose \( W_k \supset W_j \) for all \( DMU_j \in K \) and let
\[
T(k) = \{ DMU_j | W_j = W_k \text{ and } DMU_j \in K - \{ DMU_k \} \}.
\]

Now apply Lemma 16 for all \( o = j \) where \( DMU_j \subset T(k) \cup \{ DMU_k \} \), substitute the resulting expressions on the right side of (77) and we will get an expression of the form (77) where the summation is over a set \( K' \) such that (78a,b,c) all hold and for every \( DMU_j \in K' \), \( W_j \) properly contains \( W_k \). This process can be iterated until we reach a \( K^0 \) which is equal to \( E(o) \). This completes the proof of Theorem 4.

Proof of Theorem 5: Since \( E(o) = D(o) \cap E \) and \( D(o) = \{ DMU_k | W_0 \subset W_k \} \) we know that \( W_0 \subset DMU_j \subset E(o) \). Moreover, since for each \( w \in W_0^n \) we have from (56) with \( \lambda = \lambda^o \) that \( w \in W(R^0,I^0) \). Therefore
\[
W_0 \subset DMU_j \subset E(o) \cap W_j \cap W(R^0,I^0).
\]

Next, suppose that \( P_o \) satisfies (64) and that \( w \) belongs to the right hand side of (65). By the definition of \( E(o) \), all \( h_j(w) \) for \( DMU_j \in E(o) \) are equal. Let the common value be \( \beta \). Then, \( w' = \left[ \begin{array}{c} u \\ \beta v \end{array} \right] \) also belongs to the right hand side of (65) and for it
\[
h_j(w') = 1 \leq h_k(w') \text{ for all } DMU_j \in E(o), DMU_k \in J.
\]

Multiply (64) by \( w'^T \) and we get
\[
u_T y_j - \beta v^T x_j = w'^T P_o - \sum_{DMU_j \in E(o)} a_j w'^T p_j - w'^T g_j^0 = \sum a_j 0 - 0 - 0.
\]

Therefore, \( h_o(w') = 1 \) and so \( w' \) and hence also \( w \) is in \( W_0 \). This shows that
\[
W_0 \supset \bigcap_{DMU_j \in E(o)} W_j \cap W(R^0,I^0)
\]
and completes the proof of Theorem 5.
10. PROOFS OF THE PRINCIPAL CLASSIFICATION THEOREMS (PART 3)

Proof of Theorem 6: The lemmas in Section 8 take care of most of theorem 6.

(E1) is the definition of E

(E2) follows from Lemmas 9 and 11

(E3) is Lemma 10

(E4) implies (E6) and therefore (E1). Lemma 13 implies (E4)

(E5) is Lemma 11

(E6) is equivalent to (E5) by Lemma 9.

Proof of Theorem 7:

(E'1) is the definition of E'

(E'2) follows from Theorem 5

(E'3) If W_o is non-empty and s^0 = 0, then since s^0 is maximally positive in W(R^+, I^+) we must have R^+ = R, I^+ = I and hence w^0 > 0.

Therefore, (E'3) implies (E'1). The reverse implication is obvious from (49).

(E'4) A similar argument to the one for (E'3) holds for (E'4).

(E'5) If W_o contains any positive vector w then w^0 is positive and s^0 = 0 so that (E'5) implies (E'3). Conversely, if s^0 = 0 then w^0 is positive so that (E'3) implies (E'5).

Proof of Theorem 8:

(F1) is the definition of F.

(F2) If w^o has a zero component then s^0 has a positive component and therefore no w e W_o can be positive. Hence (F2) implies (F1). The converse is trivial.

(F3) and (F4) both imply that W_o can contain no positive vector and hence that (F1) holds. Next, if w^o is not positive then s^0 ≠ 0 so that (F1)
implies (F4). Finally, if $w^*$ is not positive then we cannot have both $R^+ = \mathbb{R}$ and $I^+ = \mathbb{I}$ so that $s^+(o) \neq 0$. Therefore (F1) implies (F3).

11. CLASSES OF NON SCALE EFFICIENT DMUs

Suppose that $DMU_o \in N$ and that $\lambda^*, w^*$ are optimal dual vectors for $DMU_o$. Then consider a new $DMU'_o$ defined by

$$P'_o = \begin{bmatrix} y^* \\ -s^* X_o \end{bmatrix}.$$  

(It may indeed happen that for some $j$ and $a > 0$ we have $P'_o = aP_j$. This makes no essential difference in the discussion which follows.)

