A Simple Characterization of Returns to Scale in DEA

Kaoru Tone

Graduate School of Policy Science
Saitama University
Urawa, Saitama 338, Japan

tel: 81-48-858-6096
fax: 81-48-852-0499
e-mail: tone@poli-sci.saitama-u.ac.jp

A Simple Characterization of Returns to Scale in DEA

Kaoru Tone*

February 10, 1995

Abstract

In this paper, we will present a simple method for deciding the local returns-to-scale characteristics of DMUs (Decision Making Units) in Data Envelopment Analysis. This method proceeds as follows: first, we solve the BCC (Banker-Charnes-Cooper) model and find the returns-to-scale of BCC-efficient DMUs and a reference set to each BCC-inefficient DMU. We can then decide the local returns-to-scale characteristics of each BCC-inefficient DMU by observing only the returns-to-scale characteristics of DMUs in their respective reference sets. No extra computation is required. We can also apply this method to the output oriented model.

Keywords: DEA, returns to scale, computation

1 Introduction

The standard model for analyzing returns-to-scale in DEA was first proposed by Banker (1980) and subsequently, Banker, Charnes and Cooper (1984), Banker (1984) and Färe, Grosskopf and Lovell (FGL, 1985) have extended its contents and analysis substantially.

*Graduate School of Policy Science, Saitama University, Urawa, Saitama 338, Japan. e-mail: tone@poli-sci.saitama-u.ac.jp

Then, Banker and Thrall (BT, 1992) presented extensive research with respect to returns-to-scale in DEA. Recently, Banker, Bardham and Cooper (BBC, 1995) added computational convenience and efficiency to the work of Banker and Thrall (BT, 1992).

Although the concept of returns-to-scale is unambiguous only at points on the efficient sections of the production frontier, several pieces of research extended this concept to inefficient DMUs by moving them to the efficient frontiers. Naturally, in this case, returns-to-scale depend on the method used to bring such DMUs to efficient frontiers. FGL (1985) and Banker, Chang and Cooper (BCC, 1995) address this subject. Both methods employ a two-step approach to estimate returns-to-scale. Specifically, FGL(1985) solve, in step 1, the BCC and CCR models and in step 2 solve a linear program for each nonconstant returns-to-scale DMU. Therefore, they need to solve three LPs for nonconstant returns-to-scale DMUs. On the other hand, BCC(1995) solve the CCR model in step 1 and then an LP further for each DMU with nonconstant returns-to-scale characteristics in step 2.

The method that we will propose in this paper solves the BCC model and then determines returns-to-scale of BCC-efficient DMUs. It will be demonstrated that returns-to-scale of projected BCC-inefficient DMUs can be determined automatically from their reference set. From the computational point of view, one pass of BCC computation and BT (1992) process for BCC-efficient DMUs are all that is needed for estimating returns-to-scale.

The rest of the paper is organized as follows. Sections 2 and 3 offer preliminary information, concerned with definitions and known theorems that will be used in the succeeding section. Section 4 is the main part of the paper, in which an alternative method will be presented. In Section 5, we will exhibit a numerical example of our method and then a similar analysis will be presented for the output oriented case. Finally, comparisons with other methods will be discussed in Section 7.

2 The CCR and BCC Models

We will deal with n DMUs (Decision Making Units) with the input and output matrices $X = (x_j) \in R^{m \times n}$ and $Y = (y_j) \in R^{s \times n}$, respectively. We assume that x_j and y_j are *semipositive, i.e.* $x_j \geq 0, x_j \neq 0$ and $y_j \geq 0, y_j \neq 0$ for $j = 1, \ldots, n$.

