3 THE CCR MODEL AND PRODUCTION CORRESPONDENCE

3.1 INTRODUCTION

In this chapter, we relax the *positive* data set assumption. We shall instead assume the data are semipositive. That is, we assume that some (but not all) inputs and outputs are positive. This will allow us to deal with applications which involve zero data in inputs and/or outputs. The *production possibility set* composed of these input and output data (X, Y) will also be introduced. The *dual* problem of the CCR model will then be constructed and it will be shown that the dual problem evaluates efficiency based on a linear programming problem applied to the data set (X, Y). The CCR-efficiency will be redefined, taking into account all input excesses and output shortfalls.

The observations that form the production possibility set are very fundamental in that they make it possible to assess the CCR model from a broader point of view and to extend this model to other models that will be introduced in succeeding chapters.

One version of a CCR model aims to minimize inputs while satisfying at least the given output levels. This is called the *input-oriented* model. There is another type of model called the *output-oriented* model that attempts to maximize outputs without requiring more of any of the observed input values. In the last section, the latter will be introduced along with a combination of the two models.

Computational aspects of the CCR model are also covered in this chapter. The computer code DEA-Solver accompanying this book (as mentioned in Problem 1.1 of Chapter 1) will be utilized on some of the problems we provide.

3.2 PRODUCTION POSSIBILITY SET

We have been dealing with the pairs of positive input and output vectors $(\boldsymbol{x}_j, \boldsymbol{y}_j)$ $(j = 1, \ldots, n)$ of n DMUs. In this chapter, the positive data assumption is relaxed. All data are assumed to be nonnegative but at least one component of every input and output vector is positive. We refer to this as *semipositive* with a mathematical characterization given by $\boldsymbol{x}_j \geq \mathbf{0}, \boldsymbol{x}_j \neq \mathbf{0}$ and $\boldsymbol{y}_j \geq \mathbf{0}, \boldsymbol{y}_j \neq \mathbf{0}$ for some $j = 1, \ldots, n$. Therefore, each DMU is supposed to have at least one positive value in both input and output. We will call a pair of such semipositive input $\boldsymbol{x} \in R^m$ and output $\boldsymbol{y} \in R^s$ *an activity* and express them by the notation $(\boldsymbol{x}, \boldsymbol{y})$. The components of each such vector pair can be regarded as a semipositive orthant point in $(m + s)$ dimensional linear vector space in which the superscript m and s specify the number of dimensions required to express inputs and outputs, respectively. The set of feasible activities is called *the production possibility set* and is denoted by P. We postulate the following

Properties of P (the Production Possibility Set)

(A1) The observed activities $(\boldsymbol{x}_j, \boldsymbol{y}_j)$ $(j = 1, \ldots, n)$ belong to P.

(A2) If an activity $(\boldsymbol{x}, \boldsymbol{y})$ belongs to P, then the activity $(t\boldsymbol{x}, t\boldsymbol{y})$ belongs to P for any positive scalar t. We call this property the *constant returns-to-scale* assumption.

(A3) For an activity $(\boldsymbol{x}, \boldsymbol{y})$ in P, any semipositive activity $(\bar{\boldsymbol{x}}, \bar{\boldsymbol{y}})$ with $\bar{\boldsymbol{x}} \geq \boldsymbol{x}$ and $\bar{\boldsymbol{y}} \leq \boldsymbol{y}$ is included in P. That is, any activity with input no less than $\boldsymbol{x}$ in any component and with output no greater than $\boldsymbol{y}$ in any component is feasible.

(A4) Any semipositive linear combination of activities in P belongs to P.[1]

Arranging the data sets in matrices $X = (\boldsymbol{x}_j)$ and $Y = (\boldsymbol{y}_j)$, we can define the production possibility set P satisfying (A1) through (A4) by

$$P = \{(\boldsymbol{x}, \boldsymbol{y}) \mid \boldsymbol{x} \geq X\boldsymbol{\lambda},\ \boldsymbol{y} \leq Y\boldsymbol{\lambda},\ \boldsymbol{\lambda} \geq \mathbf{0}\}, \tag{3.1}$$

where $\boldsymbol{\lambda}$ is a semipositive vector in R^n.

Figure 3.1 shows a typical production possibility set in two dimensions for the CCR model in the single input and single output case, so that $m = 1$ and $s = 1$, respectively. In this example the possibility set is determined by B and the ray from the origin through B is the efficient frontier.

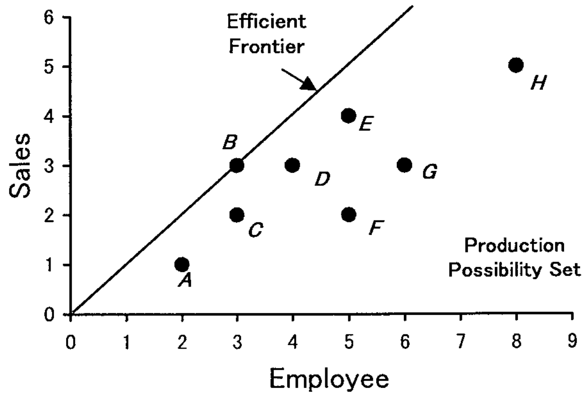

Figure 3.1. Production Possibility Set

3.3 THE CCR MODEL AND DUAL PROBLEM

Based on the matrix (X, Y), the CCR model was formulated in the preceding chapter as an LP problem with row vector $\boldsymbol{v}$ for input multipliers and row vector $\boldsymbol{u}$ as output multipliers. These multipliers are treated as variables in the following LP problem ([Multiplier form]):

$$(LP_o) \quad \max_{\boldsymbol{v},\boldsymbol{u}} \quad \boldsymbol{u}\boldsymbol{y}_o \tag{3.2}$$

$$\text{subject to} \quad \boldsymbol{v}\boldsymbol{x}_o = 1 \tag{3.3}$$

$$-\boldsymbol{v}X + \boldsymbol{u}Y \leq \mathbf{0} \tag{3.4}$$

$$\boldsymbol{v} \geq \mathbf{0}, \quad \boldsymbol{u} \geq \mathbf{0}. \tag{3.5}$$

This is the same as (2.7)-(2.11), (LP_o) in the preceding chapter, which is now expressed in vector-matrix notation.

The dual problem[2] of (LP_o) is expressed with a real variable θ and the transpose, T, of a nonnegative vector $\boldsymbol{\lambda} = (\lambda_1, \ldots, \lambda_n)^T$ of variables as follows ([Envelopment form]):

$$(DLP_o) \quad \min_{\theta,\boldsymbol{\lambda}} \quad \theta \tag{3.6}$$

$$\text{subject to} \quad \theta\boldsymbol{x}_o - X\boldsymbol{\lambda} \geq \mathbf{0} \tag{3.7}$$

$$Y\boldsymbol{\lambda} \geq \boldsymbol{y}_o \tag{3.8}$$

$$\boldsymbol{\lambda} \geq \mathbf{0}. \tag{3.9}$$

Correspondences between the primal (LP_o) and the dual (DLP_o) constraints and variables are displayed in Table 3.1.

(DLP_o) has a feasible solution $\theta = 1$, $\lambda_o = 1$, $\lambda_j = 0$ $(j \neq o)$. Hence the optimal θ, denoted by θ^*, is not greater than 1. On the other hand, due to

Table 3.1. Primal and Dual Correspondences

Multiplier form Constraint (LP_o)	Envelopment form Variable (DLP_o)	Envelopment form Constraint (DLP_o)	Multiplier form Variable (LP_o)
$\boldsymbol{v}\boldsymbol{x}_o = 1$	θ	$\theta \boldsymbol{x}_o - X\boldsymbol{\lambda} \geq \mathbf{0}$	$\boldsymbol{v} \geq \mathbf{0}$
$-\boldsymbol{v}X + \boldsymbol{u}Y \leq \mathbf{0}$	$\boldsymbol{\lambda} \geq \mathbf{0}$	$Y\boldsymbol{\lambda} \geq \boldsymbol{y}_o$	$\boldsymbol{u} \geq \mathbf{0}$

the nonzero (i.e., semipositive) assumption for the data, the constraint (3.8) forces $\boldsymbol{\lambda}$ to be nonzero because $\boldsymbol{y}_o \geq \mathbf{0}$ and $\boldsymbol{y}_o \neq \mathbf{0}$. Hence, from (3.7), θ must be greater than zero. Putting this all together, we have $0 < \theta^* \leq 1$. Now we observe the relation between the production possibility set P and (DLP_o). The constraints of (DLP_o) require the activity $(\theta \boldsymbol{x}_o,\ \boldsymbol{y}_o)$ to belong to P, while the objective seeks the minimum θ that reduces the input vector $\boldsymbol{x}_o$ radially to $\theta \boldsymbol{x}_o$ while remaining in P. In (DLP_o), we are looking for an activity in P that guarantees at least the output level $\boldsymbol{y}_o$ of DMU$_o$ in all components while reducing the input vector $\boldsymbol{x}_o$ proportionally (radially) to a value as small as possible. Under the assumptions of the preceding section, it can be said that $(X\boldsymbol{\lambda},\ Y\boldsymbol{\lambda})$ outperforms $(\theta \boldsymbol{x}_o,\ \boldsymbol{y}_o)$ when $\theta^* < 1$. With regard to this property, we define the input *excesses* $\boldsymbol{s}^- \in R^m$ and the output *shortfalls* $\boldsymbol{s}^+ \in R^s$ and identify them as "slack" vectors by:

$$\boldsymbol{s}^- = \theta \boldsymbol{x}_o - X\boldsymbol{\lambda}, \quad \boldsymbol{s}^+ = Y\boldsymbol{\lambda} - \boldsymbol{y}_o, \tag{3.10}$$

with $\boldsymbol{s}^- \geq \mathbf{0},\ \boldsymbol{s}^+ \geq \mathbf{0}$ for any feasible solution $(\theta,\ \boldsymbol{\lambda})$ of (DLP_o).

To discover the possible input excesses and output shortfalls, we solve the following two-phase LP problem.

Phase I

We solve (DLP_o). Let the optimal objective value be θ^*. By the duality theorem of linear programming,[3] θ^* is equal to the optimal objective value of (LP_o) and is the CCR-efficiency value, also called "Farrell Efficiency," after M.J. Farrell (1957). See below. This value of θ^* is incorporated in the following Phase II extension of (DLP_o).

Phase II

Using our knowledge of θ^*, we solve the following LP using $(\boldsymbol{\lambda},\ \boldsymbol{s}^-,\ \boldsymbol{s}^+)$ as variables:

$$\max_{\boldsymbol{\lambda},\boldsymbol{s}^-,\boldsymbol{s}^+} \quad \omega = \boldsymbol{e}\boldsymbol{s}^- + \boldsymbol{e}\boldsymbol{s}^+ \tag{3.11}$$

$$\text{subject to} \quad \boldsymbol{s}^- = \theta^* \boldsymbol{x}_o - X\boldsymbol{\lambda} \tag{3.12}$$

$$\boldsymbol{s}^+ = Y\boldsymbol{\lambda} - \boldsymbol{y}_o \tag{3.13}$$

$$\lambda \geq \mathbf{0}, \quad s^- \geq \mathbf{0}, \quad s^+ \geq \mathbf{0},$$

where $e = (1, \ldots, 1)$ (a vector of ones) so that $es^- = \sum_{i=1}^m s_i^-$ and $es^+ = \sum_{r=1}^s s_r^+$.

The objective of Phase II is to find a solution that maximizes the sum of input excesses and output shortfalls while keeping $\theta = \theta^*$.

We should note that we could replace the objective term in (3.11) with any weighted sum of input excesses and output shortfalls such as:

$$\omega = \boldsymbol{w}_x \boldsymbol{s}^- + \boldsymbol{w}_y \boldsymbol{s}^+, \tag{3.14}$$

where the weights $\boldsymbol{w}_x$ and $\boldsymbol{w}_y$ are positive row vectors. The modified objective function may result in a different optimal solution for Phase II. However, we can have the optimal $\omega^* > 0$ in (3.11) if and only if a nonzero value is also obtained when the objective in (3.11) is replaced with (3.14). Thus the objective in (3.11) will identify some nonzero slacks with inefficiency if and only if some nonzero (possibly different) slacks are identified with inefficiency in (3.14).

Definition 3.1 (Max-slack Solution, Zero-slack Activity)
An optimal solution $(\lambda^*,\ s^{-*},\ s^{+*})$ *of* Phase II *is called the* max-slack solution. *If the max-slack solution satisfies* $s^{-*} = \mathbf{0}$ *and* $s^{+*} = \mathbf{0}$, *then it is called* zero-slack.

Definition 3.2 (CCR-Efficiency, Radial Efficiency, Technical Efficiency)
If an optimal solution $(\theta^*,\ \lambda^*,\ s^{-*},\ s^{+*})$ *of the two LPs above satisfies* $\theta^* = 1$ *and is zero-slack* ($s^{-*} = \mathbf{0},\ s^{+*} = \mathbf{0}$), *then the* DMU_o *is called* CCR- efficient. *Otherwise, the* DMU_o *is called* CCR-inefficient, *because*

(i) $\theta^* = 1$

(ii) *All slacks are zero*

must both be satisfied if full efficiency is to be attained.

The first of these two conditions is referred to as "radial efficiency." It is also referred to as "technical efficiency" because a value of $\theta^* < 1$ means that all inputs can be simultaneously reduced without altering the mix (=proportions) in which they are utilized. Because $(1-\theta^*)$ is the maximal proportionate reduction allowed by the production possibility set, any further reductions associated with nonzero slacks will necessarily change the input proportions. Hence the inefficiencies associated with any nonzero slack identified in the above two-phase procedure are referred to as "mix inefficiencies." Other names are also used to characterize these two sources of inefficiency. For instance, the term "weak efficiency" is sometime used when attention is restricted to (i) in Definition 3.2. The conditions (i) and (ii) taken together describe what is also called "Pareto-Koopmans" or "strong" efficiency, which can be verbalized as follows,

Definition 3.3 (Pareto-Koopmans Efficiency)
A DMU is fully efficient if and only if it is not possible to improve any input or output without worsening some other input or output.

This last definition recognizes the contributions of the economists Vilfredo Pareto and Tjalling Koopmans. Implementable form was subsequently given to it by M.J. Farrell, another economist, who showed how to apply these concepts to observed data. However, Farrell was only able to carry his developments to a point which satisfied condition (*i*) but not condition (*ii*) in Definition 3.2. Hence he did not fully satisfy the conditions for Pareto-Koopmans efficiency but stopped short, instead, with what we just referred to as "weak efficiency" (also called "Farrell efficiency") because nonzero slack, when present in any input or output, can be used to effect additional improvements without worsening any other input or output. Farrell, we might note, was aware of this shortcoming in his approach which he tried to treat by introducing new (unobserved) "points at infinity" but was unable to give his concept implementable form.[4] In any case, this was all accomplished by Charnes, Cooper and Rhodes in a mathematical formulation that led to the two-phase procedure described above. Hence we also refer to this as "CCR-efficiency" when this procedure is utilized on empirical data to fulfill both (*i*) and (*ii*) in Definition 3.2.

