경영학연구논문집 제7권 1호 pp. 89~102, 울산대학교
Journal of Managment Vol. 7, No. 1, pp. 89~102, 2000
University of Ulsan.

Inefficiency Identification in IDEA (Imprecise Data Envelopment Analysis) via Additive Model

Kyung Sam Park
Assistant Professor of Management Information Systems
College of Business Administration

<Abstract>

Data Envelopment Analysis (DEA) is a mathematical programming approach to evaluating the relative efficiency of Decision Making Units (DMUs) that use multiple inputs to produce multiple outputs. While assuming exact data in ordinary DEA, development of Imprecise Data Envelopment Analysis (IDEA) makes possible to deal with imprecise data in DEA. However, IDEA only provides an aggregated measure of inefficiency for each DMU. It is thus needed to develop methods from which we can obtain specific inefficiencies such as slacks, as well as peer groups and scale sizes, as have been done in ordinary DEA evaluations. The purpose of this paper is hence on the identification of specific inefficiencies in IDEA. This is done via employing an additive model which we refer to as additive IDEA model. A point to be noted is that the original formulation becomes a nonlinear programming problem. We thus transform it into a linear programming equivalent and then present a two-stage method to identify specific inefficiencies. In the first stage, we obtain an aggregated measure of inefficiency from solving the linear version of additive IDEA model. We then retrieve exact data based upon the optimal solutions obtained in the first stage. These exact data retrieved are used in the next stage which implies that an ordinary additive DEA model is constructed. We can thus obtain the specific inefficiencies in terms of slacks as well as peer groups and scale sizes for each DMU to be considered in IDEA problem.

가산모델을 이용한 IDEA에서의 비효율성 결정

박경삼
경영학부 경영정보전공

<요 약>

DEA(Data Envelopment Analysis)는 복수의 투입-산출요소들을 갖는 DMU(Decision Making Unit)들의 효율성을 분석하기 위한 수리계획 접근법이다. 기존 DEA에서는 투입-산출요소들에 관한 정확한 데이터가 주어진다고 가정하고 있는 반면, IDEA (Imprecise DEA)는 불완전 데이터를 DEA에서 취급하기 위하여 개발되었다. 그러나, IDEA는 각각의 DMU에 대한 종합적 비효율성 만을 측정/제공한다. 따라서, 기존의 DEA 평가에서 이루어졌던 것처럼, DMU들에 관한 세부적인 비효율성(예를 들어 투입-잔여분, 산출-부족 등)을 측정할 수 있는 방법론 개발이 필요하다. 이러한 세부적인 (또는 구체적인) 비효율성을 밝혀내는 것이 본 논문의 목적이다. 한가지 중요한 점은, 세부적인 비효율성을 결정하기 위한 수리계획 모형이 비선형 문제가 된다는 것이다. 따라서 본 논문에서는 비선형 문제를 동일한 해를 갖는 선형문제로 전환한 후, 세부적인 비효율성을 결정하기 위한 two-stage 방법을 개발한다. 첫번째 단계에서는 전환된 선형문제로부터 종합적 비효율성을 먼저 얻는다. 이때 선형문제에 관한 변수들의 최적해를 도출한다. 도출된 최적해는 정확한 데이터를 포함하며, 이 정확한 데이터는 두번째 단계에서 이용되어 기존의 DEA 모델을 형성하게 된다. 따라서 본 논문의 목적인 세부적인 비효율성을 결정할 수 있다.

1. Introduction

Cooper et al. (1999) developed an IDEA (Imprecise DEA) model for dealing with imprecise data in DEA and methods for transforming the nonlinear version of IDEA model into a linear programming equivalent. Examples of imprecise data are such that input-output data are known only to lie within the upper and lower bounds and/or to obey ordinal relations. It was also shown how conditions on the (multiplier) variables as well as the data could be treated in this same manner. This included Assurance Region (AR) conditions on the variables, as in Thompson et al. (1990, 1995) and the combined variable-data transformations employed in the cone-ratio envelopment of Charnes et al. (1990) and also Brockett et al. (1997)[1]. Thus the resulting approach shows how all the above approaches can be combined into one unified approach which

1) For more detailed descriptions on AR (or weight restriction), see Allen et al. (1997) who review the developments and suggest future directions for research on the use of various AR bounds in DEA.

is referred to as AR-IDEA.