2.1 The CCR Model

The production possibility set P_C of the CCR (Charnes, Cooper and Rhodes, 1978) model is defined as a set of semipositive (x, y) as follows:

$$P_C = \{(x, y) \mid x \geq X\lambda,\ y \leq Y\lambda,\ \lambda \geq 0\}, \tag{1}$$

where λ is a semipositive vector in R^n.
The CCR model evaluates the efficiency of each DMU_o (x_o, y_o) $(o = 1, \ldots, n)$ by solving the following linear program:

$$(CCR_o) \qquad \min\ \theta_C \tag{2}$$

$$\text{subject to} \quad \theta_C x_o - X\lambda - s_x = 0 \tag{3}$$

$$Y\lambda - s_y = y_o \tag{4}$$

$$\lambda \geq 0,\ s_x \geq 0,\ s_y \geq 0. \tag{5}$$

The dual problem of (CCR_o) is described by:

$$(DCCR_o) \qquad \max\ u y_o \tag{6}$$

$$\text{subject to} \quad v x_o = 1 \tag{7}$$

$$(8) \qquad -vX + uY \leq 0$$

$$(9) \qquad v \geq 0, \;\; u \geq 0,$$

where $v \in R^m$ and $u \in R^s$ are row vectors and represent dual variables corresponding to (3) and (4), respectively.

In every optimal solution for (CCR_o) and $(DCCR_o)$, the pairs (s_x, v) and (s_y, u) are complementary each other, *i.e.*, it holds

$$(10) \qquad vs_x = 0 \;\text{ and }\; us_y = 0.$$

We use the following two-phase process with the purpose of solving (CCR_o) and determining the input surplus s_x and output shortage s_y. In Phase I, we solve (CCR_o), and in Phase II we maximize $es_x + es_y$ (the sum of input surplus and output shortage) under the added condition $\theta_C = \theta_C^*$ (the Phase I optimal objective value), where e is a row vector in which all elements are equal to 1.

Let an optimal solution in Phase II be $(\theta_C^*, \lambda^*, s_x^*, s_y^*)$, based on which we define CCR-efficiency as follows:

Definition 1 (CCR-Efficiency)
A DMU_o is called CCR-efficient *if it has $\theta_C^* = 1$, $s_x^* = 0$ and $s_y^* = 0$. Otherwise, it is called* CCR-inefficient.

If a DMU_o is CCR-efficient, it holds that $s_x^* = 0$ and $s_y^* = 0$ for every optimal solution for (CCR_o). Thus, the strong theorem of complementary slackness ensures the existence of a positive optimal solution (v^*, u^*) for the dual problem $(DCCR_o)$.

Lemma 1 *If a DMU_o is CCR-efficient, then $(DCCR_o)$ has an optimal solution (v^*, u^*) such that*

$$(11) \qquad v^* > 0 \quad \textit{and} \quad u^* > 0.$$

Suppose that $DMU_{j_1}, \ldots, DMU_{j_k}$ are CCR-efficient. Let a semipositive activity (x, y) be a nonnegative combination of them:

$$x = \sum_{l=1}^{k} \lambda_l x_{j_l} \quad \text{and} \quad y = \sum_{l=1}^{k} \lambda_l y_{j_l}. \tag{12}$$

Then, we have the following well established theorem, which will be utilized later in demonstrating our main theorems.

Theorem 1

If $DMU_{j_1}, \ldots, DMU_{j_k}$ are CCR-efficient, then the activity expressed by (12) *is also CCR-efficient.*

If a DMU_o is CCR-inefficient, a *reference set* to the DMU_o is defined by

$$E_o^C = \{j | \lambda_j^* > 0, \quad j = 1, \ldots, n\}. \tag{13}$$

The optimal solution satisfies the following relations:

$$\theta_C^* x_o = \sum_{j \in E_o^C} \lambda_j^* x_j + s_x^* \tag{14}$$

$$y_o = \sum_{j \in E_o^C} \lambda_j^* y_j - s_y^*. \tag{15}$$

The *CCR-projection* based on the optimal solution is defined by:

$$x_e^C = \theta_C^* x_o - s_x^* = \sum_{j \in E_o^C} \lambda_j^* x_j \tag{16}$$

$$y_e^C = y_o + s_y^* = \sum_{j \in E_o^C} \lambda_j^* y_j. \tag{17}$$

Lemma 2 *Every DMU in the reference set E_o^C is CCR-efficient.*

Lemma 3 *The CCR-projected activity (x_e^C, y_e^C) is CCR-efficient.*

2.2 BCC Model

The production possibility set of the BCC (Banker, Charnes and Cooper, 1984) model is described as:

$$(18) \qquad P_B = \{(x, y) \mid x \geq X\lambda,\ y \leq Y\lambda,\ e\lambda = 1,\ \lambda \geq 0\}.$$