We have already given a definition of CCR-efficiency in Chapter 2. For the data set (X, Y) under the assumption of semipositivity, we can also define CCR- efficiency by Definition 2.1 in that chapter. We now prove that the CCR-efficiency above, gives the same efficiency characterization as is obtained from Definition 2.1 in Chapter 2. This is formalized by:

Theorem 3.1 *The CCR-efficiency given in Definition 3.2 is equivalent to that given by Definition 2.1.*

Proof. First, notice that the vectors $\boldsymbol{v}$ and $\boldsymbol{u}$ of (LP_o) are dual multipliers corresponding to the constraints (3.7) and (3.8) of (DLP_o), respectively. See Table 3.1. Now the following "complementary conditions"[5] hold between any optimal solutions $(\boldsymbol{v}^*,\ \boldsymbol{u}^*)$ of (LP_o) and $(\boldsymbol{\lambda}^*,\ \boldsymbol{s}^{-*},\ \boldsymbol{s}^{+*})$ of (DLP_o).

$$\boldsymbol{v}^* \boldsymbol{s}^{-*} = 0 \ \text{ and } \ \boldsymbol{u}^* \boldsymbol{s}^{+*} = 0. \tag{3.15}$$

Known as the "complementary slackness" condition, this means that if any component of $\boldsymbol{v}^*$ or $\boldsymbol{u}^*$ is positive then the corresponding component of $\boldsymbol{s}^{-*}$ or $\boldsymbol{s}^{+*}$ must be zero, and conversely, with the possibility also allowed in which both components may be zero simultaneously.

Now we demonstrate that Definition 3.2 implies Definition 2.1

(*i*) If $\theta^* < 1$, then DMU$_o$ is CCR-inefficient by Definition 2.1, since (LP_o) and (DLP_o) have the same optimal objective value θ^*.

(*ii*) If $\theta^* = 1$ and is not zero-slack ($\boldsymbol{s}^{-*} \neq \boldsymbol{0}$ and/or $\boldsymbol{s}^{+*} \neq \boldsymbol{0}$), then, by the complementary conditions above, the elements of $\boldsymbol{v}^*$ or $\boldsymbol{u}^*$ corresponding to the positive slacks must be zero. Thus, DMU$_o$ is CCR- inefficient by Definition 2.1.

(*iii*) Lastly if $\theta^* = 1$ and zero-slack, then, by the "strong theorem of complementarity,"[6] (LP_o) is assured of a positive optimal solution $(\boldsymbol{v}^*, \boldsymbol{u}^*)$ and hence DMU$_o$ is CCR-efficient by Definition 2.1.

The reverse is also true by the complementary relation and the strong complementarity theorem between $(\boldsymbol{v}^*, \boldsymbol{u}^*)$ and $(\boldsymbol{s}^{-*}, \boldsymbol{s}^{+*})$. □

3.4 THE REFERENCE SET AND IMPROVEMENT IN EFFICIENCY

Definition 3.4 (Reference Set)
For an inefficient DMU_o, we define its reference set E_o, *based on the max-slack solution as obtained in phases one and two —see (3.11)— by*

$$E_o = \{j \mid \lambda_j^* > 0\} \ (j \in \{1, \ldots, n\}). \tag{3.16}$$

An optimal solution can be expressed as

$$\begin{aligned} \theta^* \boldsymbol{x}_o &= \sum_{j \in E_o} \boldsymbol{x}_j \lambda_j^* + \boldsymbol{s}^{-*} \\ \boldsymbol{y}_o &= \sum_{j \in E_o} \boldsymbol{y}_j \lambda_j^* - \boldsymbol{s}^{+*}, \end{aligned} \tag{3.17}$$

where $j \in E_o$ means the index j is included in the set E_o. This can be interpreted as follows,

$$\boldsymbol{x}_o \geq \theta^* \boldsymbol{x}_o - \boldsymbol{s}^{-*} = \sum_{j \in E_o} \boldsymbol{x}_j \lambda_j^*$$

which means

$$\begin{aligned} \boldsymbol{x}_o &\geq \text{technical} - \text{mix inefficiency} \\ &= \text{a positive combination of observed input values.} \end{aligned} \tag{3.18}$$

Also

$$\boldsymbol{y}_o \leq \boldsymbol{y}_o + \boldsymbol{s}^{+*} = \sum_{j \in E_o} \boldsymbol{y}_j \lambda_j^*$$

means

$$\begin{aligned} \boldsymbol{y}_o &\leq \text{observed outputs} + \text{shortfalls} \\ &= \text{a positive combination of observed output values.} \end{aligned} \tag{3.19}$$

These relations suggest that the efficiency of $(\boldsymbol{x}_o, \ \boldsymbol{y}_o)$ for DMU_o can be improved if the input values are reduced radially by the ratio θ^* and the input excesses recorded in $\boldsymbol{s}^{-*}$ are eliminated. Similarly efficiency can be attained if the output values are augmented by the output shortfalls in $\boldsymbol{s}^{+*}$. Thus, we have a method for improving an inefficient DMU that accords with Definition 3.2. The gross input improvement $\Delta \boldsymbol{x}_o$ and output improvement $\Delta \boldsymbol{y}_o$ can be calculated from:

$$\Delta \boldsymbol{x}_o = \boldsymbol{x}_o - (\theta^* \boldsymbol{x}_o - \boldsymbol{s}^{-*}) = (1 - \theta^*) \boldsymbol{x}_o + \boldsymbol{s}^{-*} \tag{3.20}$$
$$\Delta \boldsymbol{y}_o = \boldsymbol{s}^{+*}. \tag{3.21}$$

Hence, we have a formula for improvement, which is called *the CCR projection:*[7]

$$\widehat{\boldsymbol{x}}_o = \boldsymbol{x}_o - \Delta \boldsymbol{x}_o = \theta^* \boldsymbol{x}_o - \boldsymbol{s}^{-*} \leq \boldsymbol{x}_o \tag{3.22}$$
$$\widehat{\boldsymbol{y}}_o = \boldsymbol{y}_o + \Delta \boldsymbol{y}_o = \boldsymbol{y}_o + \boldsymbol{s}^{+*} \geq \boldsymbol{y}_o. \tag{3.23}$$

However, note that there are other formulae for improvement as will be described later.

In Theorems 3.2, 3.3 and 3.4 in the next section, we will show that the improved activity $(\widehat{\boldsymbol{x}}_o, \widehat{\boldsymbol{y}}_o)$ projects DMU$_o$ into the reference set E_o and any nonnegative combination of DMUs in E_o is efficient.

3.5 THEOREMS ON CCR-EFFICIENCY

Theorem 3.2 *The improved activity $(\widehat{\boldsymbol{x}}_o,\ \widehat{\boldsymbol{y}}_o)$ defined by (3.22) and (3.23) is CCR-efficient.*

Proof. The efficiency of $(\widehat{\boldsymbol{x}}_o,\ \widehat{\boldsymbol{y}}_o)$ is evaluated by solving the LP problem below:

$$(DLP_e) \qquad \min_{\theta,\boldsymbol{\lambda},\boldsymbol{s}^-,\boldsymbol{s}^+} \quad \theta \tag{3.24}$$

$$\text{subject to} \quad \theta\widehat{\boldsymbol{x}}_o - X\boldsymbol{\lambda} - \boldsymbol{s}^- = \boldsymbol{0} \tag{3.25}$$

$$Y\boldsymbol{\lambda} - \boldsymbol{s}^+ = \widehat{\boldsymbol{y}}_o \tag{3.26}$$

$$\boldsymbol{\lambda} \geq \boldsymbol{0},\quad \boldsymbol{s}^- \geq \boldsymbol{0},\quad \boldsymbol{s}^+ \geq \boldsymbol{0}. \tag{3.27}$$

Let an optimal (max-slack) solution for (DLP_e) be $(\widehat{\theta}, \widehat{\boldsymbol{\lambda}}, \widehat{\boldsymbol{s}}^-, \widehat{\boldsymbol{s}}^+)$. By inserting the formulae (3.22) and (3.23) into the constraints, we have

$$\widehat{\theta}\theta^*\boldsymbol{x}_o = X\widehat{\boldsymbol{\lambda}} + \widehat{\boldsymbol{s}}^- + \widehat{\theta}\boldsymbol{s}^{-*}$$
$$\boldsymbol{y}_o = Y\widehat{\boldsymbol{\lambda}} - \widehat{\boldsymbol{s}}^+ - \boldsymbol{s}^{+*}.$$

Now we can also write this solution as

$$\widetilde{\theta}\boldsymbol{x}_o = X\widehat{\boldsymbol{\lambda}} + \widetilde{\boldsymbol{s}}^-$$
$$\boldsymbol{y}_o = Y\widehat{\boldsymbol{\lambda}} - \widetilde{\boldsymbol{s}}^+$$

where $\widetilde{\theta} = \widehat{\theta}\theta^*$ and $\widetilde{\boldsymbol{s}}^- = \widehat{\boldsymbol{s}}^- + \widehat{\theta}\boldsymbol{s}^{-*} \geq \boldsymbol{0}$, $\widetilde{\boldsymbol{s}}^+ = \widehat{\boldsymbol{s}}^+ + \boldsymbol{s}^{+*} \geq \boldsymbol{0}$. However, θ^* is part of an optimal solution so we must have $\widetilde{\theta} = \widehat{\theta}\theta^* = \theta^*$ so $\widehat{\theta} = 1$. Furthermore, with $\widehat{\theta} = 1$ we have

$$\boldsymbol{e}\widetilde{\boldsymbol{s}}^- + \boldsymbol{e}\widetilde{\boldsymbol{s}}^+ = (\boldsymbol{e}\widehat{\boldsymbol{s}}^- + \boldsymbol{e}\boldsymbol{s}^{-*}) + (\boldsymbol{e}\widehat{\boldsymbol{s}}^+ + \boldsymbol{e}\boldsymbol{s}^{+*}) \leq \boldsymbol{e}\boldsymbol{s}^{-*} + \boldsymbol{e}\boldsymbol{s}^{+*}$$

since $\boldsymbol{e}\boldsymbol{s}^{-*} + \boldsymbol{e}\boldsymbol{s}^{+*}$ is maximal. It follows that we must have $\boldsymbol{e}\widehat{\boldsymbol{s}}^- + \boldsymbol{e}\widehat{\boldsymbol{s}}^+ = 0$ which implies that all components of $\widehat{\boldsymbol{s}}^-$ and $\widehat{\boldsymbol{s}}^+$ are zero. Hence conditions (i) and (ii) of Definition 3.2 are both satisfied and CCR-efficiency is achieved as claimed. □

Corollary 3.1 (Corollary to Theorem 3.2)
The point with coordinates $\widehat{\boldsymbol{x}}_o$, $\widehat{\boldsymbol{y}}_o$ defined by (3.22) and (3.23), viz.,

$$\widehat{\boldsymbol{x}}_o = \theta^*\boldsymbol{x}_o - \boldsymbol{s}^{-*} = \sum_{j\in E_o} \boldsymbol{x}_j\lambda_j^* \tag{3.28}$$

$$\widehat{\boldsymbol{y}}_o = \boldsymbol{y}_o + \boldsymbol{s}^{+*} = \sum_{j\in E_o} \boldsymbol{y}_j\lambda_j^* \tag{3.29}$$

is the point on the efficient frontier used to evaluate the performance of DMU_o.

In short the CCR projections identify the point either as a positive combination of other DMUs with $\boldsymbol{x}_o \geq \widehat{\boldsymbol{x}}_o$ and $\widehat{\boldsymbol{y}}_o \geq \boldsymbol{y}_o$ unless $\theta^* = 1$ and all slacks are zero in which case $\boldsymbol{x}_o = \widehat{\boldsymbol{x}}_o$ and $\widehat{\boldsymbol{y}}_o = \boldsymbol{y}_o$ so the operation in (3.28) performed on the observation for DMU_o identifies a new DMU positioned on the efficient frontier. Conversely, the point associated with the thus generated DMU evaluates the performance of DMU_o as exhibiting input excesses $\boldsymbol{x}_o - \widehat{\boldsymbol{x}}_o$ and output shortfalls $\widehat{\boldsymbol{y}}_o - \boldsymbol{y}_o$.

We notice that the improvement by the formulae (3.22) and (3.23) should be achieved by using the *max-slack* solution. If we do it based on another (not max-slack) solution, the improved activity ($\widehat{\boldsymbol{x}}_o$, $\widehat{\boldsymbol{y}}_o$) is not necessarily CCR-efficient. This is shown by the following example with 4 DMUs A, B, C and D, each with 3 inputs x_1, x_2 and x_3 and all producing 1 output in amount $y = 1$.

	A	B	C	D
x_1	2	2	2	1
x_2	1	1.5	2	1
x_3	1	1	1	2
y	1	1	1	1

Here the situation for DMU C is obvious. We can observe, for instance, that DMU C has two possible slacks in x_2, 1 against A and 0.5 against B. If we choose B as the reference set of C, then the improved activity coincides with B, which still has a slack of 0.5 in x_2 against A. Therefore, the improved activity is not CCR-efficient. However, if we improve C by using the max-slack solution, then we move to the CCR-efficient DMU A directly.

Lemma 3.1 *For the improved activity ($\widehat{\boldsymbol{x}}_o$, $\widehat{\boldsymbol{y}}_o$), there exists an optimal solution ($\widehat{\boldsymbol{v}}_o, \widehat{\boldsymbol{u}}_o$) for the problem ($LP_e$), which is dual to ($DLP_e$), such that*

$$\widehat{\boldsymbol{v}}_o > \mathbf{0} \text{ and } \widehat{\boldsymbol{u}}_o > \mathbf{0}$$

$$\widehat{\boldsymbol{v}}_o \boldsymbol{x}_j = \widehat{\boldsymbol{u}}_o \boldsymbol{y}_j \quad (j \in E_o) \tag{3.30}$$

$$\widehat{\boldsymbol{v}}_o X \geq \widehat{\boldsymbol{u}}_o Y \tag{3.31}$$

Proof. Since ($\widehat{\boldsymbol{x}}_o, \widehat{\boldsymbol{y}}_o$) is zero-slack, the strong theorem of complementarity means that there exists a positive optimal solution ($\widehat{\boldsymbol{v}}_o, \widehat{\boldsymbol{u}}_o$) for ($LP_e$). The equality (3.30) is the complementarity condition between primal-dual optimal solutions. The inequality (3.31) is a part of the constraints of (LP_e). □

Theorem 3.3 *The DMUs in E_o as defined in (3.16) are CCR-efficient.*

Proof. As described in Lemma 3.1, there exists a positive multiplier ($\widehat{\boldsymbol{v}}_o, \widehat{\boldsymbol{u}}_o$). These vectors also satisfy

$$\widehat{\boldsymbol{v}}_o \boldsymbol{x}_j = \widehat{\boldsymbol{u}}_o \boldsymbol{y}_j \quad (j \in E_o) \tag{3.32}$$

$$\widehat{\boldsymbol{v}}_o X \geq \widehat{\boldsymbol{u}}_o Y \tag{3.33}$$

For each $j \in E_o$, we can adjust $(\widehat{\boldsymbol{v}}_o, \widehat{\boldsymbol{u}}_o)$ by using a scalar multiplier so that the relation $\widehat{\boldsymbol{v}}_o \boldsymbol{x}_j = \widehat{\boldsymbol{u}}_o \boldsymbol{y}_j = 1$ holds, while keeping (3.33) satisfied. Thus, activity $(\boldsymbol{x}_j, \boldsymbol{y}_j)$ is CCR-efficient by Definition 2.1. (A proof via the envelopment model may be found in Problem 7.5 of Chapter 7, and this proof comprehends the BCC as well as the CCR model). □

Theorem 3.4 *Any semipositive combination of DMUs in E_o is CCR-efficient.*

Proof. Let the combined activity be

$$\boldsymbol{x}_c = \sum_{j \in E_o} c_j \boldsymbol{x}_j \text{ and } \boldsymbol{y}_c = \sum_{j \in E_o} c_j \boldsymbol{y}_j \text{ with } c_j \geq 0 \quad (j \in E_o). \tag{3.34}$$

The multiplier $(\widehat{\boldsymbol{v}}_o, \widehat{\boldsymbol{u}}_o)$ in Lemma 3.1 satisfies

$$\widehat{\boldsymbol{v}}_o \boldsymbol{x}_c = \widehat{\boldsymbol{u}}_o \boldsymbol{y}_c \tag{3.35}$$

$$\widehat{\boldsymbol{v}}_o X \geq \widehat{\boldsymbol{u}}_o Y \tag{3.36}$$

$$\widehat{\boldsymbol{v}}_o > \mathbf{0}, \ \widehat{\boldsymbol{u}}_o > \mathbf{0} \tag{3.37}$$

Thus, $(\boldsymbol{x}_c, \boldsymbol{y}_c)$ is CCR-efficient by Definition 2.1. □

3.6 COMPUTATIONAL ASPECTS OF THE CCR MODEL

In this section, we discuss computational procedures for solving linear programs for the CCR model in detail. Readers who are not familiar with the terminology of linear programs and are not interested in the computational aspects of DEA can skip this section.