There have been applications of IDEA and AR-IDEA to the efficiency evaluation of telephone offices (Kim et al., 1999) and to a mobile telecommunication company (Cooper et al., 2000a). In particular, Cooper et al. (2000a) showed that the original AR conditions in AR-IDEA model should be adjusted in accordance with rescaling the input-output data, for use them correctly in the linear programming form. Cooper et al. (2000a) also provided important points to help the potential uses of IDEA, including how to treat strict as well as weakly ordered data in order to effect proper efficiency discriminations.

Still further extensions have been made. Cooper et al. (2000b) extended the transformation developed in Cooper et al. (1999) to a more general situation of imprecise data via introducing dummy variables. This hence removed a limitation underlying the transformation in Cooper et al. (1999) and also formalized the adjustment of original AR conditions as shown in Cooper et al. (2000a). More recently, Zhu (2000a) and Park (2000) independently showed that the transformations could be done in the simpler manner via only employing variable alterations without rescaling and introducing dummy variables as was done in Cooper et al. (2000b).

Note that the papers mentioned above deal with CCR (Charnes, Cooper and Rhodes, 1978) model in which some (or all) input-output data are known imprecisely in arbitrary linear forms. We know, however, that the transformation developed can also apply to other (multiplier) DEA models, such as BCC (Banker, Charnes and Cooper, 1984) model and additive model (Charnes et al., 1985), involving imprecise data. We thus show an application of the transformation to an additive model based IDEA problem viz., transforming a nonlinear additive model, due to data imprecision, into a linear programming equivalent[2].

Nevertheless, we have questions upon how we can obtain the information necessary for analyzing (technical) efficiency in the linear IDEA model transformed. In other word, we need to obtain the amount of inefficiencies in terms of slacks as well as the radial efficiency in terms of proportional reduction for all inputs to be considered in IDEA, if the IDEA is based on CCR or BCC model. It is not easy to do this, as far as we know, because the IDEA based on CCR or BCC model involves a non-Archimedean element which restricts the (multiplier) variable values to be positive.

Choosing an additive model[3] for IDEA, so we can thus avoid employing non-Archimedean in the model. This implies that the variable values can be restricted to be positive without using non-Archimedean element. It is essential for us to achieve the goal of this paper identification of inefficiencies in IDEA problems, from which we can also obtain peer groups and scale sizes as have been done in ordinary DEA evaluations. In the present paper, this is done in a two-stage manner which is

2) Of course, this is not the main purpose of this paper.

3) See Cooper et al. (1999) for the additive models in detail, where various points are revealed for the use of additive models including relations to other models and measures in DEA.

different from two-phase procedures as used in ordinary CCR and BCC models.

The plane of development is as follows. First, an additive model with imprecise data is shown, referred to as additive model based IDEA or briefly additive IDEA model. We then show how the nonlinear version of additive IDEA model can be transformed into an ordinary linear programming problem. This is followed by the identification of inefficiencies in terms of slacks, including specifications of peer groups and scale sizes in additive IDEA model transformed into linear programming equivalent. Numerical illustrations are then provided. Finally a summary and sketch of further research opportunities conclude this paper.

2. Additive IDEA model

We consider an additive model in which the input-output data are known imprecisely in arbitrary linear forms. The following model will help to make this more precise:

$$\left.\begin{aligned} \max\ z_0 &= \sum_{r=1}^{s}\mu_r y_{r0} - \sum_{i=1}^{m}\omega_i x_{i0} \\ \text{s.t.}\quad & \sum_{r=1}^{s}\mu_r y_{rj} - \sum_{i=1}^{m}\omega_i x_{ij} \le 0, \quad j=1,\dots,n \end{aligned}\right\} \quad (1.1)$$

$$\left.\begin{aligned} \mathbf{y}_r &= (y_{rj}) \in D_r^+, \quad r=1,\dots,s \\ \mathbf{x}_i &= (x_{ij}) \in D_i^-, \quad i=1,\dots,m \end{aligned}\right\} \quad (1.2)$$