The BCC model evaluates the efficiency of each DMU_o (x_o, y_o) $(o = 1, \ldots, n)$ by solving the following linear program:

$$
\begin{array}{lrrcl}
(19) & (BCC_o) & \min & \theta_B & \\
(20) & \text{subject to} & \theta_B x_o - X\lambda - s_x & = & 0 \\
(21) & & Y\lambda - s_y & = & y_o \\
(22) & & e\lambda & = & 1 \\
(23) & & \lambda \geq 0, \; s_x & \geq & 0, \; s_y \geq 0.
\end{array}
$$

We express the dual program of (BCC_o) as:

$$
\begin{array}{lrrcl}
(24) & (DBCC_o) \quad \max & z & = & u y_o - u_0 \\
(25) & \text{subject to} & v x_o & = & 1 \\
(26) & & -vX + uY - u_0 e & \leq & 0 \\
(27) & & v \geq 0, \; u & \geq & 0,
\end{array}
$$

where u_0 is free in sign.

As with the CCR case, we employ the two phase process for solving (BCC_o). Let an optimal solution in Phase II be $(\theta_B^*, \lambda^*, s_x^*, s_y^*)$, based on which we define BCC-efficiency as follows:

Definition 2 (BCC-Efficiency)

A DMU_o is called BCC-efficient, *if it has $\theta_B^* = 1$, $s_x^* = 0$ and $s_y^* = 0$. Otherwise, it is called* BCC-inefficient.

If a DMU_o is BCC-efficient, then there exists an optimal solution (v^*, u^*, u_0^*) with $v^* > 0$ and $u^* > 0$.

If a DMU_o is BCC-inefficient, the *reference set* E_o^B and the *BCC-projection* (x_e^B, y_e^B) based on the reference set, are defined as:

$$E_o^B = \{j | \lambda_j^* > 0, \ \ j = 1, \ldots, n\} \tag{28}$$

$$x_e^B = \theta_B^* x_o - s_x^* = \sum_{j \in E_o^B} \lambda_j^* x_j \tag{29}$$

$$y_e^B = y_o + s_y^* = \sum_{j \in E_o^B} \lambda_j^* y_j. \tag{30}$$

Corresponding to Lemmas 2 and 3, we have:

Lemma 4 *Every DMU in the reference set E_o^B is BCC-efficient.*

Lemma 5 *The BCC projected activity (x_e^B, y_e^B) is BCC-efficient.*

3 Returns to Scale of BCC-Efficient DMUs

Banker and Thrall (1992) demonstrated the following theorem on returns-to-scale of BCC efficient DMUs.

Theorem 2 (Returns-to-Scale)
Suppose DMU_o is BCC-efficient and let the sup *and* inf *of u_0 in the optimal solution for $(DBCC_o)$ be $\bar{u}_0$ and $\underline{u}_0$, respectively. Then, we have:*

1. *If $0 > \bar{u}_0$, then increasing returns-to-scale prevail in the DMU_o.*

2. *If $\bar{u}_0 \geq 0 \geq \underline{u}_0$, then constant returns-to-scale prevail in the DMU_o.*

3. *If $\underline{u}_0 > 0$, then decreasing returns-to-scale prevail in the DMU_o.*

We will denote *increasing, constant* and *decreasing* returns-to-scale by IRS, CRS and DRS, respectively.

Corollary 1 *DMU_o is CCR-efficient if and only if it is BCC-efficient and displays CRS.*

Usually, we solve (BCC_o) by the simplex method of linear programming and obtain an optimal dual solution (v^*, u^*, u_0^*) as the simplex multiplier of the Phase I optimal tableau. Thus, if $u_0^* > 0$, then we need to solve the lower bound $\underline{u}_0$, and if $u_0^* < 0$, then we need to solve the upper bound $\bar{u}_0$ for deciding the returns-to-scale characteristics of the DMU. If $u_0^* = 0$, the DMU shows CRS. The computation of $\underline{u}_0$ (or $\bar{u}_0$) is carried out in the primal side of LP.