3.6.1 Computational Procedure for the CCR Model

As described in Section 3.3, the computational scheme of the CCR model for DMU$_o$ results in the following two stage LP problem.

$$
\begin{array}{lrcll}
(DLP_o) & & & & \\
\text{Phase I objective} & \min & & \theta & (3.38)\\
\text{Phase II objective} & \min & & -\boldsymbol{e}\boldsymbol{s}^- - \boldsymbol{e}\boldsymbol{s}^+ & (3.39)\\
\text{subject to} & \theta \boldsymbol{x}_o & = & X\boldsymbol{\lambda} + \boldsymbol{s}^- & (3.40)\\
 & \boldsymbol{y}_o & = & Y\boldsymbol{\lambda} - \boldsymbol{s}^+ & (3.41)\\
\theta \geq 0, \ \boldsymbol{\lambda} \geq \mathbf{0}, & \boldsymbol{s}^- & \geq & \mathbf{0}, \ \boldsymbol{s}^+ \geq \mathbf{0}, & (3.42)
\end{array}
$$

where Phase II replaces the variable θ with a fixed value of min $\theta = \theta^*$.

Using the usual LP notation, we now rewrite (DLP_o) as follows,

$$
\begin{array}{lrcll}
(DLP'_o) & & & & \\
\text{Phase I objective} & \min \ z_1 & = & \boldsymbol{c}\boldsymbol{x} & (3.43)\\
\text{Phase II objective} & \min \ z_2 & = & \boldsymbol{d}\boldsymbol{x} & (3.44)\\
\text{subject to} & A\boldsymbol{x} & = & \boldsymbol{b} & (3.45)\\
 & \boldsymbol{x} & \geq & \mathbf{0}, & (3.46)
\end{array}
$$

where $\boldsymbol{c}$ and $\boldsymbol{d}$ are row vectors.

The correspondences between (DLP_o) and (DLP'_o) are:

$$\boldsymbol{x} = (\theta,\ \boldsymbol{\lambda}^T,\ \boldsymbol{s}^{-T},\ \boldsymbol{s}^{+T})^T \tag{3.47}$$

$$\boldsymbol{c} = (1,\ \boldsymbol{0},\ \boldsymbol{0},\ \boldsymbol{0}) \tag{3.48}$$

$$\boldsymbol{d} = (0,\ \boldsymbol{0},\ -\boldsymbol{e},\ -\boldsymbol{e}) \tag{3.49}$$

$$A = \begin{pmatrix} \boldsymbol{x}_o & -X & -I & O \\ \boldsymbol{0} & Y & O & -I \end{pmatrix} \tag{3.50}$$

$$\boldsymbol{b} = \begin{pmatrix} \boldsymbol{0} \\ \boldsymbol{y}_o \end{pmatrix}, \tag{3.51}$$

where $\boldsymbol{e}$ is the vector with all elements unity. See the discussion immediately following (3.13).

Phase I

First, we solve the LP problem with the Phase I objective. Letting an optimal basis be B we use this matrix B as follows. First we compute several values, as noted below, where R is the nonbasic part of the matrix A and the superscript B (or R) shows the columns of A corresponding to B (or R).

$$\text{basic solution} \quad \boldsymbol{x}^B = \bar{\boldsymbol{b}} = B^{-1}\boldsymbol{b} \tag{3.52}$$

$$\text{simplex multiplier} \quad \pi = \boldsymbol{c}^B B^{-1} \tag{3.53}$$

$$\text{Phase I simplex criterion} \quad \boldsymbol{p}^R = \pi R - \boldsymbol{c}^R \tag{3.54}$$

$$\text{Phase II simplex criterion} \quad \boldsymbol{q}^R = \pi R - \boldsymbol{d}^R, \tag{3.55}$$

where B^{-1} is the inverse of B, π is a vector of "multipliers" derived (from the data) as in (3.53), and $\boldsymbol{p}^R$ and $\boldsymbol{q}^R$ are referred to as "reduced costs."

From optimality of the basis B, we have:

$$\bar{\boldsymbol{b}} \geq \boldsymbol{0}, \tag{3.56}$$

because the conditions for non-negativity are satisfied by $\boldsymbol{x}^B$ in (3.52) and

$$\boldsymbol{p}^R \leq \boldsymbol{0}, \tag{3.57}$$

as required for optimality in (3.54) and, for the columns of the basis B, we have: $\boldsymbol{p}^B = \pi B - \boldsymbol{c}^B = \boldsymbol{0}$.

Phase II

We exclude the columns with $p_j < 0$ in the optimal simplex tableau at the end of Phase I from the tableaux for further consideration. The remainder is called the *restricted problem* or *restricted tableau* and dealt with in the next computational step. At Phase II, we solve the LP problem with the second objective. If the simplex criterion of Phase II objective satisfies $\boldsymbol{q}^R \leq \boldsymbol{0}$, we then halt the iterations. The basic solution thus obtained is a max-slack solution. This is the procedure for finding θ^* and a max-slack solution $(\boldsymbol{s}^{-*},\ \boldsymbol{s}^{+*})$ for DMU$_o$. As has been already pointed out, (LP_o) in (3.2)-(3.5) is dual to (DLP_o) in (3.6)-(3.9), and we will discuss its solution procedure in the next section.

3.6.2 Data Envelopment Analysis and the Data

The characterizations given by (3.52)-(3.55) represent borrowings from the terminology of linear programming in which the components of π are referred to as "simplex multipliers" because they are associated with a use of the simplex method which generates such multipliers in the course of solving LP problems such as (DLP'_o). "Simplex" and "dual simplex" methods are also used in DEA. We therefore refer to the problem (LP_o) as being in "multiplier form." See (3.2)-(3.5). Problem (DLP_o) is then said to be in "envelopment form." This is the source of the name "Data Envelopment Analysis."

Reference to Figure 3.2 in the next section will help to justify this usage by noting, first, that all data are inside the frontier stretching from F through R. Also, at least one observation is touching the frontier. Hence, in the manner of an envelope all data are said to be "enveloped" by the frontier which achieves its importance because it is used to analyze and evaluate the performances of DMUs associated with observations like those portrayed in Figure 3.2. More compactly, we say that DEA is used to evaluate the performances of each observation relative to the frontier that envelops all of the observations.

3.6.3 Determination of Weights (=Multipliers)

The simplex multipliers in π at the end of Phase I are associated with an optimal solution of its dual problem (LP_o)[8] as given in (3.2)-(3.5). In fact, π is an $(m+s)$ vector in which the first m components assign optimal weights $\boldsymbol{v}^*$ to the inputs and the remaining s components assign optimal weights $\boldsymbol{u}^*$ to the outputs. By observing the dual problem of (DLP'_o) in (3.38)-(3.42), it can be shown that $\boldsymbol{v}^*$ and $\boldsymbol{u}^*$ satisfy (3.3) through (3.5). The symbol I, for the identity matrix in A of (3.50), relates to the (input and output) slack. Let the simplex criteria to be applied to these columns be represented by "pricing vectors" $\boldsymbol{p}^{s^-}$ and $\boldsymbol{p}^{s^+}$. These vectors relate to $\boldsymbol{v}^*$ and $\boldsymbol{u}^*$ in the following way:

$$\boldsymbol{v}^* = -\boldsymbol{p}^{s^-} \quad (\geq \boldsymbol{0}) \tag{3.58}$$

$$\boldsymbol{u}^* = -\boldsymbol{p}^{s^+} \quad (\geq \boldsymbol{0}). \tag{3.59}$$

3.6.4 Reasons for Solving the CCR Model Using the Envelopment Form

It is not advisable to solve (LP_o) directly. The reasons are:

(1) The computational effort of LP is apt to grow in proportion to powers of the number of constraints. Usually in DEA, n, the number of DMUs is considerably larger than $(m+s)$, the number of inputs and outputs and hence it takes more time to solve (LP_o) which has n constraints than to solve (DLP_o) which has $(m+s)$ constraints. In addition, since the memory size needed for keeping the basis (or its inverse) is the square of the number of constraints, (DLP_o) is better fitted for memory saving purposes.

(2) We cannot find the pertinent max-slack solution by solving (LP_o).

(3) The interpretations of (DLP_o) are more straightforward because the solutions are characterized as inputs and outputs that correspond to the original

data whereas the multipliers provided by solutions to (LP_o) represent evaluations of these observed values. These values are also important, of course, but they are generally best reserved for supplementary analyses after a solution to (DLP_o) is achieved.

3.7 EXAMPLE

We will apply (DLP_o) to a sample problem and comment on the results. For this purpose we use Example 3.1 as shown in Table 3.2, which is Example 2.2 with an added activity G. All computations were done using the DEA-Solver which comes with this book and the results are stored in the Excel 97 Workbook "Sample-CCR-I.xls" in the sample file. So, readers interested in using DEA-Solver should try to trace the results by opening this file or running DEA-Solver for this problem.

Table 3.2. Example 3.1

	DMU	A	B	C	D	E	F	G
Input	x_1	4	7	8	4	2	10	3
	x_2	3	3	1	2	4	1	7
Output	y	1	1	1	1	1	1	1

(1) (DLP_A) for A is:

$$
\begin{aligned}
&\text{Phase I} \quad \min\ \theta \\
&\text{Phase II} \quad \min\ -s_1^- - s_2^- - s^+ \\
&\text{subject to} \\
&4\theta - 4\lambda_A - 7\lambda_B - 8\lambda_C - 4\lambda_D - 2\lambda_E - 10\lambda_F - 3\lambda_G - s_1^- = 0 \\
&3\theta - 3\lambda_A - 3\lambda_B - \lambda_C - 2\lambda_D - 4\lambda_E - \lambda_F - 7\lambda_G - s_2^- = 0 \\
&\lambda_A + \lambda_B + \lambda_C + \lambda_D + \lambda_E + \lambda_F + \lambda_G - s^+ = 1
\end{aligned}
$$

where the variables are restricted to nonnegative values in the vectors $\boldsymbol{\lambda}$, $\boldsymbol{s}^-$ and $\boldsymbol{s}^+$.
The optimal solution for (DLP_A) is:

$\theta^* = 0.8571$ (in Worksheet "Sample-CCR-I.Score")
$\lambda_D^* = 0.7143$, $\lambda_E^* = 0.2857$, other $\lambda_j^* = 0$ (in Worksheet "Sample-CCR-I.Score")
$s_1^{-*} = s_2^{-*} = s^{+*} = 0$ (in Worksheet "Sample-CCR-I.Slack").

Since $\lambda_D^* > 0$ and $\lambda_E^* > 0$, the reference set for A is

$$E_A = \{D,\ E\}.$$

And $\lambda_D^* = 0.7143$, $\lambda_E^* = 0.2857$ show the proportions contributed by D and E to the point used to evaluate A. Hence A is technically inefficient. No mix inefficiencies are present because all slacks are zero. Thus removal of all inefficiencies is achieved by reducing all inputs by 0.1429 or, approximately, 15% of their observed values. In fact, based on this reference set and $\boldsymbol{\lambda}^*$, we can express the input and output values needed to bring A into efficient status as

$$0.8571 \times (\text{Input of } A) = 0.7143 \times (\text{Input of } D) + 0.2857 \times (\text{Input of } E)$$
$$(\text{Output of } A) = 0.7143 \times (\text{Output of } D) + 0.2857 \times (\text{Output of } E).$$

From the magnitude of coefficients on the right hand side, A has more similarity to D than E. A can be made efficient either by using these coefficients, $\lambda_D^* = 0.7143$, $\lambda_E^* = 0.2857$ or by reducing both of its inputs—viz., by reducing the input value radially in the ratio 0.8571. It is the latter (radial contraction) that is used in DEA-Solver. Thus, as seen in Worksheet "Sample-CCR-I.Projection," the CCR-projection of (3.32) and (3.33) is achieved by,

$$\begin{aligned}
\widehat{x}_1 &\leftarrow \theta^* x_1 = 0.8571 \times 4 = 3.4286 \quad (14.29\% \text{ reduction}) \\
\widehat{x}_2 &\leftarrow \theta^* x_2 = 0.8571 \times 3 = 2.5714 \quad (14.29\% \text{ reduction}) \\
\widehat{y} &\leftarrow y = 1 \quad (\text{no change}).
\end{aligned}$$

The optimal solution for the multiplier problem (LP_A) can be found in Worksheet "Sample-CCR-I.Weight" as follows,

$$v_1^* = 0.1429, \quad v_2^* = 0.1429, \quad u^* = 0.8571.$$

This solution satisfies constraints (3.3)-(3.5) and maximizes the objective in (3.2), i.e., $u^* y = 0.8571 \times 1 = 0.8571 = \theta^*$ in the optimal objective value of (DLP_A). (See Problem 3.4 for managerial roles of these optimal multipliers (weights) for improving A.)

The Worksheet "Sample-CCR-I.WeightedData" includes optimal weighted inputs and output, i.e.,

$$\begin{aligned}
v_1^* x_1 &= 0.1429 \times 4 = 0.5714 \\
v_2^* x_2 &= 0.1429 \times 3 = 0.4286 \\
u^* y &= 0.8571 \times 1 = 0.8571.
\end{aligned}$$

The sum of the first two terms is 1 which corresponds to the constraint (3.3). The last term is the optimal objective value in this single output case.