$$\mu = (\mu_r) \ge 1; \quad \omega = (\omega_i) \ge 1. \quad (1.3)$$

Here, y_{rj}, x_{ij} respectively represent the observed or recorded amounts of the rth output (r = 1,...,s) and the ith input (i = 1,...,m) for each Decision Making Unit, DMUj (j = 1,...,n). The y_{r0}, x_{i0} data represent the outputs and inputs for DMU_0, the DMU_j to be evaluated The variables , are multipliers associated with outputs and inputs and we restrict these variable values to be greater than or equal to unity as in an ordinary additive DEA model[4].

The sets D_r^+, D_i^- in (1.2) represent imprecise data for the vector of output variables $\boldsymbol{yr}$ = $(y_{r1},\cdots,y_{rn})$ and input variables x_i = $(x_{i1},\cdots,x_{in})$. Examples of imprecise data for outputs are

4) Note that we can also use AR bounds in place of (1.3) so that model (1) becomes an AR-IDEA model as shown in Cooper et al. (1999). However, we do not deal with AR bounds in the present paper because our focus is mainly on analyzing technical efficiency in DEA with imprecise data.

Fixed bounds :	$y_{rj}^{-} \le y_{rj} \le y_{rj}^{+}$	(2.1)
Strict orders :	$y_{rj} - y_{r,j+1} \le -\alpha_{rj}$	(2.2)
Ratio bounds :	$y_{rj}^{-} \le y_{rj} / y_{r1} \le y_{rj}^{+}$	(2.3)
Weak orders :	$y_{rj} - y_{r,j+1} \le 0$	(2.4)
Multiplied orders :	$\gamma_{rj} y_{rj} \le y_{r,j+1}$	(2.5)
Difference ranks :	$y_{rj} - y_{r,j+1} \le y_{r,j+1} - y_{r,j+2}$,	(2.6)

where y_{rj}^{-}, y_{rj}^{+} and α_{rj}, γ_{rj} are positive constants to be given in advance. The set D_r^{+} can include each of the above constraints, mixtures of them, or other available forms if any. The set D_i^{-} follows similarly for inputs.

Without loss of generality, we thus represent D_r^{+}, D_i^{-} to include arbitrary linear forms as follows:

$$D_r^{+} = \{\mathbf{y}_r \in \Re^n : \mathbf{H}_r^{+} \mathbf{y}_r^{\mathrm{T}} \le \mathbf{h}_r^{+}\}, \quad r = 1, \dots, s \tag{3.1}$$

$$D_i^{-} = \{\mathbf{x}_i \in \Re^n : \mathbf{H}_i^{-} \mathbf{x}_i^{\mathrm{T}} \le \mathbf{h}_i^{-}\}, \quad i = 1, \dots, m. \tag{3.2}$$

The $\mathbf{H}_r^{+}, \mathbf{H}_i^{-}$ $k_r^{+} \times n, k_i^{-} \times n$ k_r^{+}, k_i^{-} $\mathbf{h}_r^{+}, \mathbf{h}_i^{-}$ $\mathbf{h}_r^{+} \in \Re^{k_r^{+}}, \mathbf{h}_i^{-} \in \Re^{k_i^{-}}$. D_r^{+}, D_i^{-} represent the permissible values of output and input data variables satisfying the systems of linear constraints in (3.1) and (3.2), respectively. We assume, throughout this paper, that these constraints on data are consistent and closed so that the optimal objective value z0* is attained in model (1).

3. Transformation to linear programming equivalents

As shown in Zhu (2000a) and Park (2000), we introduce new variables Y_{rj}, X_{ij} to reduce additive IDEA model (1) to an ordinary linear programming problem. We then define

$$\begin{aligned} Y_{rj} &= y_{rj}\mu_r, \quad r = 1, \dots, s; \; j = 1, \dots, n \\ X_{ij} &= x_{ij}\omega_i, \quad i = 1, \dots, m; \; j = 1, \dots, n. \end{aligned} \tag{4}$$