From the primal equations we have

$$(79) \quad P'_o = \sum_{j=1}^{n} \lambda_j^* P_j - s^* \leq \sum_{j=1}^{n} \lambda_j^* P_j.$$  

Since $DMU_o \in N$, $t^*_o > 0$ and hence $\lambda^*_o = 0$, therefore (79) shows that $P'_o \in K(D)$ (the conical hull of D). Moreover, (79) shows that $DMU'_o$ is scale efficient in the domain $D^+$ (for $p = P'_o$) and hence, that $DMU'_o$ is in the frontier $FR$ of $D$, and therefore belongs to (exactly) one of the classes $E, E', F$ in $D^+$. Based on this we partition $N$ into three classes as follows:

$$(80) \quad DMU_o \in NE \text{ or } NE' \text{ or } NF \text{ according as } DMU_o \in E \text{ or } E' \text{ or } F \text{ in } D^+.$$  

We now have a partition of the DMUs into six classes.

For randomly generated data we might conjecture that the classes $E', F$ and $IE$ are very small as contrasted to $E, IE'$ and $I$. It would be interesting to check this conjecture for data domains which have been studied.

(The 15 hospital case in [6] was contrived to enhance the likelihood that $F$ would not be empty, but what about $E'$ and $IE$?)
12. THE ROLE OF ZERO INPUTS AND OUTPUTS

We have established the classification theorems for DMUs without the earlier DEA assumptions that all \( x_{ij} \) and \( y_{-j} \) were non-zero. However, further discussion of the influence of zero inputs and outputs is warranted. In particular, it follows from the universal feasibility of the programs in Table 1 that no \( DMU_j \) can become scale efficient only because of some \( w \) with \( \sum x_{ij} v_j = 0 \).

In discussing zero inputs it is convenient to introduce subsets \( I_1, \ldots, I_n \) of \( I \) by the definition

\[
I_j = \{ i | x_{ij} = 0 \}, \quad j = 1, \ldots, n.
\]

Theorem 18. Let \( I' \) be a non-empty subset of \( I \) for which

\[
I' \subseteq I_j \iff j \in \{1, \ldots, k\}
\]

and let \( D^* \) be the reduced data domain consisting of \( DMU_1, \ldots, DMU_k \). Then if \( o \in \{1, \ldots, k\} \) the category of \( DMU_o \) is the same in \( D^* \) as in \( D \).

Consider the primal equation

\[
\delta x_{io} = \sum_j \lambda_j x_{ij} + s_i^-. 
\]

If \( i \in I' \) then \( x_{io} = 0 \) so that \( s_i^- = 0 \) and for each \( j \) either \( \lambda_j = 0 \) or \( x_{ij} = 0 \). Consider some \( j > k \). Then by (82) there will be some \( i \in I' \) for which \( x_{ij} \neq 0 \) and, therefore, \( \lambda_j = 0 \) (\( j = k + 1, \ldots, n \)). This means that the primal system for \( DMU_o \) in \( D \) reduces to that in \( D^* \). Now, since by Theorems 6, 7, 8 and the arguments of Section 11 the category of \( DMU_o \) is determined by the primal system, the theorem is established.

The next corollaries all follow in turn from Theorem 18.

Corollary 19. There is at least one \( j \neq k \) for which \( DMU_j \) belongs to \( E \).

Corollary 20. Suppose that for some \( DMU_o \) the intersections \( I_o \cap I_j \) are proper subsets of \( I_o \) for all \( j \neq j \). Then \( DMU_o \) belongs to \( E \).
Corollary 21. Suppose that for some DMU and some \( i \) we have \( x_{10} = 0 \) and \( x_{ij} \neq 0 \) for all \( j \neq 0 \). Then DMU belongs to \( E \).

Theorem 18 and its corollaries shed light on the importance and roles of zeros in the input coefficients \( x_{ij} \). Any zero input coefficient will lead the way to an element of \( E \).