(2) Moving from DMU A to DMU B, the optimal solution for B is:

$$\begin{aligned}
&\theta^* = 0.6316 \\
&\lambda_A^* = \lambda_B^* = 0,\ \lambda_C^* = 0.1053,\ \lambda_D^* = 0.8947,\ \lambda_E^* = \lambda_F^* = \lambda_G^* = 0 \\
&s_1^{-*} = s_2^{-*} = s^{+*} = 0 \\
&v_1^* = 0.0526,\ v_2^* = 0.2105,\ u^* = 0.6316.
\end{aligned}$$

Since $\lambda_C^* > 0,\ \lambda_D^* > 0$, the reference set for B is:

$$E_B = \{C,\ D\}.$$

B can be expressed as:

$$0.6316 \times (\text{Input of } B) = 0.1053 \times (\text{Input of } C) + 0.8947 \times (\text{Input of } D)$$
$$(\text{Output of } B) = 0.1053 \times (\text{Output of } C) + 0.8947 \times (\text{Output of } D).$$

That is, it can be expressed in ratio form, as shown on the left or as a nonnegative combination of $\lambda_j^* > 0$ values as shown on the right. Using the expression on the left, the CCR-projection for B is

$$\begin{aligned}
&\widehat{x}_1 \leftarrow \theta^* x_1 = 0.6316 \times 7 = 4.4211 \quad (36.84\% \text{ reduction})\\
&\widehat{x}_2 \leftarrow \theta^* x_2 = 0.6316 \times 3 = 1.8974 \quad (36.84\% \text{ reduction})\\
&\widehat{y} \leftarrow y = 1 \quad (\text{no change}).
\end{aligned}$$

(3) C, D and E

These 3 DMUs are found to be efficient. (See Worksheet "Sample-CCR-I.Score.")

(4) The optimal solution of the LP problem for F is:

$$\begin{aligned}
&\theta^* = 1\\
&\lambda_A^* = \lambda_B^* = 0,\ \lambda_C^* = 1,\ \lambda_D^* = \lambda_E^* = \lambda_F^* = \lambda_G^* = 0\\
&s_1^{-*} = 2,\ s_2^{-*} = s^{+*} = 0\\
&v_1^* = 0,\ v_2^* = 1,\ u^* = 1
\end{aligned}$$

where, again, the last set of values refer to the multiplier problem. For the envelopment model we have $\lambda_C^* > 0$ as the only positive value of λ. Hence the reference set for F is:

$$E_F = \{C\}.$$

Considering the excess in Input 1 ($s_1^{-*} = 2$), F can be expressed as:

$$\begin{aligned}
(\text{Input 1 of } F) &= (\text{Input 1 of } C) + 2\\
(\text{Input 2 of } F) &= (\text{Input 2 of } C)\\
(\text{Output of } F) &= (\text{Output of } C).
\end{aligned}$$

Although F is "radial-efficient," it is nevertheless "CCR-inefficient" due to this excess (mix inefficiency) associated with $s_1^{-*} = 2$. Thus the performance of F can be improved by subtracting 2 units from Input 1. This can be accomplished by subtracting 2 units from input 1 on the left and setting $s_1^{-*} = 0$ on the right without worsening any other input and output. Hence condition (*ii*) in Definition 3.2 is not satisfied until this is done, so F did *not* achieve Pareto-Koopmans efficiency in its performance. See Definition 3.3.

(5) The optimal solution of the LP problem for G is:

$$\begin{aligned}
&\theta^* = 0.6667\\
&\lambda_A^* = \lambda_B^* = \lambda_C^* = \lambda_D^* = 0,\ \lambda_E^* = 1,\ \lambda_F^* = \lambda_G^* = 0\\
&s_1^{-*} = 0,\ s_2^{-*} = 0.6667,\ s^{+*} = 0\\
&v_1^* = 0.3333,\ v_2^* = 0,\ u^* = 0.6667
\end{aligned}$$

Since $\lambda_E^* > 0$, the reference set for G is:

$$E_G = \{E\}.$$

Considering the excess in Input 2 ($s_2^{-*} = 0.6667$), G can be expressed by:

$$\begin{aligned}
0.6667 \times (\text{Input 1 of } G) &= (\text{Input 1 of } E)\\
0.6667 \times (\text{Input 2 of } G) &= (\text{Input 2 of } E) + 0.6667\\
(\text{Output of } G) &= (\text{Output of } E).
\end{aligned}$$

One plan for the improvement of G is to reduce all input values by multiplying them by 0.6667 and further subtracting 0.6667 from Input 2. When this is done the thus altered values coincide with the coordinates of E. Geometrically then, the CCR-projection for G is

$$\begin{aligned}
&\widehat{x}_1 \leftarrow \theta^* x_1 - s_1^{-*} = 0.6667 \times 3 - 0 = 2 \quad (33.33\% \text{ reduction})\\
&\widehat{x}_2 \leftarrow \theta^* x_2 - s_2^{-*} = 0.6667 \times 7 - 0.6667 = 4 \quad (42.86\% \text{ reduction})\\
&\widehat{y} \leftarrow y = 1 \quad (\text{no change}),
\end{aligned}$$

where $\widehat{x}_1 = 2,\ \widehat{x}_2 = 4,\ \widehat{y} = 1$ which values are the same as for E in Table 3.2.

The above observations are illustrated by Figure 3.2, which depicts Input 1 and Input 2 values of all DMUs. Since the output value is 1 for all the DMUs, we can compare their efficiencies via the input values.

The efficient frontier consists of the bold line CDE and the production possibility set is the region enclosed by this efficient frontier line plus the vertical line going up from E and the horizontal line extending to the right from C. Let the intersection point of OA and DE be Q. The activity Q has input proportional to that of A (4, 3) and is the least input value point on OA in the production possibility set and

$$\frac{OQ}{OA} = 0.8571$$

corresponds to radial (or ratio) efficiency of A. [9] Also, we have:

$$\begin{aligned}
\text{Input 1 of } Q &= 0.8571 \times 4 \ (\text{Input 1 of } A) = 3.428 = \widehat{x}_1\\
\text{Input 2 of } Q &= 0.8571 \times 3 \ (\text{Input 2 of } A) = 2.571 = \widehat{x}_2.
\end{aligned}$$

However, Q is the point that divides D and E in the ratio 0.7143 to 0.2857.

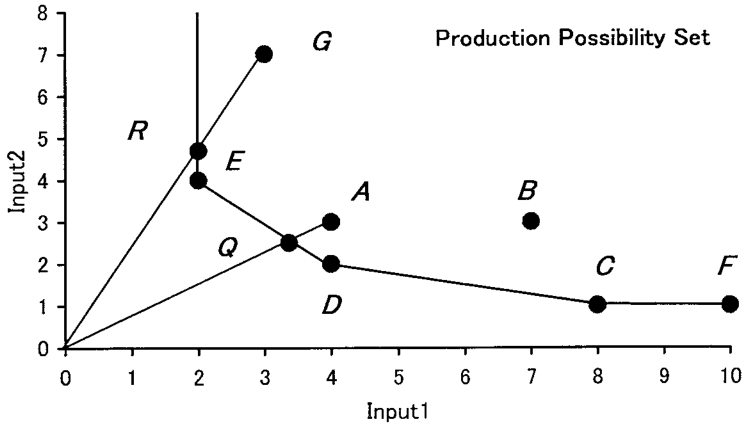

Figure 3.2. Example 3.1

Hence its values are calculated again as:

$$
\begin{aligned}
\text{Input 1 of } Q &= 0.7143 \times 4 \text{ (Input 1 of } D) + 0.2857 \times 2 \text{ (Input 1 of } E) &= 3.428 \\
\text{Input 2 of } Q &= 0.7143 \times 2 \text{ (Input 2 of } D) + 0.2857 \times 4 \text{ (Input 2 of } E) &= 2.571,
\end{aligned}
$$

where $\lambda_D^* = 0.7143$ and $\lambda_E^* = 0.2857$. Comparing these results we see that the coordinates of Q, the DMU used to evaluate A, can be derived in either of these two ways. There are also still more ways of effecting such improvements, as we will later see when changes in mix proportions are permitted.

This brings us to the role of nonzero slacks which we can illustrate with F and G. From Figure 3.2, it is evident that if we reduce Input 1 of F by the nonzero slack value of 2, then F coincides with C and is efficient. The presence of this nonzero slack means that F is not Pareto-Koopmans efficient even though its radial value is $\theta^* = 1$.

As a further illustration of the two conditions for efficiency in Definition 3.2 we turn to the evaluation of G. The CCR-efficiency of G is calculated by $OR/OG = 0.6667$. Thus G is not radially (or weakly) efficient as evaluated at R. However R is also not efficient. We can make it efficient by further reducing Input 2 by 0.6667 and shifting to E as we have seen in the CCR-projection of G. Hence G fails both of the conditions specified in Definition 3.2.

Table 3.3 summarizes the results obtained by applying DEA-Solver to all of the data in Table 3.2. Only C, D and E are fully efficient. A and B fail because $\theta^* < 1$. Their intersection with the frontier gives their radial (weak) inefficiency score with zero slacks because they intersect an efficient portion of the frontier radially. F has a value of $\theta^* = 1$ because it is on the frontier. This portion of frontier is not efficient, however, as evidenced by the nonzero slack for F under s_1^- in Table 3.3. Finally, G fails to be efficient both because $\theta^* < 1$ *and* nonzero slacks are involved in its evaluation by E, a point on the efficient portion of the frontier.

Table 3.3. Results of Example 3.1

DMU	CCR-Eff	Ref Set	Excess		Shortfall
	θ^*		s_1^-	s_2^-	s^+
A	0.8571	D E	0	0	0
B	0.6316	C D	0	0	0
C	1.0000	C	0	0	0
D	1.0000	D	0	0	0
E	1.0000	E	0	0	0
F	1.0000	C	2	0	0
G	0.6667	E	0	.6667	0

3.8 THE OUTPUT-ORIENTED MODEL

Up to this point, we have been dealing mainly with a model whose objective is to minimize inputs while producing at least the given output levels. This type of model is called *input-oriented.* There is another type of model that attempts to maximize outputs while using no more than the observed amount of any input. This is referred to as the *output-oriented* model, formulated as:

$$(DLPO_o) \quad \max_{\eta,\boldsymbol{\mu}} \quad \eta \tag{3.60}$$

$$\text{subject to} \quad \boldsymbol{x}_o - X\boldsymbol{\mu} \geq \mathbf{0} \tag{3.61}$$

$$\eta\boldsymbol{y}_o - Y\boldsymbol{\mu} \leq \mathbf{0} \tag{3.62}$$

$$\boldsymbol{\mu} \geq \mathbf{0}. \tag{3.63}$$

An optimal solution of $(DLPO_o)$ can be derived directly from an optimal solution of the input-oriented CCR model given in (3.6)-(3.9) as follows. Let us define

$$\boldsymbol{\lambda} = \boldsymbol{\mu}/\eta, \quad \theta = 1/\eta. \tag{3.64}$$

Then $(DLPO_o)$ becomes

$$\begin{aligned} (DLP_o) \quad & \min_{\theta,\boldsymbol{\lambda}} && \theta \\ & \text{subject to} && \theta\boldsymbol{x}_o - X\boldsymbol{\lambda} \geq \mathbf{0} \\ & && \boldsymbol{y}_o - Y\boldsymbol{\lambda} \leq \mathbf{0} \\ & && \boldsymbol{\lambda} \geq \mathbf{0}, \end{aligned}$$

which is the input-oriented CCR model. Thus, an optimal solution of the output-oriented model relates to that of the input-oriented model via:

$$\eta^* = 1/\theta^*, \quad \boldsymbol{\mu}^* = \boldsymbol{\lambda}^*/\theta^*. \tag{3.65}$$

The slack ($\boldsymbol{t}^{-}$, $\boldsymbol{t}^{+}$) of the output-oriented model is defined by:

$$X\boldsymbol{\mu} + \boldsymbol{t}^{-} = \boldsymbol{x}_o$$
$$Y\boldsymbol{\mu} - \boldsymbol{t}^{+} = \eta \boldsymbol{y}_o.$$

These values are also related to the input-oriented model via

$$\boldsymbol{t}^{-*} = \boldsymbol{s}^{-*}/\theta^*, \quad \boldsymbol{t}^{+*} = \boldsymbol{s}^{+*}/\theta^*. \tag{3.66}$$

Now, $\theta^* \leq 1$, so returning to (3.64), η^* satisfies

$$\eta^* \geq 1. \tag{3.67}$$

The higher the value of η^*, the less efficient the DMU. θ^* expresses the input reduction rate, while η^* describes the output enlargement rate. From the above relations, we can conclude that an input-oriented CCR model will be efficient for any DMU if and only if it is also efficient when the output-oriented CCR model is used to evaluate its performance.

The dual problem of $(DLPO_o)$ is expressed in the following model, with components of the vectors $\boldsymbol{p}$ and $\boldsymbol{q}$ serving as variables.

$$(LPO_o) \quad \min_{\boldsymbol{p},\boldsymbol{q}} \quad \boldsymbol{p}\boldsymbol{x}_o \tag{3.68}$$
$$\text{subject to} \quad \boldsymbol{q}\boldsymbol{y}_o = 1 \tag{3.69}$$
$$-\boldsymbol{p}X + \boldsymbol{q}Y \leq \boldsymbol{0} \tag{3.70}$$
$$\boldsymbol{p} \geq \boldsymbol{0}, \quad \boldsymbol{q} \geq \boldsymbol{0}. \tag{3.71}$$

On the multiplier side we have:

Theorem 3.5 *Let an optimal solution of (LP_o) be $(\boldsymbol{v}^*, \boldsymbol{u}^*)$, then an optimal solution of the output-oriented model (LPO_o) is obtained from*

$$\boldsymbol{p}^* = \boldsymbol{v}^*/\theta^*, \qquad \boldsymbol{q}^* = \boldsymbol{u}^*/\theta^*. \tag{3.72}$$

Proof. It is clear that $(\boldsymbol{p}^*, \boldsymbol{q}^*)$ is feasible for (LPO_o). Its optimality comes from the equation below.

$$\boldsymbol{p}^*\boldsymbol{x}_o = \boldsymbol{v}^*\boldsymbol{x}_o/\theta^* = \eta^*. \tag{3.73}$$

□

Thus, the solution of the output-oriented CCR model may be obtained from that of the input oriented CCR model. The improvement using this model is expressed by:

$$\widehat{\boldsymbol{x}}_o \Leftarrow \boldsymbol{x}_o - \boldsymbol{t}^{-*} \tag{3.74}$$
$$\widehat{\boldsymbol{y}}_o \Leftarrow \eta^*\boldsymbol{y}_o + \boldsymbol{t}^{+*}. \tag{3.75}$$

Carrying this a stage further we note that (LPO_o) is equivalent to the following fractional programming problem:

$$\min_{\pi,\rho} \quad \frac{\pi x_o}{\rho y_o} \tag{3.76}$$

$$\text{subject to} \quad \frac{\pi x_j}{\rho y_j} \geq 1 \quad (j = 1, \ldots, n) \tag{3.77}$$

$$\pi \geq \mathbf{0}, \quad \rho \geq \mathbf{0}. \tag{3.78}$$

That is, we exchanged the numerator and the denominator of (2.3) and (2.4) as given in Chapter 2 and minimized the objective function. It is therefore quite natural that the solutions are found to be linked by a simple rule. This mathematical transformation does not imply that there is no managerial significance to be assigned to the choice of models since, *inter alia*, different corrections may be associated with output maximization and input minimization. The difference can be substantial so this choice always deserves consideration. Furthermore, later in this book, we shall also introduce other models, where outputs are maximized and inputs are simultaneously minimized so that still further choices may need to be considered.