4 Characterization of Returns to Scale

Theorem 3
If a DMU (x_o, y_o) is BCC-inefficient, the reference set E_o^B to (x_o, y_o) defined by (28), *does not include both IRS and DRS DMUs.*

Proof. Let an optimal solution for $(DBCC_o)$ to the BCC-projected activity (x_e^B, y_e^B) be (v_e, u_e, u_{0e}), with $v_e > 0$ and $u_e > 0$. Since (x_e^B, y_e^B) is BCC-efficient, we obtain the following relations:

$$u_e y_e^B - u_{0e} = 1 \tag{31}$$

$$v_e x_e^B = 1 \tag{32}$$

$$-v_e X + u_e Y - u_{0e} e \leq 0 \tag{33}$$

$$v_e \geq 0, \quad u_e \geq 0. \tag{34}$$

Hence, for $j \in E_o^B$, we have:

$$(35) \qquad -v_e x_j + u_e y_j - u_{0e} \leq 0.$$

From (31) and (32), it holds:

$$(36) \qquad -v_e x_e^B + u_e y_e^B - u_{0e} = 0.$$

By substituting the righthand side of equations (29) and (30) for (36), we obtain:

$$(37) \qquad -v_e(\sum_{j \in E_o^B} \lambda_j^* x_j) + u_e(\sum_{j \in E_o^B} \lambda_j^* y_j) - u_{0e} = 0.$$

This equation can be transformed, using $\sum_{j \in E_o^B} \lambda_j^* = 1$, into:

$$(38) \qquad \sum_{j \in E_o^B} \lambda_j^*(-v_e x_j + u_e y_j - u_{0e}) = 0.$$

From (35), (38) and $\lambda_j^* > 0$ ($j \in E_o^B$), we have the equation:

$$(39) \qquad -v_e x_j + u_e y_j - u_{0e} = 0. \quad (\forall j \in E_o^B)$$

Thus, (v_e, u_e, u_{0e}) is a coefficient of a supporting hyperplane at (x_j, y_j) for every $j \in E_o^B$. Let $t_j = 1/v_e x_j$ (> 0), then $(t_j v_e, t_j u_e, t_j u_{0e})$ is an optimal solution for $(DBCC_j)$ for $j \in E_o^B$.

Suppose that E_o^B contains an IRS DMU α and a DRS DMU β. Then, u_{0e} must be negative, since DMU α shows IRS. At the same time, u_{0e} must be positive, since DMU β has DRS. This leads to a contradiction. □

Corollary 2 *Let a reference set to a BCC-inefficient DMU (x_o, y_o) be E_o^B. Then, E_o^B consists of one of the following combinations of BCC-efficient DMUs.*

- (i) *All DMUs have IRS.*

- (*ii*) *Mixture of DMUs with IRS and CRS.*
- (*iii*) *All DMUs have CRS.*
- (*iv*) *Mixture of DMUs with CRS and DRS.*
- (*v*) *All DMUs show DRS.*

Theorem 4 (Characterization of Return-to-Scale)
Let the BCC-projected activity of a BCC-inefficient DMU (x_o, y_o) *be* (x_e^B, y_e^B) *and the reference set to* (x_o, y_o) *be* E_o^B*. Then,* (x_e^B, y_e^B) *belongs to*

1. *IRS, if* E_o^B *consists of DMUs in categories* (*i*) *or* (*ii*) *of Corollary 2,*
2. *CRS, if* E_o^B *consists of DMUs in category* (*iii*)*, and*
3. *DRS, if* E_o^B *consists of DMUs in categories* (*iv*) *or* (*v*).

Proof. In the case of (*i*) or (*ii*), E_o^B contains at least one DMU with IRS and any supporting hyperplane at (x_e^B, y_e^B) is also a supporting hyperplane at the IRS DMU, as shown in the proof of Theorem 3. Thus, the upper bound of u_{0e} must be negative. By the same reasoning, in the case of (*iv*) or (*v*), the projected activity is DRS. In the case of (*iii*), every DMU j ($j \in E_o^B$) is CCR-efficient by Corollary 1. Since (x_e^B, y_e^B) is a convex combination of CCR-efficient DMUs, it is CCR-efficient, too. Thus, it shows CRS by Corollary 1. □