A tempting early thought in developing DEA would be to assign the value "0" to any formal quotient \( f_j(u)/g_j(v) \) with numerator positive and denominator zero and then to call DMU scale efficient. The example

\[
(34) \quad Y = [1 \ 1 \ 1], \quad X = \begin{bmatrix} 0 & 0 & 1 \\ 1 & 2 & 0 \end{bmatrix}
\]

with \( n = 3 \) shows the fallacy of this approach. The virtual multiplier \( w \) with \( u_1 = 1, v_1 = 1, v_2 = 0 \) gives both \( h_1(w) \) and \( h_2(w) \) the form \(+/0\). However, from the equations

\[
y_2 = y_1 + 0y_2 + 0y_3 \\
\frac{1}{2} x_2 = x_1 + 0x_2 + 0x_3
\]

We see that for "0" = "2", the optimal \( \theta^* \) is less than or equal to \( \frac{1}{2} \) and hence DMU is not scale efficient. (For this example \( E = \{DMU_1, DMU_3\} \)).

One might ask why we couldn't avoid the necessity for considering D-improper multipliers \( w \) of the type \( f_j(u) = 0 \) for all \( j \) by deleting any output \( r \) for which all \( y_{rj} = 0 \), i.e., forbid zero rows in \( Y \)? One reason is that in applying Theorem 18 we might obtain a reduced domain \( D^* \) for which both \( X^* \) and \( Y^* \) had zero rows even though this was not true for the \( X \) and \( Y \) of the parent domain \( D \). The following example illustrates this point. Let \( D \) be defined by a matrix of the form

\[
\begin{bmatrix} Y \\ -X \end{bmatrix} = \begin{bmatrix} Y^1 \\ 0 \\ -X^1 \\ 0 \end{bmatrix} \begin{bmatrix} 0^2 \\ 0^2 \end{bmatrix}
\]
where $Y^1$, $Y^2$, $X^1$, $X^2$ are matrices of appropriate degrees and without zero entries. Then both

$$D_1^* = \begin{bmatrix} Y^1 \\ 0 \\ -X^1 \\ 0 \end{bmatrix} \quad \text{and} \quad D_2^* = \begin{bmatrix} 0 \\ Y^2 \\ 0 \\ -X^2 \end{bmatrix}$$

can arise in applying Theorem 18. We arrive at $D$ by combining the two data domains $D_1$ and $D_2$ defined by

$$D_1 = \begin{bmatrix} Y^1 \\ -X^1 \end{bmatrix}, \quad D_2 = \begin{bmatrix} Y^2 \\ -X^2 \end{bmatrix}$$

and which have no common outputs or inputs.

The following example is more realistic

$$\begin{bmatrix} Y^1 & 0 & Y^{1,3} \\ 0 & Y^2 & Y^{2,3} \\ -X^1 & 0 & -X^{1,3} \\ 0 & -X^2 & -X^{2,3} \\ -X^{*,1} & -X^{*,2} & -X^{*,3} \end{bmatrix}$$

and illustrates how permitting zero input and output coefficients extends the theory to situations where the DMUs do not have the same active inputs and outputs.

Some idea of the importance of being in a position to treat these zero inputs and/or outputs is apparent by reference to issues in divestiture such as were involved in the recent dissolution of AT&T. See [14]. In such cases, the issue turns on the complete removal of the entity from certain activities so that as in the AT&T case, zero outputs appear for local calls in one part of the proposed new entity and zero outputs appear for the toll call
components for other parts of the new entities arising from the divestitur

13. THEORETICAL APPLICATION OF THE CLASSIFICATION THEOREMS.

We turn now to methods for application of the classification theorems to a reduced data domain $D$. and consider three problems.

Problem 1. Partition the set $J$ of DMU's into the six classes $E, E', F, NE, NE', NF$.

Problem 2. Identify all inclusion relations among the multiplier sets $W_1, \ldots, W_m$.

Problem 3. For each DMU, construct optimal primal and dual vectors $\lambda^j$ and $w^j$ which satisfy the strong complimentary conditions (48) and (49).

We proceed as follows. First, (Phase 1) consider Problems 1 and 2 under the assumption that Problem 3 has been solved. Next, (Phase 2) see how to handle Problems 1 and 2 without having optimal, strongly complimentary solutions available. Then (Phase 3) turn to Problem 3.