3.9 DISCRETIONARY AND NON-DISCRETIONARY INPUTS

Up to this point we have assumed that all inputs and outputs can be varied at the discretion of management or other users. These may be called "discretionary variables." "Non-discretionary variables," not subject to management control, may also need to be considered. In evaluating performances of different bases for the Fighter Command of U.S. Air Forces, for instance, it was necessary to consider weather as an input since the number of "sorties" (successfully completed missions) and "aborts" (non-completed mission)[10] treated as outputs, could be affected by the weather (measured in degree days and numbers of "flyable" days) at different bases.

Even though Non-Discretionary, it is important to take account of such inputs in a manner that is reflected in the measures of efficiency used. We follow the route provided by Banker and Morey (1986)[11] who refer to such variables as "exogenously fixed," in forms like "age of store," in their use of DEA to evaluate the performances of 60 DMUs in a network of fast-food restaurants. Reverting to algebraic notation, we can represent their formulation by the following modification of the CCR model.

$$\min \quad \theta - \varepsilon \left(\sum_{i \in D} s_i^- + \sum_{r=1}^{s} s_r^+ \right) \tag{3.79}$$

$$\text{subject to} \quad \theta x_{io} = \sum_{j=1}^{n} x_{ij}\lambda_j + s_i^-, \quad i \in D \tag{3.80}$$

$$x_{io} = \sum_{j=1}^{n} x_{ij}\lambda_j + s_i^-, \quad i \in ND \tag{3.81}$$

$$y_{ro} = \sum_{j=1}^{n} y_{rj}\lambda_j - s_r^+, \quad r = 1, \ldots, s. \tag{3.82}$$

where all variables (except θ) are constrained to be nonnegative.

Here the symbols $i \in D$ and $i \in ND$ refer to the sets of "Discretionary" and "Non-Discretionary" inputs, respectively. To be noted in the constraints is the fact that the variable θ is not applied to the latter inputs because these values are exogenously fixed and it is therefore not possible to vary them at the discretion of management. This is recognized by entering all x_{io}, $i \in ND$ at their fixed (observed) value. Turning to the objective (3.79) we utilize the symbol[12] $\varepsilon > 0$ to mean that the slack variables (shown in the parenthses) are to be handled at a second stage where, as previously described, they are to be maximized in a manner that does not disturb the previously determined first-stage minimization of θ to achieve $\theta = \theta^*$. Finally, we note that the slacks $s_i^-, i \in ND$ are omitted from the objective. Hence these Non-Discretionary inputs do not enter directly into the efficiency measures being optimized in (3.79). They can, nevertheless, affect the efficiency evaluations by virtue of their presence in the constraints.

We can further clarify the way these Non-Discretionary variables affect the efficiency scores by writing the dual of (3.79)-(3.82) in the form of the following (modified) multiplier model,

$$\max \quad \sum_{r=1}^{s} u_r y_{ro} - \sum_{i \in ND} v_i x_{io} \tag{3.83}$$

$$\text{subject to} \quad \sum_{r=1}^{s} u_r y_{rj} - \sum_{i \in ND} v_i x_{ij} - \sum_{i \in D} v_i x_{ij} \leq 0, \ j = 1, \ldots, n \tag{3.84}$$

$$\sum_{i \in D} v_i x_{io} = 1 \tag{3.85}$$

$$v_i \geq \varepsilon, \ i \in D \tag{3.86}$$

$$v_i \geq 0, \ i \in ND \tag{3.87}$$

$$u_r \geq \varepsilon, \ r = 1, \ldots, s. \tag{3.88}$$

As can be seen, the Non-Discretionary but not the Discretionary inputs, enter into the objective (3.83). The multiplier values associated with these Non-Discretionary inputs may be zero, as in (3.87), but the other variables must always be positive, as in (3.86) and (3.88). The interpretations we now provide flow from the "complementary slackness principle" of linear programming. If (3.81) is satisfied strictly at an optimum then $v_i^* = 0$ is associated with this constraint and this x_{io} does not affect the evaluation recorded in (3.79). On the other hand, if $v_i^* > 0$ for any $i \in ND$ then the efficiency score recorded in (3.79) is reduced by the multiplier of x_{io} for this DMU_o.

This follows from the dual theorem of linear programming—viz.,

$$\theta^* - \varepsilon\left(\sum_{i\in D} s_i^{-*} + \sum_{r=1}^{s} s_r^{+*}\right) = \sum_{r=1}^{s} u_r^* y_{ro} - \sum_{i\in ND} v_i^* x_{io}. \tag{3.89}$$

Via this same relation we see that a *decrease* in this same x_{io} will *increase* the efficiency score recorded in the expression or the left of the equality (3.89).

For managerial use, the sense of this mathematical characterization may be interpreted in the following manner. The output achieved, as recorded in the y_{ro}, deserve a higher efficiency rating when they have been achieved under a relatively tighter constraint and a lower efficiency score when this constraint is loosened by increasing this x_{io}.

This treatment of Non-Discretionary variables must be qualified, at least to some extent, since, *inter alia,* allowance must be made for ranges over which the values of v_i^* and u_r^* remain unaltered. Modifications to allow for $v_i^* \geq \varepsilon \geq 0$ must also be introduced, as described in Problem 3.2 at the end of this chapter and effects like those treated in Problem 3.3 need to recognized when achievement of the efficient frontier is a consideration.

After additional backgrounds have been supplied, Chapter 7, later in this book, introduces a variety of extensions and alternate ways of treating Non-Discretionary outputs as well as inputs. Here we only need to note that modifications of a relatively straightforward variety may also be made as in the following example.

A study of 638 public secondary schools in Texas was undertaken by a consortium of 3 universities in collaboration with the Texas Education Agency (TEA). The study was intended to try to develop improved methods for accountability and evaluation of school performances. In addition to discretionary inputs like teacher salaries, special instruction, etc., the following Non-Discretionary inputs had to be taken into account,

1. **Minority** : Number of minority students, expressed as a percentage of total student enrollment.

2. **Disadvantage** : Number of economically disadvantaged students, expressed as a percentage of total student enrollment.

3. **LEP** : Number of Limited English Proficiency students, expressed as a percentage of total student enrollment.

These inputs differ from the ones that Banker and Morey had in mind. For instance, a regression calculation made at an earlier stage of the consortium study yielded *negative* coefficients that tested statistically significant for every one of these 3 inputs in terms of their effects on academic test scores. In the subsequent DEA study it was therefore deemed desirable to reverse the sign associated with the x_{io} in the expression on the right of (3.89)

This was accomplished by reversing (3.81)

$$\text{from } \sum_{j=1}^{n} x_{ij}\lambda_j \leq x_{io} \text{ to } \sum_{j=1}^{n} x_{ij}\lambda_j \geq x_{io} \tag{3.90}$$

for each of these $i \in ND$. In this manner an ability to process more of these inputs, when critical, was recognized in the form of higher (rather than lower) efficiency scores. Full details are supplied in I.R. Bardhan (1995)[13] and summarized in Arnold *et al.* (1997).[14] Here we note that the results were sufficiently satisfactory to lead to a recommendation to accord recognition to "efficiency" as well as "effectiveness" in state budgetary allocations to different school districts. Although not accepted, the recommendation merits consideration at least to the extent of identifying shortcomings in the usual statements (and reward structures) for academic performance. The following scheme can help to clarify what is being said,[15]

Effectiveness implies

- Ability to state desired goals
- Ability to achieve desired goals

Efficiency relates to

- Benefits realized
- Resources used

Consider, for instance, the State-mandated Excellence Standards for Texas recorded in Table 3.4. These represent statements of desired goals and schools are rewarded (or not rewarded) on the basis of their achievements. Nothing is

Table 3.4. State-mandated Excellence Standards on Student Outcomes

	Outcome Indicator	State-mandated Excellence Standard
1.	Texas Assessment of Academic Skills (TAAS) Test	90% of students passing on all standardized tests
2.	Attendance	97% of total enrollment in the school
3.	Dropout Rate	Less than or equal to 1% of total enrollment
4.	Graduation Rate	99% of graduating class
5.	College Admission Tests	• 35% of graduates scoring above the criterion score which is equal to 25 on the ACT[a] and 1000 on the SAT[b] • 70% of graduates taking either the ACT or the SAT

[a] ACT = American Collegiate Tests
[b] SAT = Scholastic Aptitude Tests

said about the amounts (or varieties) of resources used. Hence it should be no surprise that only 1 of the excellent-rated schools included in this study was found to be efficient. The other schools rated as excellent by the State of Texas had all expended excessive resources in achieving these goals. On the other hand many schools that failed to achieve excellence were nevertheless found to be efficient in producing desired outputs under very difficult conditions. Therefore some way of recognizing this kind of achievements is needed if "efficiency" as well as "effectiveness" is to be rewarded.[16]

Table 3.5 as taken from Bardhan (1994), contains portions of the printout from a DEA study involving one of these "excellent" schools. The row labelled Disadv (= Economically Disadvantaged) shows a slack value of 15 students. Judging from the performances of its peer group of efficient schools, this school

Table 3.5. Non-Discretionary Inputs

	Current Level	*Slack*	*Value if Efficient*
Minority	47.0	—	47.0
Disadv	14.0	15.0	29.0
LEP	3.5	—	3.5

should therefore have been able to more than double the number of such students that it processed *without affecting its efficiency score* because (a) the slack for $i \in ND$ is not in the objective of (3.79) and (b) the presence of nonzero slack for any $i \in ND$ means the associated $v_i^* = 0$. In addition, the other ND inputs are associated with positive multiplier values so this school would have been able to *increase* its efficiency score by *increasing* the Minority and LEP (Low English Proficiency = Mainly Hispanic) students it processed. As shown in Arnold *et al.* (1997) this can be done by introducing constraints to insure that no worsening of any already achieved excellence is avoided.

This is not the end of the line for what can be done. See Section 7.3 in Chapter 7, below, for further developments. Additional extensions could proceed to a two-stage operation in which school "outputs" are transformed into "outcomes" where the latter includes things like success in securing employment *after* school has been completed.[17] For instance, see C.A.K. Lovell *et al.* (1994)[18] who employ cohort data in such a two-stage analysis where "outputs" are converted to "outcomes" to find that the record for American Public Schools is substantially better for the latter than the former.

3.10 SUMMARY OF CHAPTER 3

In this chapter we described the CCR model in some detail in both its input-oriented and output-oriented versions.

1. We also relaxed assumptions of a positive data set to semipositivity.

2. We defined the production possibility set based on the constant returns-to-scale assumption and developed the CCR model under this assumption.

3. The dual problem of the original CCR model was introduced as (DLP_o) in (3.6)-(3.9) and the existence of input excesses and output shortfalls clarified by solving this model. (To avoid confusion and to align our terminology with the DEA literature, we referred to the dual as the "envelopment model" and the primal (LP_o) introduced in Chapter 2 as the "multiplier model.")

4. In Definition 3.2 we identified a DMU as *CCR-efficient* if and only if it is (i) radial-efficient and (ii) has zero-slack in the sense of Definition 3.1. Hence a DMU is CCR-efficient if and only if it has no input excesses and no output shortfalls.

5. Improvement of inefficient DMUs was discussed and formulae were given for effecting the improvements needed to achieve full CCR efficiency in the form of the CCR-projections given in (3.22) and (3.23).

6. Detailed computational procedures for the CCR model were presented in Section 3.6 and an optimal multiplier values $(\boldsymbol{v},\ \boldsymbol{u})$ were obtained as the simplex multipliers for an optimal tableau obtained from the simplex method of linear programming.

3.11 NOTES AND SELECTED BIBLIOGRAPHY

As noted in Section 3.3 the term "Pareto-Koopmans" efficiency refers to Vilfredo Pareto and Tjalling Koopmans. The former, i.e., Pareto, was concerned with welfare economics which he visualized in terms of a vector-valued function with its components representing the utilities of all consumers. He then formulated the so-called Pareto condition of welfare maximization by noting that such a function could not be at a maximum if it was possible to increase one of its components without worsening other components of such a vector valued function. He therefore suggested this as a criterion for judging any proposed social policy: "the policy should be adopted if it made some individuals better off without decreasing the welfare (measured by their utilities) of other individuals." This, of course, is a necessary but not a sufficient condition for maximizing such a function. Proceeding further, however, would involve comparing utilities to determine whether a decrease in the utilities of some persons would be more than compensated by increases in the utilities of other individuals. Pareto, however, wanted to proceed as far as possible without requiring such interpersonal comparisons of utilities. See Vilfredo Pareto, *Manuel d'economie politique,* deuxieme edition, Appendix, pp. 617 ff., Alfred Bonnet, ed.(Paris: Marcel Giard, 1927).

Tjalling Koopmans adapted these concepts to production. In an approach that he referred to as "activity analysis," Koopmans altered the test of a vector optimum by reference to whether it was possible to increase any output without worsening some other output under conditions allowed by available resources such as labor, capital, raw materials, etc. See pp. 33-97 in T.C. Koopmans, ed., *Activity Analysis of Production and Allocation,* (New York: John Wiley & Sons, Inc., 1951).

The approaches by Pareto and Koopmans were entirely conceptual. No empirical applications were reported before the appearance of the 1957 article by M.J. Farrell in the *Journal of the Royal Statistical Society* under the title "The Measurement of Productive Efficiency." This article showed how these methods could be applied to data in order to arrive at relative efficiency evaluations. This, too, was in contrast with Koopmans and Pareto who conceptualized matters in terms of theoretically known efficient responses without much (if any) attention to how inefficiency, or more precisely, technical inefficiency, could be identified. Koopmans, for instance, assumed that producers would respond optimally to prices which he referred to as "efficiency prices." Pareto assumed that all consumers would (and could) maximize their utilities under the social policies being considered. The latter topic, i.e., the identification of inefficiencies, seems to have been, first, brought into view in an article published as "The Coefficient of Resource Utilization," by G. Debreu in *Econometrica* (1951) pp.273-292. Even though their models took the form of linear programming problems both Debreu and Farrell formulated their models in the tradition of "activity analysis." Little, if any, attention had been paid to computational implementation in the activity analysis literature. Farrell, therefore, undertook a massive and onerous series of matrix inversions in his first efforts. The alternative of linear programming algorithms was called to Farrell's attention by A.J. Hoffman, who served as a commentator in this same issue of the *Journal of the Royal Statistical Society.* Indeed, the activity analysis approach had already been identified with linear programming and reformulated and extended in the 1957 article A. Charnes and W.W. Cooper published in "On the Theory and Computation of Delegation-Type Models: K-Efficiency, Functional Efficiency and Goals," *Proceedings of the Sixth International Meeting of The Institute of Management Science* (London: Pergamon Press). See also Chapter IX in A. Charnes and W.W. Cooper, *Management Models and Industrial Applications of Linear Programming* (New York: John Wiley & Sons, Inc., 1961).