5 A Numerical Example

Table 1 exhibits the data of 14 general hospitals, each with 2 inputs and 2 outputs. Our method is as follows. First, we solve the BCC model; the results are shown in the columns 'BCC (RTS)' and 'Reference Set'. In the 'BCC (RTS)' column, the returns-to-scale of BCC-efficient DMUs ($\theta_B^* = 1$) are evaluated by BT (1992) and denoted by (I), (C) and (D), which mean IRS, CRS and DRS, respectively. The returns-to-scale characteristics of BCC-inefficient DMUs are decided by those in the reference set, using Theorem 4. For example, H4 has the reference set consisting of H1, which has IRS. Hence, the BCC-projected activity of H4 shows IRS. The reference set of H7 consists of H1(I), H3(C) and H8(C) and we can conclude that H7 has IRS. Since H13 has H10(C), H12(D) and H14(D) as reference, H13 has DRS.

Table 1: Data of General Hospital and Returns-to-Scale

	Input		Output		BCC	Reference	RTS
Hospital	Doctor	Nurse	Outpatient	Inpatient	(RTS)	Set	
H1	3008	20980	97775	101225	1 (I)		
H2	3985	25643	135871	130580	1 (C)		
H3	4324	26978	133655	168473	1 (C)		
H4	3534	25361	46243	100407	0.851	H1	I
H5	8836	40796	176661	215616	0.845	H2, H3, H10	C
H6	5376	37562	182576	217615	1 (C)		
H7	4982	33088	98880	167278	0.862	H1, H3, H8	I
H8	4775	39122	136701	193393	1 (C)		
H9	8046	42958	225138	256575	0.996	H2, H10	C
H10	8554	48955	257370	312877	1 (C)		
H11	6147	45514	165274	227099	0.919	H6, H12	D
H12	8366	55140	203989	321623	1 (D)		
H13	13479	68037	174270	341743	0.794	H10, H12, H14	D
H14	21808	78302	322990	487539	1 (D)		

6 The Output Oriented Case

The output oriented BCC model can be dealt with similarly. This model is described as:

$$
\begin{aligned}
&(40) \qquad (BCCO_o) \quad \max \quad \tau_B \\
&(41) \qquad \text{subject to} \quad X\lambda + s_x = x_o \\
&(42) \qquad \tau_B y_o - Y\lambda + s_y = 0 \\
&(43) \qquad e\lambda = 1 \\
&(44) \qquad \lambda \geq 0, \; s_x \geq 0, \; s_y \geq 0.
\end{aligned}
$$

Let an optimal solution for $(BCCO_o)$ be $(\tau_B^*, \lambda^*, s_x^*, s_y^*)$. A DMU (x_o, y_o) is defined to be *output oriented BCC-efficient* (BCCO-efficient) if and only if it holds that $\tau_B^* = 1, s_x^* = 0$ and $s_y^* = 0$, and hence if and only if it is BCC-efficient.

The returns-to-scale characteristics of a BCCO-efficient DMU are the same as those in the BCC case. For a BCCO-inefficient DMU, the BCCO-projected activity (x_e, y_e) is defined by

$$
\begin{aligned}
&(45) \qquad x_e = x_o - s_x^* \\
&(46) \qquad y_e = \tau_B^* y_o + s_y^*.
\end{aligned}
$$

The returns-to-scale of such activity are generally different from those of the BCC-projected case. However, we can decide the characteristics in a similar way as in the BCC case, using the reference set.

7 Comparisons with Other Methods

We will briefly survey two representative methods related to returns-to-scale and compare them with our method.

7.1 Färe, Grosskopf and Lovell (1985)

FGL (1985) suggest the following two-step method to estimate returns-to-scale. In step 1, the BCC and the CCR models are solved to determine the optimal objective values θ_B^* and θ_C^* and define the scale efficiency θ_S^* as $\theta_S^* = \theta_C^*/\theta_B^*$. A value $\theta_S^* = 1$ indicates that the DMU has CRS and a value $\theta_S^* < 1$ indicates IRS or DRS. In step 2, when $\theta_S^* < 1$, the following linear program is solved to determine whether the scale inefficiency is associated with IRS or DRS.

$$
\begin{aligned}
&(47) & (LP_E) \quad \theta_E^* &= \min\ \theta_E \\
&(48) & \text{subject to} \quad \theta_E x_o - X\lambda - s_x &= 0 \\
&(49) & Y\lambda - s_y &= y_o \\
&(50) & e\lambda &\le 1 \\
&(51) & \lambda \ge 0, \quad s_x &\ge 0, \quad s_y \ge 0.
\end{aligned}
$$

FGL (1985) state that if $\theta_E^* < 1$ and $\theta_E^* = \theta_B^*$, then the DMU has IRS and if $\theta_E^* < 1$ and $\theta_E^* < \theta_B^*$, then the DMU shows DRS.