Phase 1

Let $\lambda^o$, $w^o$ be a strong complementary optimal pair for a selected DMU, and let $\theta^* = z^* = \sum_r y^o_r \mu^o_r$ be the optimal values for the objective functions (cf. Table 1).

Then it follows from the classification theorem that DMU belongs to:

- $E$ iff $\theta^* = 1$, $s^o_j = 0$, and $t^o_j > 0$ for all $j \neq 0$
- $E'$ iff $\theta^* = 1$, $s^o_j = 0$, and $t^o_j = 0$ for some $j \neq 0$
- $F$ iff $\theta^* = 1$, $s^o_j \neq 0$
- $NE$ iff $\theta^* < 1$, $s^o_j = 0$, and for some DMU $o'$

$\lambda^o_o, \neq 0$ and $\lambda^o_j = 0$ for all $j \neq o'$
NE' iff \( \theta^* < 1 \), \( s^o = 0 \), and more than one \( \lambda_j^o \) is non-zero
NF iff \( \theta^* < 1 \), \( s^o \neq 0 \).

We observe that \( DMU_o \) is in NE iff there exists \( a > 0 \) such that

\[
\begin{bmatrix}
y^o_r \\
-x^o_r
\end{bmatrix} = \begin{bmatrix}
0^o \\
-\alpha x^o_r
\end{bmatrix}
\]

where \( DMU_o \in E \) the stringency of this condition makes it likely that NE will be empty.

We complete Phase 1 by noting that \( W_o \subseteq W_j \) iff \( \lambda_j^o \neq 0 \). For, if \( \lambda_j^o \neq 0 \), \( t_j = -\sum y_j^r u^r + \sum x_{ij}^r v_1 = 0 \) for all optimal virtual multipliers \( w = (u,v) \in W_o^m \) and therefore \( W \subseteq W_j \) (not necessarily to \( W_j^m \)). It follows that \( \lambda_j^o \neq 0 \) implies that \( W_o \subseteq W_j \). Conversely, if \( \lambda_j^o = 0 \) then \( t_j^o \neq 0 \). so that \( w^o \notin W_j \) and hence \( W_o \notin W_j \).

Phase 2.

We begin with any optimal primal, dual pair \( (\lambda^*, w^*) \) for \( DMU_o \). Suppose that \( DMU_o \) belongs to RE, i.e., \( \theta^* = z^* = 1 \). Now consider the modified primal program

\[
\max \sum_{r=1}^{s} s_r^+ + \sum_{i=1}^{m} s_i^-
\]

subject to

\[
\lambda_j \geq 0, s_r^+ \geq 0, s_i^- \geq 0 \quad (j = 1, \ldots, m, r = 1, \ldots, s, i = 1, \ldots, m)
\]

and

\[
P_o = \sum_{j=1}^{m} \lambda_j P_j + s^o.
\]

Since \( \theta^* = 1 \), \( \lambda^* \) is a feasible solution of the modified system. If the maximum is positive \( DMU_o \in F \), otherwise \( DMU_o \in EUE^1 \).
In this case, we again solve a modified form of the primal problem; this time with the objective function \( \min \lambda_o \), but again with \( \theta \) fixed at \( \theta = 1 \). If the minimum value is 1 then DMU\(_o\) belongs to E; otherwise (i.e., if the minimum is less than one) it belongs to E'.

If DMU\(_o\) \( \in N \) (i.e., \( \theta^* < 1 \)) then we apply this same procedure to the related DMU\(_r\), as defined in Section II above to determine which class NE, NE', or NF, contains DMU\(_o\).

One minor point needs to be mentioned. If, perchance, P\(_0\)' is proportional to some P\(_j\) for DMU\(_j\) \( \in RE \) then the data domain \( D \cup DMU\(_o\) \) is not reduced. But, then DMU\(_o\) belongs to IE, IE' or EF according as DMU\(_j\) belongs to E, E', or F.

Problem 2 can be handled similarly. If \( \lambda_j^* \neq 0 \) then \( W_o \subset W_j \).

If \( t_j^* \neq 0 \) then \( W_o \subset W_j \). If \( \lambda_j^* = t_j^* = 0 \) modify the primal program as follows. Fix \( \theta \) at \( \theta^* \) and change the objective function from \( \min \theta \) to \( \max \lambda_j \). Then \( W_o \subset W_j \) iff the optimal value of \( \lambda_j \) is positive. This completes Phase 2.