The modern version of DEA originated in two articles by A. Charnes, W.W. Cooper and E. Rhodes: (1) "Measuring the Efficiency of Decision Making Units," *European Journal of Operational Research* 2, 1978, pp.429-444 and (2) "Evaluating Program and Managerial Efficiency: An Application of Data Envelopment Analysis to Program Follow Through," *Management Science* 27, 1981, pp.668-697. The latter article not only introduced the name Data Envelopment Analysis for the concepts introduced in the former article, it also exploited the duality relations as well as the computational power that the former had made available. In addition, extensions were made that included the CCR projection operations associated with (3.32)-(3.33) and used these projections to evaluate programs (such as the U.S. Office of Education's "Program Follow Through") that made it possible to identify a "program's efficiency" separately from the way the programs had been managed. (Hence the distinction between the "program" and "managerial" efficiencies had been confounded in the observations generated from this education study.) Equally important, the latter article started a tradition in which applications were used to guide subsequent research in developing new concepts and methods of analyses. This, in turn, led to new applications, and so on, with many hundreds now reported in the literature.

The first of the above two articles also introduced the ratio form of DEA that is represented as (FP_o) in Section 2.3 of Chapter 2. This ratio form with its enhanced interpretive power and its contacts with other definitions of efficiency in engineering and science was made possible by prior research in which Charnes and Cooper opened the field of fractional programming. See Problem 3.1, below. It also led to other new models and measures of efficiency which differ from (DLP_o) in Section 3.3, above, which had been the only model form that was previously used.

3.12 RELATED DEA-SOLVER MODELS FOR CHAPTER 3

CCR-I (Input-oriented Charnes-Cooper-Rhodes model).

This code solves the CCR model expressed by (3.6)-(3.9) or by (3.38)-(3.42). The data set should be prepared in an Excel Workbook by using an appropriate Workbook name prior to execution of this code. See the sample format displayed in Figure B.1 in Section B.5 of Appendix B and refer to explanations above the figure. This style is the most basic one and is adopted in other models as the standard main body of data. The main results will be obtained in the following Worksheets as displayed in Table 3.6. The worksheet "Summary" includes statistics on data — average, standard deviation of each input and output, and correlation coefficients between observed items. It also reports DMUs with inappropriate data for evaluation and summarizes the results.

CCR-O (Output-oriented CCR model).

This code solves the output-oriented CCR model expressed by (3.60)-(3.63). In this model the optimal efficiency score η^* describes the output enlargement rate and satisfies $\eta^* \geq 1$. However, we display this value by its inverse as $\theta^* = 1/\eta^* (\leq 1)$ and call it the "CCR-O efficiency score." This will facilitate comparisons of scores between the input-oriented and the output-oriented models. In the CCR model, both models have related efficiency values as shown by (3.64). The other results are exhibited in the Worksheets in Table 3.6. In Worksheet "Weight," $\boldsymbol{v}$ and $\boldsymbol{u}$ correspond to $\boldsymbol{p}$ and $\boldsymbol{q}$ in (3.68)-(3.71), respectively. "Projection" is based on the formulas in (3.74) and (3.75).

Table 3.6. Worksheets Containing Main Results

Worksheet name	Contents
Summary	Summary on data and results.
Score	The efficiency score θ^*, the reference set($\boldsymbol{\lambda}^*$), ranking, etc.
Rank	The descending order ranking of efficiency scores.
Weight	The optimal (dual) multipliers $\boldsymbol{v}^*$, $\boldsymbol{u}^*$ in (3.2)-(3.5).
WeightedData	The weighted data $\{x_{ij}v_i^*\}$ and $\{y_{rj}u_r^*\}$.
Slack	The input excesses $\boldsymbol{s}^-$ and the output shortfalls $\boldsymbol{s}^+$ in (3.10).
Projection	Projection onto the efficient frontiers by (3.22)-(3.23).
Graph1	The bar chart of the CCR scores.
Graph2	The bar chart of scores in ascending order.

3.13 PROBLEM SUPPLEMENT FOR CHAPTER 3

Problem 3.1 (Ratio Forms and Strong and Weak Disposal)

The suggested response to Problem 2.3 in Chapter 2 showed how the "ratio of ratios" definition of efficiency in engineering could be subsumed under the CCR ratio form of DEA given for (FP_o) in Section 2.3 of that chapter. Can you now extend this to show how the linear programming formulation given in (LP_o) relates to the ratio form given in (FP_o)?

Suggested Response : A proof of equivalence is provided with Theorem 2.1 in Section 2.4 of Chapter 2. The following proof is adopted from A. Charnes and W.W. Cooper, "Programming with Linear Fractional Functionals" *Naval Research Logistics Quarterly* 9, 1962, pp.181-185, which initiated (and named) the field of fractional programming. We use this as an alternative because it provides contact with the (now) extensive literature on fractional programming. See S. Schaible (1994) "Fractional Programming" in S. Gass and C.M. Harris, eds., *Encyclopedia of Operations Research and Management Science* (Norwell, Mass. Kluwer Academic Publishers) who notes 900 articles that have appeared since Charnes-Cooper (1962). To start we reproduce (FP_o) from section 2.3 of Chapter 2, as follows

$$\begin{aligned}
\max \quad & \theta = \frac{\sum_{r=1}^{s} u_r y_{ro}}{\sum_{i=1}^{m} v_i x_{io}} && (3.91)\\
\text{subject to} \quad & \frac{\sum_{r=1}^{s} u_r y_{rj}}{\sum_{i=1}^{m} v_i x_{ij}} \le 1, \quad j = 1, \ldots, n \\
& u_r, v_i \ge 0. \quad \forall r, i
\end{aligned}$$

Now we can choose a new variable t in such a way that

$$t \sum_{i=1}^{m} v_i x_{io} = 1, \qquad (3.92)$$

which implies $t > 0$. Multiplying all numerators and denominators by this t does not change the value of any ratio. Hence setting

$$\begin{aligned} \mu_r &= tu_r, \quad r = 1, \ldots, s \\ \nu_i &= tv_i, \quad i = 1, \ldots, m \end{aligned} \tag{3.93}$$

we have replaced the above problem by the following equivalent,

$$\begin{aligned} \max \quad & \theta = \sum_{r=1}^{s} \mu_r y_{ro} && (3.94) \\ \text{subject to} \quad & \sum_{i=1}^{m} \nu_i x_{io} = 1 \\ & \sum_{r=1}^{s} \mu_r y_{rj} - \sum_{i=1}^{m} \nu_i x_{ij} \leq 0 \;\; j = 1, ..., n \\ & \mu_r, \; \nu_i \geq 0. \;\; \forall r, \; i \end{aligned}$$

This is the same as (LP_o) in Section 2.4 which we have transformed using what is referred to as the "Charnes-Cooper transformation" in fractional programming. This reduction of (3.91)to the linear programming equivalent in (3.94) also makes available (DLP_o) the dual problem which we reproduce here as

$$\begin{aligned} \min \quad & \theta && (3.95) \\ \text{subject to} \quad & \theta x_{io} = \sum_{j=1}^{n} x_{ij} \lambda_j + s_i^-, \quad i = 1, \ldots, m \\ & y_{ro} = \sum_{j=1}^{n} y_{rj} \lambda_j - s_r^+. \;\; r = 1, \ldots, s \end{aligned}$$

This is the form used by Farrell (1978) which we employed for Phase I, as described in Section 3.3, followed by a Phase II in which the slacks are maximized in the following problem

$$\begin{aligned} \max \quad & \sum_{i=1}^{m} s_i^- + \sum_{r=1}^{s} s_r^+ && (3.96) \\ \text{subject to} \quad & \theta^* x_{io} = \sum_{j=1}^{n} x_{ij} \lambda_j + s_i^-, \quad i = 1, \ldots, m \\ & y_{ro} = \sum_{j=1}^{n} y_{rj} \lambda_j - s_r^+, \quad r = 1, \ldots, s \\ & 0 \leq \lambda_j, s_i^-, s_r^+, \quad \forall j, i, r \end{aligned}$$

where θ^* is the value obtained by solving (3.95) in Phase I. As noted in the text the solution to (3.95) is referred to as "Farrell efficiency." It is also referred to as "weak efficiency," as measured by θ^*, since this measure does not comprehend the non-zero slacks that may be present. In the economics literature (3.95) is said to assume "strong disposal." If we omit the s_i^- in the first m constraints in (3.95) we then have what is called "weak disposal" which we can write in the following form

$$\begin{aligned} \min \quad & \theta \\ \text{subject to} \quad & \theta x_{io} = \sum_{j=1}^{n} x_{ij}\lambda_j, \quad i = 1, \ldots, m \\ & y_{ro} = \sum_{j=1}^{n} y_{rj}\lambda_j - s_r^+, \quad r = 1, \ldots, s \end{aligned} \tag{3.97}$$

which means that the input inequalities are replaced with equalities so there is no possibility of positive input slacks that may have to be disposed of. Sometimes this is referred to as the assumption of "weak" and "strong" *input* disposal in order to distinguish it from corresponding formulations in *output* oriented models. In either case, these weak and strong disposal assumptions represent refinements of the "free disposal" assumption introduced by T.C. Koopmans (1951) for use in activity analysis. This assumption means that there is no cost associated with disposing of excess slacks in inputs or outputs. That is, slacks in the objective are all to be assigned a zero coefficient. Hence, all nonzero slacks are to be ignored, whether they occur in inputs or outputs. For a fuller treatment of "weak" and "strong" disposal, see R. Färe, S. Grosskopf and C.A.K. Lovell, *The Measurement of Efficiency of Production* (Boston: Kluwer Academic Publishers Group, 1985).

Problem 3.2

Can you provide a mathematical formulation that will serve to unify the Phase I and Phase II procedures in a single model?

Suggested Response : One way to do this is to join (3.6) and (3.11) together in a single objective as follows:

$$\begin{aligned} \min \quad & \theta - \varepsilon \left(\sum_{i=1}^{m} s_i^- + \sum_{r=1}^{s} s_r^+ \right) \\ \text{subject to} \quad & \theta x_{io} = \sum_{j=1}^{n} x_{ij}\lambda_j + s_i^-, \quad i = 1, \ldots, m \\ & y_{ro} = \sum_{j=1}^{n} y_{rj}\lambda_j - s_r^+. \quad r = 1, \ldots, s \\ & 0 \leq \lambda_j, s_i^-, s_r^+. \quad \forall j, i, r \end{aligned} \tag{3.98}$$

It is tempting to represent $\varepsilon > 0$ by a small (real) number such as $\varepsilon = 10^{-6}$. However, this is not advisable. It can lead to erroneous results and the situation may be *worsened* by replacing $\varepsilon = 10^{-6}$ by even smaller values. See I. Ali and L. Seiford, "The Mathematical Programming Approach to Efficiency Analysis" in H.O. Fried, C.A.K. Lovell and S.S. Schmidt, ed., *The Measurement of Productive Efficiency* (New York: Oxford University Press, 1993) or Ali and Seiford (1993) "Computational Accuracy and Infinitesimals in Data Envelopment Analysis," *INFOR* 31, pp.290-297.

As formulated in the Charnes, Cooper and Rhodes article in the *European Journal of Operational Research*, cited in the Notes and Selected Bibliography, above, $\varepsilon > 0$ is formulated as a "non-Archimedean infinitesimal." That is, $\varepsilon > 0$ is smaller than *any* positive real number and, in fact, the product of ε by *any* real number, so that, however large the multiplier, $k > 0$, the value of $k\varepsilon > 0$ remains smaller than *any* positive real number. This means that $\varepsilon > 0$ is not a real number because the latter all have the Archimedean property — viz., given any real number $n > 0$ there exists another real number $n/2$ such that $n > n/2 > 0$. Thus, to deal with non-Archimedean elements it is necessary to embed the field of real numbers in a still larger field. However, it is not necessary to go into the further treatments of this kind of (non-standard) mathematics. It is not even necessary to specify a value of $\varepsilon > 0$ explicitly. The two-phase procedure described in this chapter accomplishes all that is required. Phase I accords priority to $\min \theta = \theta^*$ with $\theta^* > 0$ when the data are semipositive. Fixing $\theta = \theta^*$ as is done in (3.12) for Phase II in Section 3.3 of this chapter, we must then have, by definition,

$$0 < \theta^* - \varepsilon \left(\sum_{i=1}^{m} s_i^{-*} + \sum_{r=1}^{s} s_r^{+*} \right).$$

A more detailed treatment of these non-Archimedean elements and their relations to mathematical programming may be found in V. Arnold, I. Bardhan, W.W. Cooper and A. Gallegos, "Primal and Dual Optimality in Computer Codes Using Two-Stage Solution Procedures in DEA" in J.E. Aronson and S. Zionts, ed., *Operations Research: Methods, Models and Applications* (Westport, Conn.: Quorum Books, 1998). Here we only note that the further implications of this non-Archimedean element, which appears in the objective of (3.98) can be brought forth by writing its dual as

$$\begin{aligned}
\max \quad & \theta = \sum_{r=1}^{s} \mu_r y_{ro} && (3.99) \\
\text{subject to} \quad & \sum_{r=1}^{s} \mu_r y_{rj} - \sum_{i=1}^{m} \nu_i x_{ij} \leq 0 \quad j = 1, \ldots, n \\
& \sum_{i=1}^{m} \nu_i x_{io} = 1 \\
& \mu_r,\ \nu_i \geq \varepsilon > 0. \quad \forall r,\ i
\end{aligned}$$

This means that all variables are constrained to positive values. Hence, they are to be accorded "some" worth, even though it is not specified explicitly.

Finally, as shown in Arnold *et al.*, the principle of complementary slackness used in Section 3.3 of this chapter is modified to the following

$$0 \le s_i^{-*}\nu_i^* \le s_i^{-*}\varepsilon \tag{3.100}$$
$$0 \le s_r^{+*}\mu_r^* \le s_r^{+*}\varepsilon. \tag{3.101}$$

Hence, unlike what is done with free disposal, one cannot justify ignoring nonzero slacks by assuming that corresponding multiplier values will be zero. Indeed, the stage two optimization maximizes the slacks, as is done in the Phase II procedure associated with (3.96), in order to try to wring out the maximum possible inefficiency values associated with nonzero slacks.

Problem 3.3

We might note that (3.99), above, differs from (3.94) by its inclusion of the non-Archimedean conditions $\mu_r, \nu_i \ge \varepsilon > 0$. Can you show how these conditions should be reflected in a similarly altered version of the ratio model in (3.91)?