This method requires 3 LP solutions (BCC, CCR and the above (LP_E)) for IRS and DRS DMUs, and is said to be a three-pass method. We can simplify this method by applying Theorem 4 described in this paper, since solving (LP_E) is necessary only for DMUs with $\theta_S^* < 1$ and $\theta_B^* = 1$.

7.2 Banker, Chang and Cooper (1995)

BCC(1995) solve the CCR model in step 1, and if $e\lambda^* = 1$, then the DMU has CRS, and if $e\lambda^* < 1$, they proceed to step 2. In step 2, the following linear program is solved to determine whether the DMU is associated with

CRS or IRS.

$$(LP_Z) \quad z^* = \max e\lambda \tag{52}$$

$$\text{subject to} \quad X\lambda + s_x = \theta_C^* x_o \tag{53}$$

$$Y\lambda - s_y = y_o \tag{54}$$

$$e\lambda \leq 1 \tag{55}$$

$$\lambda \geq 0, \; s_x \geq 0, \; s_y \geq 0, \tag{56}$$

where θ_C^* is the optimal objective value for the CCR model. IF $z^* = 1$, then the DMU has CRS and if $z^* < 1$, it shows IRS. The case $e\lambda^* > 1$ in the CCR model can be handled in an obvious modification of the above linear program (LP_Z).

This method requires 2 LP solutions for DMUs with $e\lambda^* \neq 1$ and is concerned only with the returns-to-scale characteristics and the information obtained from the CCR model. The BCC-projection is not explicitly described in this method.

8 Conclusion

In this paper we presented a simple alternative method for deciding the returns-to-scale characteristics of BCC (BCCO)-projected activities. This method is 'one' pass in the sense that BCC software equipped with a procedure for deciding returns-to-scale of BCC-efficient DMUs (*e.g.* Banker and Thrall (1992)), is sufficient for this purpose. No other software is needed. Since the number of BCC-efficient DMUs is considerably less than that of BCC-inefficient ones, this method will contribute to save computation time. Also, Theorems 3, 4 and Corollary 2 contribute to the development of theory and algorithms in DEA.

Acknowledgment

The author would like to express his sincere gratitude to Professor Cooper for supplying returns-to-scale related papers and for his continual guidance in DEA research.

References

[1] Banker, R.D., 1980, "Studies in Cost Allocation and Efficiency Evaluation," DBA Thesis, Harvard University Graduate School of Business.

[2] Banker, R.D., 1984, "Estimating Most Productive Scale Size Using Data Envelopment Analysis," *European Journal of Operations Research,* 17, 35-44.

[3] Banker, R.D., I. Bardhan and W.W. Cooper, 1995, "A Note on Returns of Scale in DEA," *European Journal of Operations Research,* (to appear)

[4] Banker, R.D., H. Chang and W.W. Cooper, 1995, "Equivalence and Implementation of Alternative Methods for Determining Returns to Scale in Data Envelopment Analysis," *European Journal of Operations Research,* (to appear)

[5] Banker, R.D., A. Charnes and W.W. Cooper, 1984, "Models for the Estimation of Technical and Scale Inefficiencies in Data Envelopment Analysis," *Management Science,* 30, 1078-1092.

[6] Banker, R.D. and R.M. Thrall, 1992, "Estimation of Returns to Scale using Data Envelopment Analysis," *European Journal of Operations Research,* 62, 74-84.

[7] Charnes, A., W.W. Cooper and E. Rhodes, 1978, "Measuring the Efficiency of Decision Making Units," *European Journal of Operations Research,* 2, 429-444.

[8] Färe, R.,S. Grosskopf and C.A.K. Lovell, 1985, *The Measurement of Efficiency of Production,* Boston, Kluwer Nijhoff.