**Phase 3**

The strong complementary slackness theorem appeared in 1961 in the Charnes - Cooper text [7] pp. 441-443 where it is called the "Extended Theorem of the Alternative." Two years later Dantzig [17] gave the result in its slightly weaker homogenous form. In 1968 Balinski [1] Th. 8 p. 59, gave an elegant proof of the theorem under the title "Full Complementary Slackness." The alternate title used here "Strong Complementary Slackness" was introduced in 1970 in [18] Th. 1, p. 308. Since this theorem has remained relatively unnoticed by other writers we provide a brief discussion of a theoretical (not necessarily practical) computational procedure for obtaining \( \lambda^o, w^o \) using a standard simplex algorithm. We note that,
in general, we cannot expect $\lambda^0$, $w^0$ to be basic solutions. However, we know that every optimal solution is a convex combination of basic solutions and we base our procedure on that property of $\lambda^0$ and $w^0$.

Suppose now that $\text{DMU}_o \subseteq \text{RE}$ and that $\lambda^*$, $w^*$ are basic optimal solutions of the primal, dual systems with corresponding slack vectors

$$s^{*'} = \begin{bmatrix} s_{1}^{*} \\ s_{2}^{*} \\ \vdots \end{bmatrix}, t^*.$$  

Then, since $\theta^* = 1$, any solution $(\lambda, s^*)$ of the modified primal constraint system

$$P_o = P\lambda + s^*, \lambda \geq 0, s^* \geq 0$$

will be optimal for $\text{DMU}_o$.

Similarly, since $z^* = 1$, any solution $(w, t)$ of the modified dual constraint system

$$t + P^Tw = 0, t_o = 0, x^Tv = 1, w \geq 0, t \geq 0$$

will be optimal for $\text{DMU}_o$.

Now suppose for some $i$ that $v_i^* = s_i^* = 0$. Then exactly one of the problems

$$\max s_i \text{ subject to (87)}$$

or

$$\max v_i \text{ subject to (88)}$$

will have a positive objective value; denote a corresponding optimizing vector by $\lambda(i)$, $s^*(i)$ if $s_i^-$ can be positive and by $v(i)$, $t(i)$ if $v_i^-$ can be positive. Let $I^* = \left\{ i | v_i^* = s_i^* = 0, v_i^+ \text{ can be positive} \right\}$ and let $I^{**} = I - I^*$.

Next, let $v(I)$, $t(I)$ be the average of all pairs $v(i)$, $t(i)$ for $i \in I^*$ and let $(I)$, $s(I)$ be the average of all pairs $\lambda(i)$, $s^*(i)$ for $i \in I^{**}$.
Then do the same for \( R \) and \( J \) to get \( v(R), t(R), \lambda(R), s(R) \) and \\
v(J), t(J), \lambda(J), s(J). \) Now, the choice \\

\[(91) \quad (\lambda^0, s^0) = \frac{1}{4} \left[ \left( \lambda^*, s^* \right) + (\lambda(I), s(I)) + (\lambda(R), s(R)) + (\lambda(J), s(J)) \right] \]

and

\[(92) \quad (w^0, t^0) = \frac{1}{4} \left[ (w^*, t^*) + w(I), t(I) + (w(R), t(R)) + (w(J), t(J)) \right] \]

will give optimal primal and dual solutions which satisfy (49), i.e., \\
which satisfy the strong complementary slackness condition.

We have now solved the three problems stated at the opening \\
of this section. These solutions involve repeated use of simplex \\
solutions to the dual problems in Table 1 plus closely related modifi-

\cations. Although each such problem is simple, the number of the 
problems that needs to be addressed is very large. This leads to 
the problem of developing a more practical approach for securing 
numerical results in actual applications.