Suggested Answer : Multiply and divide the objective in (3.99) by $t > 0$. Then multiply all constraints by the same t to obtain

$$\max \quad \frac{\sum_{r=1}^{s}(t\mu_r)y_{ro}}{t} \tag{3.102}$$
$$\text{subject to} \quad \sum_{r=1}^{s}(t\mu_r)y_{rj} - \sum_{i=1}^{m}(t\nu_i)x_{ij} \le 0 \quad j = 1,\ldots,n$$
$$\sum_{i=1}^{m}(t\nu_i x_{io}) = t$$
$$(t\mu_r),\ (t\nu_i) \ge \varepsilon > 0. \ \forall r,\ i$$

Set $u_r = t\mu_r$, $v_i = t\nu_i$ for each r and i. Then substitute in these last expressions to obtain

$$\max \quad \frac{\sum_{r=1}^{s} u_r y_{ro}}{\sum_{i=1}^{m} v_i x_{io}} \tag{3.103}$$
$$\text{subject to} \quad \frac{\sum_{r=1}^{s} u_r y_{rj}}{\sum_{i=1}^{m} v_i x_{ij}} \le 1, \quad j = 1,\ldots,n$$
$$u_r \Big/ \sum_{i=1}^{m} v_i x_{io} \ge \varepsilon \quad \forall r$$
$$v_i \Big/ \sum_{i=1}^{m} v_i x_{io} \ge \varepsilon. \quad \forall i$$

Note that we have reflected the condition $\sum_{i=1}^{m} v_i x_{io} = t$ in the above objective but not in the constraints because it can *always* be satisfied. We can also now write

$$\frac{\sum_{r=1}^{s} u_r^* y_{ro}}{\sum_{i=1}^{m} v_i^* x_{io}} = \sum_{r=1}^{s} u_r^* y_{ro} = \theta^* - \varepsilon\left(\sum_{i=1}^{m} s_i^{-*} + \sum_{r=1}^{s} s_r^{+*}\right). \tag{3.104}$$

The equality on the right follows from the dual theorem of linear programming. We have added the equality on the left by virtue of our derivation and, of course, the values of these expressions are bounded by zero and one.

Problem 3.4

The optimal solution for (LP_o) , the multiplier problem, for DMU A in Example 3.1 is

$$v_1^* = 0.1429, v_2^* = 0.1429, u^* = 0.8571.$$

Show how these multiplier values could be used as a management guide to improve efficiency in A's performance.

Suggested Response : Applying the above solution to the data in Table 3.2 we utilize the multiplier model to write the constraints for DMU A as follows:

	DMU	Output Value	Output Variable	Input Variable	Input Value
	A	0.8571	$= 1u^*$	$\le 4v_1^* + 3v_2^*$	=1.000
	B	0.8571	$= 1u^*$	$\le 7v_1^* + 3v_2^*$	=1.429
	C	0.8571	$= 1u^*$	$\le 8v_1^* + v_2^*$	=1.284
$\longrightarrow$	D	0.8571	$= 1u^*$	$\le 4v_1^* + 2v_2^*$	=0.8571
$\longrightarrow$	E	0.8571	$= 1u^*$	$\le 2v_1^* + 4v_2^*$	=0.8571
	F	0.8571	$= 1u^*$	$\le 10v_1^* + v_2^*$	=1.5719
	G	0.8571	$= 1u^*$	$\le 3v_1^* + 7v_2^*$	=1.429

$$\text{Add constraint : } \quad 1 = 4v_1^* + 3v_2^*.$$

Note, first, that $1u^* = \theta^* = 0.8571$ so the value of the optimal solution of the multiplier model is equal to the optimal value of the envelopment model in accordance with the duality theory of linear programming and, similarly, the reference set is $\{D, E\}$, as indicated by the arrows. Here we have $v_1^* = v_2^* = 0.1429$ so a unit reduction in the $4+3 = 7$ units of input used by A would bring it into $4 + 2 = 6$ units used by D and E. Thus, $v^* = v_1^* = v_2^* = 0.1429 = 1/7$ is the reduction required to bring A to full efficiency. We can provide added perspective by reducing $x_1 = 4$ to $\widehat{x}_1 = 3$ which positions A half way between D and E on the efficient frontier in Figure 3.2. It also produces equality of the relations for A as in the following expression

$$u^* = 3v_1^* + 3v_2^* \doteq 0.8575.$$

Division then produces

$$\frac{u^*}{3v_1^* + 3v_2^*} = 1,$$

which is the condition for efficiency prescribed for (FP_o) as given in (2.3)-(2.6) in Chapter 2. Here we have $v_1^* = v_2^* = 0.1429$ so a use of either x_1 or x_2 produces equal results per unit change en route to achieving full efficiency. More generally we will have $v_1^* \neq v_2^*$, etc., so these values may be used to guide managerial priorities for the adjustments to the model.

There are two exceptions that require attention. One exception is provided by Definition 2.1 in Chapter 2, which notes that all components of the vectors $\boldsymbol{v}^*$

and $\boldsymbol{u}^*$ must be positive in at least one optimal solution. We use the following solution for G as an illustration,

$$v_1^* = 0.3333, \quad v_2^* = 0, \quad u^* = 0.6666.$$

From the data of Example 3.1, we then have

$$0.6666 = 1u^* < 3v_1^* + 7v_2^* = 1.00.$$

Replacing $x_1 = 3$ by $\widehat{x}_1 = 2$ produces equality, but efficiency cannot be claimed because $v_2^* = 0$ is present.

Reference to Figure 3.2 shows that the inefficiency associated with the nonzero slack in going from R to E is not attended to and, indeed, the amount of slack associated with x_2 is worsened by the horizontal movement associated with the reduction in going from $x_1 = 3$ to $\widehat{x}_1 = 2$. Such a worsening would not occur for F in Figure 3.2, but in either case full efficiency associated with achieving a ratio of unity does not suffice unless all multipliers are positive in the associated adjustments.

The second exception noted above involves evaluations of the DMUs that are fully efficient and hence are used to evaluate *other* DMUs. This complication is reflected in the fact that the data in the basis matrix B is changed by any such adjustment and so the inverse B^{-1} used to evaluate other DMUs, as given in (3.53), is also changed. This topic involves complexities which cannot be treated here, but is attended to later in this book after the requisite background has been supplied. It is an important topic that enters into further uses of DEA and hence has been the subject of much research which we summarize as follows.

The problem of sensitivity to data variations was first addressed in the DEA literature by A. Charnes, W.W. Cooper, A.Y. Lewin, R.C. Morey and J. Rousseau "Sensitivity and Stability Analysis in DEA," *Annals of Operations Research* 2, 1985, pp.139-156. This paper was restricted to sensitivity analyses involving changes in a single input or output for an efficient DMU. Using the CCR ratio model, as in (3.91), above, this was subsequently generalized to simultaneous changes in all inputs and outputs for any DMU in A. Charnes and L. Neralic, "Sensitivity Analysis of the Proportionate Change of Inputs (or Outputs) in Data Envelopment Analysis," *Glasnik Matematicki* 27, 1992, pp.393-405. Also using the ratio form of the multiplier model, R.G. Thompson and R.M. Thrall provided sensitivity analyses in which *all* data for *all* DMUs are varied simultaneously. The basic idea is as follows. All inputs and all outputs are worsened for the efficient DMUs until at least one DMU changes its status. The multipliers and ratio values are then recalculated and the process is reiterated until all DMUs become inefficient. See R.G. Thompson, P.S. Dharmapala, J. Diaz, M. Gonzalez-Lima and R.M. Thrall, "DEA Multiplier Analytic Center Sensitivity Analysis with an Illustrative Application to Independent Oil Companies," *Annals of Operations Research* 66, 1996, pp.163-167. An alternate and more exact approach is given in L. Seiford and J. Zhu, "Sensitivity Analysis of DEA Models for Simultaneous Changes in All the Data," *Journal of the Operational Research Society* 49, 1998, pp.1060-1071. Finally, L. Seiford and

J. Zhu in "Stability Regions for Maintaining Efficiency in Data Envelopment Analysis," *European Journal of Operational Research* 108, 1998, pp. 127-139 develop a procedure for determining exact stability regions within which the efficiency of a DMU remains unchanged. More detail on the topic of sensitivity is covered in Chapter 9. See also W.W. Cooper, L.M. Seiford and J. Zhu (2004) Chapter 3, "Sensitivity Analysis in DEA," in W.W. Cooper, L.M. Seiford and J. Zhu, eds., *Handbook on Data Envelopment Analysis* (Norwell, Mass., Kluwer Academic Publishers).

Problem 3.5 (Complementary Slackness and Sensitivity Analysis)

Table 3.7 displays a data set of 5 stores (A, B, C, D and E) with two inputs (the number of employees and the floor area) and two outputs (the volume of sales and profits).

Assignment: (i) Using the code CCR-I in the DEA-Solver, obtain the efficiency score (θ^*), reference set ($\boldsymbol{\lambda}^*$), the optimal weights ($\boldsymbol{v}^*, \boldsymbol{u}^*$) and slacks ($\boldsymbol{s}^{-*}, \boldsymbol{s}^{+*}$) for each store. (ii) Interpret the complementary slackness conditions (3.15) between the optimal weights and the optimal slacks. (iii) Discuss the meaning of u_1^* of C, and use sensitivity analysis to verify your discussion. (iv) Check the output shortfall s_2^{+*} of B and identify the influence of its increase on the value of the corresponding multiplier (weight) u_2 using the CCR-I code. (v) Regarding the store B, discuss the meaning of v_1^*.

Suggested Response : (i) The efficiency score (θ^*) and reference set ($\boldsymbol{\lambda}^*$) are listed in Table 3.7 under the heading "Score." Store E is the only efficient DMU and is referent to all other stores. The optimal weights and slacks are displayed in Table 3.8.

Table 3.7. Data and Scores of 5 Stores

	Data				Score		
Store	Employee	Area	Sales	Profits	θ^*	Rank	Reference
A	10	20	70	6	0.933333	2	E (0.77778)
B	15	15	100	3	0.888889	3	E (1.11111)
C	20	30	80	5	0.533333	5	E (0.88889)
D	25	15	100	2	0.666667	4	E (1.11111)
E	12	9	90	8	1	1	E (1)

(ii) The complementary slackness conditions in (3.15) assert that, for an optimal solution we have

$$v_i^* s_i^{-*} = 0 \text{ (for } i = 1, 2) \text{ and } u_r^* s_r^{+*} = 0 \text{ (for } r = 1, 2) .$$

Table 3.8. Optimal Weights and Slacks

	Weights				Slacks			
Store	Emply. v_1^*	Area v_2^*	Sales u_1^*	Profits u_2^*	Emply. s_1^{-*}	Area s_2^{-*}	Sales s_1^{+*}	Profits s_s^{+*}
A	0.1	0	0.0133	0	0	11.6667	0	0.2222
B	0.0667	0	0.00889	0	0	3.3333	0	5.8889
C	0.05	0	0.00667	0	0	8	0	2.1111
D	0	0.0667	0.00667	0	3.3333	0	0	6.8889
E	0.0702	0.0175	0.00526	0.0658	0	0	0	0

This means that if $v_i^* > 0$ then $s_i^{-*} = 0$ and if $s_i^{-*} > 0$ then $v_i^* = 0$, and the same relations hold between $\boldsymbol{u}^*$ and $\boldsymbol{s}^{+*}$. This can be interpreted thus: if a DMU has an excess (s^{-*}) in an input against the referent DMU, then the input item has no value in use to the DMU$_o$ being evaluated, so the optimal (multiplier) solution assigns a value of zero to the corresponding weight (v^*). Similarly we can see that for a positive weight to be assigned to a multiplier, the corresponding optimal slack must be zero.
(iii) C has $u_1^* = 0.00667$ and this value can be interpreted in two ways. First, with recourse to the fractional program described by (2.3)-(2.6) in Chapter 2, we can interpret u_1^* (the optimal weight to "sales") as the degree of influence which one unit of sales has on the optimal efficiency score. Thus, if C increases its sales by one unit, then it is expected that θ^* will increase by 0.00667 ($=u_1^*$) to 0.533333+0.00667=0.54, since the denominator of (2.3) does not change and retains the value 1. On the other hand, $u_2^* = 0$ implies that a one unit increase in "profits" has no effect on the efficiency score of C. The second interpretation is that in (3.59) we introduced $\boldsymbol{u}^*$ as "pricing vectors" and in the envelopment form of LP, u_r^* is the reduced cost induced by one unit change in output r. Notice that $\boldsymbol{y}_o$ appears as the constant in (3.41). Hence $u_1^* = 0.00667$ shows the degree of contribution that one unit change of sales can make to use the efficiency score. The above observations contribute to identifying which output has the largest influence on the efficiency score. However, such analysis of influence is valid only in a limited range of observations. To check this range you can change the efficiency score of C by adding to the data set a virtual store C', with sales of 101 and other data equal to C, and then applying the CCR-I code again.
(iv) B has the slack $s_2^{+*} = 5.889$ (shortfall in "profits") and, by the complementary slackness theorem, $u_2^{+*} = 0$, showing that this level of profits has no effect on the efficiency evaluation. Adding this value to current profits will raise profits to 8.889. We augmented the data set by adding B', which has profits of 8.889, and other observations equal to B and tried CCR-I for this data set. As a result, we have, for B', $s_2^{+*} = 0$ and $u_2^{+*} = 0.1$. This means that, at this level

of profits, this output has a positive influence on the efficiency score, which in this case is a 0.1 per unit increase.
(v) B has the optimal weight $v_1^* = 0.0667$. It can therefore be said that changes in the number of employees affects the efficiency score. In this case, reduced costs analysis using LP is not as straightforward as in u^*. Hence, dealing with the fractional program in (2.3)-(2.6) is appropriate for the sensitivity analysis of data on employees. If B reduces its employees by one to 14, then the denominator of (2.3) decreases by $v_1^* = 0.0667$. Since the other terms are unchanged, the objective function (2.3) takes the value

$$(0.00889 \times 100 + 0 \times 3)/(0.0667 \times 14 + 0 \times 15) = 0.952.$$

Since the value is still less than one and all constraints in (2.4) are satisfied by this set of $(\boldsymbol{v}^*, \boldsymbol{u}^*)$, the above objective value gives a lower bound of θ^* for this adjusted problem. Application of CCR-I code for this new problem showed that 0.952 is the optimal objective value. Thus, it may be expected that an input item with a large v_i^* value greatly affects the efficiency value.

Problem 3.6

Solve Example 3.1 in Table 3.2 (Section 3.7) using the output-oriented model (CCR-O model in DEA-Solver) and compare the corresponding CCR-projection with the input-oriented case.

Suggested Response : For ease of comparison with the input-oriented case, the optimal objective value η^* is demonstrated by its inverse θ^* $(= 1/\eta^*)$.

The optimal solution to A reads

$$\begin{aligned}
&\theta^* = 0.8571 \quad (\eta^* = 1.1667) \\
&\lambda_D^* = 0.8333, \ \lambda_E^* = 0.3333, \ \text{other } \lambda_j^* = 0 \\
&s_1^{-*} = s_2^{-*} = s^{+*} = 0
\end{aligned}$$

The output-oriented CCR-projection was performed following (3.74)-(3.75), which resulted in

$$\begin{aligned}
&\widehat{x}_1 \leftarrow x_1 = 4 \quad \text{(no change)} \\
&\widehat{x}_2 \leftarrow x_2 = 3 \quad \text{(no change)} \\
&\widehat{y} \leftarrow \eta^* y = 1.167 \times 1 = 1.167 \quad \text{(16.7\% increase)}.
\end{aligned}$$

These results differ from the input-oriented case, reflecting the difference in model orientation. Table 3.9 exhibits the CCR-projection in both input and output orientations.

Problem 3.7

Suppose the activity $(\boldsymbol{x}_o,\ \boldsymbol{y}_o)$ is inefficient. Let the improved activity obtained by the input-oriented model (3.22) and (3.23) be $(\widehat{\boldsymbol{x}}_o,\ \widehat{\boldsymbol{y}}_o)$,

$$\widehat{\boldsymbol{x}}_o = \theta^* \boldsymbol{x}_o - \boldsymbol{s}^{-*} = X\boldsymbol{\lambda}^* \tag{3.105}$$

Table 3.9. CCR-projection in Input and Output Orientations

DMU	Data			Score	Input orientation			Output orientation		
	x_1	x_2	y	θ^*	x_1	x_2	y	x_1	x_2	y
A	4	3	1	0.8571	3.43	2.57	1	4	3	1.17
B	7	3	1	0.6316	4.42	1.89	1	7	3	1.58
C	8	1	1	1.0000	8	1	1	8	2	1
D	4	2	1	1.0000	4	2	1	4	2	1
E	2	4	1	1.0000	2	4	1	2	4	1
F	10	1	1	1.0000	8	1	1	8	1	1
G	3	7	1	0.6667	2	4	1	3	6	1.5

$$\widehat{\boldsymbol{y}}_o = \boldsymbol{y}_o + \boldsymbol{s}^{+*} = Y\boldsymbol{\lambda}^*, \tag{3.106}$$

while that of the output-oriented model (3.74) and (3.75) be $(\breve{\boldsymbol{x}}_o,\ \breve{\boldsymbol{y}}_o)$.

$$\breve{\boldsymbol{x}}_o = \boldsymbol{x}_o - \boldsymbol{t}^{-*} = X\boldsymbol{\mu}^* \tag{3.107}$$

$$\breve{\boldsymbol{y}}_o = \eta^*\boldsymbol{y}_o + \boldsymbol{t}^{+*} = Y\boldsymbol{\mu}^*. \tag{3.108}$$

An activity on the line segment connecting these two points can be expressed by

$$(\boldsymbol{x}'_o,\ \boldsymbol{y}'_o) = \alpha_1\ (\widehat{\boldsymbol{x}}_o,\ \widehat{\boldsymbol{y}}_o)\ +\ \alpha_2\ (\breve{\boldsymbol{x}}_o,\ \breve{\boldsymbol{y}}_o) \tag{3.109}$$

$$\alpha_1 + \alpha_2 = 1,\quad \alpha_1 \geq 0,\quad \alpha_2 \geq 0.$$

Prove the following

Proposition 3.1 *The activity $(\boldsymbol{x}'_o,\ \boldsymbol{y}'_o)$ is CCR-efficient.*

Suggested Answer :
Proof. Since by (3.65) $\boldsymbol{\mu}^* = \boldsymbol{\lambda}^*/\theta^*$, we have

$$\begin{aligned}(\boldsymbol{x}'_o,\ \boldsymbol{y}'_o) &= \alpha_1\ (X,\ Y)\boldsymbol{\lambda}^*\ +\ \alpha_2\ (X,\ Y)\boldsymbol{\lambda}^*/\theta^* \\ &= (\alpha_1 + \alpha_2/\theta^*)\ (X\boldsymbol{\lambda}^*,\ Y\boldsymbol{\lambda}^*)\end{aligned}$$

By applying Theorem 3.4, we can see that $(\boldsymbol{x}'_o,\ \boldsymbol{y}'_o)$ is CCR-efficient. □
This proposition is valid for *any* nonnegative combination of two points ($\alpha_1 \geq 0$ and $\alpha_2 \geq 0$). That is, it holds even when relaxing the convex-combination condition $\alpha_1 + \alpha_2 = 1$. Also, the case $\alpha_1 = 1,\ \alpha_2 = 0$ corresponds to the input-oriented improvement and the case $\alpha_1 = 0,\ \alpha_2 = 1$ to the output-oriented one. The case $\alpha_1 = 1/2,\ \alpha_2 = 1/2$ is a compromise between the two points.

Problem 3.8

Solve the hospital example in Chapter 1 (see Table 1.5) by input-oriented and output-oriented CCR models. Compare the CCR-projections of hospital C in both cases. Based on the previous Proposition 3.1, determine the combination of both projections using (3.109) with $\alpha_1 = \alpha_2 = 1/2$.

Suggested Answer : The data sheet for this problem is recorded in Figure B.1 in Appendix B. Prepare two Excel Workbooks, e.g., "Hospital-CCR-I.xls" and "Hospital-CCR-O.xls," each containing Figure B.1 as the data sheet. Then run CCR-I (the input-oriented CCR model) and CCR-O (the output-oriented CCR model). After the computations, the CCR-projections are stored in "Projection" sheets as follows:
For input orientation:

Doctor	25	$\rightarrow$	20.9	(16% reduction)
Nurse	160	$\rightarrow$	141	(12% reduction)
Outpatient	160	$\rightarrow$	160	(no change)
Inpatient	55	$\rightarrow$	55	(no change)

For output orientation:

Doctor	25	$\rightarrow$	23.6	(5% reduction)
Nurse	160	$\rightarrow$	160	(no change)
Outpatient	160	$\rightarrow$	181	(13% increase)
Inpatient	55	$\rightarrow$	62.3	(13% increase)

We have a compromise of the two models as the averages of the projected values as follows (this is equivalent to the case $\alpha_1 = \alpha_2 = 1/2$ in Proposition 3.1):

Doctor	25	$\rightarrow$	22	(12% reduction)
Nurse	160	$\rightarrow$	150	(6% reduction)
Outpatient	160	$\rightarrow$	171	(7% increase)
Inpatient	55	$\rightarrow$	59	(7% increase)

These improvements will put C on the efficient frontier.

Problem 3.9

Assume that the data set (X, Y) is semipositive (see the definition in Section 3.2) and let an optimal solution of (DLP_o) in Section 3.3 be $(\theta^*, \boldsymbol{\lambda}^*, \boldsymbol{s}^{-*}, \boldsymbol{s}^{+*})$. Prove that the reference set defined by $E_o = \{j|\lambda^* > 0\}$ $(j \in \{1, \ldots, n\})$ is not empty.
Suggested Answer :
We have the equation:

$$\theta^* \boldsymbol{x}_o = X\boldsymbol{\lambda}^* + \boldsymbol{s}^{-*}.$$

Let an optimal solution of (LP_o) in Section 3.3 be $(\boldsymbol{v}^*, \boldsymbol{u}^*)$. By multiplying $\boldsymbol{v}^*$ from left to the above equation, we have

$$\theta^* \boldsymbol{v}^* \boldsymbol{x}_o = \boldsymbol{v}^* X\boldsymbol{\lambda}^* + \boldsymbol{v}^* \boldsymbol{s}^{-*}.$$

Since $\boldsymbol{v}^* \boldsymbol{x}_o = 1$ is a constraint in (LP_o) and $\boldsymbol{v}^* \boldsymbol{s}^{-*} = 0$ by the complementarity condition, it holds

$$\theta^* = v^* X \lambda^*.$$

Here, θ^* is positive as claimed in Section 3.3. Thus, λ^* must be semipositive and E_o is not empty. □

Problem 3.10

Part 1. Using model (3.95) show that the output inequality will always be satisfied as an equation in the single output case. That is, the output slack will be zero at an optimum.

Proof :

It is easy to see what is happening if we start with the case of no outputs. In this case the solution will be $\theta_o^* = 0$ and all $\lambda_j^* = 0$ since only the output constraints keep this from happening. It follows that for the case of one output a solution with $y_o < \sum_{j=1}^n y_j \lambda_j$ cannot be optimal since this choice of λs would prevent θ from achieving its minimum value. To see that this is so note that a choice of the minimizing value of θ is determined by

$$\theta = \max_i \left\{ \frac{\sum_{j=1}^n x_{ij}\lambda_j}{x_{io}} | i = 1, \ldots, m \right\} = \frac{\sum_{j=1}^n x_{kj}\lambda_j}{x_{ko}}$$

and this maximum value can be lowered until $y_o = \sum_{j=1}^n y_j \lambda_j$. (Here for simplicity we are assuming that all data are positive.) Hence optimality requires $y_o = \sum_{j=1}^n y_j \lambda_j^*$ so the output slack is zero in an optimum solution. □

Part 2. Extend the above to show that at least one output inequality must be satisfied as an equation in the case of multiple outputs $r = 1, \ldots, s$, like the ones represented in (3.95).

Proof :

Replace (3.95) by the following model

$$\min_{\lambda, \theta} \max_r \left\{ \theta, \sum_{j=1}^n y_{rj}\lambda_j \geq y_{ro} | r = 1, \ldots, s \right\} = \left(\theta^*, \sum_{j=1}^n y_{kj}\lambda_j^* = y_{ko} \right)$$

$$\text{subject to} \tag{3.110}$$

$$\theta x_{io} \geq \sum_{j=1}^n x_{ij}\lambda_j, \ i = 1, \ldots, m$$

$$0 \leq \theta \leq 1, \ \lambda_j \geq 0, \ j = 1, \ldots, n.$$

This is a multiple criteria programming problem which seeks to minimize the maximum of the $r = 1, \ldots, s$ inequalities representing the output values set by the y_{ro} as lower limits for each of the s outputs. Note that the maximum output values are limited by the input constraints. The condition $0 \leq \theta \leq 1$ which eliminates the possibility of infinite solutions is not needed in (3.95) since its minimizing objective guarantees its fulfillment.

Now the maximal values may be decreased in (3.110) until equality is achieved in at least one of the output inequalities and the minimization of θ eliminates the possibility of being misled by the choice of a set of $\boldsymbol{\lambda}$ values that does not minimize this maximum because of the presence of alternate optima. The optimizing values $\theta^*, \boldsymbol{\lambda}^*$ represented on the right of (3.110) therefore minimize θ with

$$\theta^* = \max_i \left\{ \frac{\sum_{j=1}^n x_{ij}\lambda_j^*}{x_{io}},\ i = 1, \ldots, m \right\}$$

and all constraints are satisfied in (3.95) as well as in (3.110). □

Corollary 3.2 *At least one input as well as one output constraint will be satisfied as an equality at an optimum so these constraints have zero slack in (3.95) and (3.110).*

We also have following

Theorem 3.6 *Model (3.110) is equivalent to model (3.95).*

Similar results hold for the "output-oriented" version of the CCR model with a "Farrell" measure of efficiency that is represented by the value of θ^* in (3.95). Such equivalences may also be established for other DEA models, like the BCC model that will be presented in the next chapter. Finally, we also note that we earlier established the relation of (3.95) to a fractional programming model so our linear programming DEA formulation provides a link between this multiple objective nonlinear programming problem and a nonlinear nonconvex fractional programming problem. See also Cooper (2005)[19] for a discussion of relations between multiple criteria programming and goal programming formulations which are also equivalent to a nonlinear problem directed to minimizing a sum of absolute values.

Notes

1. It might be noted that Postulate (A2) is included in (A4) but is separated out for special attention.

2. See Appendix A.4.

3. See Appendix A.4.

4. Farrell also restricted his treatments to the single output case. See M.J. Farrell (1957) "The Measurement of Production Efficiency, " *Journal of the Royal Statistical Society* A, 120, pp.253-281.

5. In the linear programming literature this is called the "complementary slackness" condition. This terminology is due to A.W. Tucker who is responsible for formulating and proving this. See E.D. Nering and A.W. Tucker *Linear Programs and Related Problems,* (Harcourt Brace, 1993). See our Appendix A.6 for a detailed development.

6. See Appendix A.8.

7. Although empirical studies show the uniqueness of the reference set for most DMUs, there may be multiple reference sets and improvement plans (projections) in the presence of multiple optimal solutions.

8. See Appendix A.4.

9. We notice that OQ and OA are measured by some "distance measure." If we employ the "Euclidian measure" – also called the "l_2 metric" – we have

$$\frac{d(OQ)}{d(OA)} = \frac{\sqrt{3.428^2 + 2.571^2}}{\sqrt{4^2 + 3^2}} = \frac{\sqrt{18.36}}{\sqrt{25}} = 8.57.$$

However, the measure is not restricted to Euclidian measure. Any l_k measure gives the same result. See Appendix A in Charnes and Cooper, *Management Models and Industrial Applications of Linear Programming* (New York, John Wiley, Inc., 1961). See also W.W. Cooper, L.M. Seiford, K. Tone and J. Zhu "DEA: Past Accomplishments and Future Prospects," *Journal of Productivity Analysis* (submitted, 2005).

10. "Aborts" were treated as reciprocals so that an increase in their output would reduce the value of the numerator in the (FP_o) objective represented in (2.3). They could also have been subtracted from a dominatingly large positive constant. See A. Charnes, T. Clark, W.W. Cooper and B. Golany "A Development Study of Data Envelopment Analysis in Measuring the Efficiency of Maintenance Units in the U.S. Air Forces," *Annals of Operational Research* 2, 1985, pp.59-94.

11. R.D. Banker and R.C. Morey (1986), "Efficiency Analysis for Exogenously Fixed Inputs and Outputs," *Operations Research* 34, 1986, pp.513-521. See also Chapter 10 in R. Färe, S. Grosskopf and C.A. Knox Lovell *Production Frontiers* (Cambridge University Press, 1994) where this is referred to as "sub-vector optimizations."

12. A fuller treatment of these $\varepsilon > 0$ values is provided in Problem 3.2 at the end of this chapter.

13. I.R. Bardhan (1995), "DEA and Stochastic Frontier Regression Approaches Applied to Evaluating Performances of Public Secondary Schools in Texas," Ph.D. Thesis. Austin Texas: Graduate School of Business, the University of Texas at Austin. Also available from University Microfilms, Inc. in Ann Arbor, Michigan.

14. V.L. Arnold, I.R. Bardhan and W.W. Cooper "A Two-Stage DEA Approach for Identifying and Rewarding Efficiency in Texas Secondary Schools" in W.W. Cooper, S. Thore, D. Gibson and F. Phillips, eds., IMPACT: *How IC^2 Research Affects Public Policy and Business Practices* (Westport, Conn.: Quorum Books, 1997)

15. An additional category may also be identified as *Property* which applies to (a) Choice of goals or objectives, and (b) Choice of means to achieve goals or objectives.

16. See Arnold *et al.* (1997) for further suggestions and discussions.

17. See Governmental Accounting Standards Board (GASB) Research Report: "Service Efforts and Accomplishments Reporting: Its Time Has Come," H.P. Hatry, J.M. Sullivan, J.M. Fountain and L. Kremer, eds. (Norwell, Conn., 1990).

18. C.A.K. Lovell, L.C. Walters and Lisa Wood, "Stratified Models of Education Production Using Modified DEA and Regression Analysis," in A. Charnes, W.W. Cooper, A.Y. Lewin and L.M. Seiford, eds., *Data Envelopment Analysis: Theory, Methodology and Applications* (Norwell, Mass.: Kluwer Academic Publishers, 1994).

19. W.W. Cooper (2005), "Origin, Uses of and Relations Between Goal Programming and DEA (Data Envelopment Analysis)" *Journal of Multiple Criteria Decision Analysis* (to appear).