By an alternate formulation which is slightly more complex, 
one can arrive at an instant characterization of a DMU as an EUE', 
F or N. This new formulation suggests a strategy that may lead 
to more rapid and efficient solution procedures.
In preceding sections, the characterizations of $E$, $E'$ and $F$ are made in terms of the cone $W$ and its cone subsets $W_j$. For evaluation of any particular DMU, designated as $DMU_o$, connection is made to the dual linear programming problems:

\[
\begin{align*}
\min \theta & \quad \max u^T y_o \\
\text{subject to} & \quad \text{subject to} \\
Y \lambda \geq Y_o & \quad u^T y - v^T x \leq 0 \\
\theta x_o - X \lambda \geq 0 & \quad v^T x_o = 1 \\
\lambda \geq 0 & \quad u^T \geq 0, v^T \geq 0
\end{align*}
\]

But $E$, $E'$ are characterized by $W_o$ having a strictly positive vector set of dimension $m = s$ or less than $m + m$, respectively, whereas $F$ has no strictly positive vector. These are not elements which are immediately related to the extreme points in (93) and thereby not to the computational methods e.g., simplex, etc. which rest on extreme points in (93).

Another linear programming problem which is extreme point computable and can distinguish $F$ from $E \cup E'$ is:
Note the maximization problem is feasible for any scale efficient DMU and is not feasible otherwise. Further, any feasible solution must have $v^T x^* = 1$. For if $v^T x^* < 1$, then $u^T y^* = v^T x^* < 1$, and infeasibility. The minimization problem is always feasible, however, e.g., $\psi^+ = 0$, $\psi^- = x^0/v^T x^0$, $\lambda^* = \theta^+$, $\lambda_j = 0$ for $j \neq 0$, $\theta^* = \theta^+ + 1/v^T x^0$ satisfies the constraints.

As in past work by Charnes, Cooper and Seiford in non-linear programming [12], we can distinguish the character of $DMU_o$ by an optimal solution (designated with asterisks) as follows

Theorem 19. $DMU_o$ is $\left\{ E \text{ or } E^* \right\}$ iff $-\theta^{**} + \theta^{**} \text{ is } \begin{cases} > 0, & \text{if } -\theta^{**} + \theta^{**} = 0 \\ = 0, & \text{otherwise} \end{cases}$

Proof: Consider the maximal problem. When consistent, $DMU_o$ is scale efficient. Clearly, $w^* = (u^*, v^*) > 0$ iff $\eta^* = \min (u^*_x, v^*_1) > 0$. But $\eta^* = -\theta^{**} + \theta^{**}$. Hence the first statement.
Next, note that $\eta^*$ is the maximum of the minimal value of the components of a feasible $w$ for scale efficient $DMU_0$. Then $0 = \eta^* - \theta^{**} + \theta^{*-}$ means every $w^*$ has at least one zero component, or, $DMU_0$ is in $F$.

Finally, since the minimal problem is always feasible, by the duality states theorem for linear programming (see p. 486, Charnes, Cooper, Duffuaa, Kress, [12]) the maximal problem is infeasible iff the minimal problem has $-\theta^{**} + \theta^{**} = -\eta^*$.

14. CONCLUDING REMARKS

This completes what was to be done in the present paper. We have now provided a general theoretical basis and characterization of DEA solution sets which makes it possible to handle zero entries in the data and in the variables. We have also provided a basis for developing practical numerical algorithms with which to give these ideas and concepts operational form for use in actual applications. We have indicated some of the widened class of applications that this makes possible for DEA.
REFERENCES


A Structure for Classifying and Characterizing Efficiencies and Inefficiencies in Data Envelopment Analysis

A. Charnes, W.W. Cooper, R.M. Thrall

Center for Cybernetic Studies
The University of Texas at Austin
Austin, TX 78712

Office of Naval Research (Code 434)
Washington, DC

This document has been approved for public release and sale; its distribution is unlimited.

dual linear programs, multicriterion efficiency analysis, scale efficiency, strong complementary slackness, technical efficiency, virtual multipliers

DEA (Data Envelopment Analysis) attempts to identify sources and estimate amounts of inefficiencies contained in the outputs and inputs generated by managed entities called DMUs (= Decision Making Units). Explicit formulation of underlying functional relations with specified parametric forms relating inputs to outputs is not required. An overall (scalar) measure of efficiency is also obtained for each DMU from the observed magnitudes of its multiple inputs and outputs without requiring use of a
priori weights or relative value assumptions. A partition of DMUs into six classes is established via primal and dual representation theorems and three classification theorems which do not depend on non-archimedian analysis. Earlier theory is extended to explain the consequences of zero inputs and outputs and to utilize zero virtual multipliers (shadow prices). Three of the six classes are scale inefficient and two of the three scale efficient classes are also technically (= zero waste) efficient.