Since all multipliers are to be positive, these equations can be changed to

$$\begin{aligned} y_{rj} &= Y_{rj} / \mu_r, \quad r = 1, \dots, s; \; j = 1, \dots, n \\ x_{ij} &= X_{ij} / \omega_i, \quad i = 1, \dots, m; \; j = 1, \dots, n. \end{aligned} \tag{5}$$

This implies that the nonlinear version of additive IDEA model as in (1) is converted into a linear programming equivalent. To complete this, we employ the following theorem:

Theorem 1. (i) Let $\mathbf{Y}_r = (Y_{r1},\ldots,Y_{rn})$ and $\mathbf{X}_i = (X_{i1},\ldots,X_{in})$. The constraints on the data as in (3) can then be converted intor

$$B_r^+ = \{(\mathbf{Y}_r\mu_r) \in \Re^{n+1} : \mathbf{H}_r^+\mathbf{Y}_r^{\mathrm{T}} \le \mu_r\mathbf{h}_r^+\},\quad r=1,\ldots,s$$
$$B_i^- = \{(\mathbf{X}_i,\omega_i) \in \Re^{n+1} : \mathbf{H}_i^-\mathbf{X}_i^{\mathrm{T}} \le \omega_i\mathbf{h}_i^-\},\quad i=1,\ldots,m.$$

(ii) Thus, model (1) can be transformed into the following LP problem:

$$\left.\begin{aligned} \max\ z_0 &= \sum_{r=1}^{s} Y_{r0} - \sum_{i=1}^{m} X_{i0} \\ \text{s.t.}\quad & \sum_{r=1}^{s} Y_{rj} - \sum_{i=1}^{m} X_{ij} \le 0,\quad j=1,\ldots,n \end{aligned}\right\} \tag{6.1}$$

$$\left.\begin{aligned} &(\mathbf{Y}_r,\mu_r) \in B_r^+,\quad r=1,\ldots,s \\ &(\mathbf{X}_i,\omega_i) \in B_i^-,\quad i=1,\ldots,m \end{aligned}\right\} \tag{6.2}$$

$$\mu = (\mu_r) \ge 1;\quad \omega = (\omega_i) \ge 1. \tag{6.3}$$

We omit the proof of Theorem 1 because it can be proven simply using the equations in (4) and (5). Instead, we give concrete examples of the new data constraints transformed for the original constraints given in (2) as follows:

Fixed bounds :	$y_{rj}^-\mu_r \le Y_{rj} \le y_{rj}^+\mu_r$	(7.1)
Strict orders :	$Y_{rj} - Y_{r,j+1} \le -\alpha_{rj}\mu_r$	(7.2)
Ratio bounds :	$y_{rj}^- Y_{r1} \le Y_{rj} \le y_{rj}^+ Y_{r1}$	(7.3)
Weak orders :	$Y_{rj} - Y_{r,j+1} \le 0$	(7.4)
Multiplied orders :	$\gamma_{rj} Y_{rj} \le Y_{r,j+1}$	(7.5)
Difference ranks :	$Y_{rj} - Y_{r,j+1} \le Y_{r,j+1} - Y_{r,j+2}$.	(7.6)

Therefore, we achieve a linear programming equivalent to additive IDEA model (1) that is including arbitrary linear imprecise data. Moreover, no change is made for the original multiplier variables μ, ω and their conditions as in (1.3).

4. Technical efficiency in additive IDEA model

In fact, we can have measures of technical efficiency for every DMU_0 from solving the linear version of additive IDEA model as in (6) and then obtaining the optimal objective value z_0^*. We can then classify DMUs into two groups: technically efficient when $z_0^* = 0$ and technically inefficient when $z_0^* < 0$.

However, the obtained z_0^* represents total sum of inefficiencies (or total sum of slacks) that DMU_0 has under model (6). We thus need to separate this total sum into individual slacks for each of inputs and outputs under consideration in evaluating every DMU_0. We also need to see the referent DMUs and returns to scale of DMU_0 as have been done in ordinary DEA evaluations. This is done via two-stage ways like those we develop below.

First, we obtain the optimal objective value z_0^* from solving model (6). At the same time, we also have the optimal solutions for the variables Y_{rj}, X_{ij} (including Y_{r0}, X_{i0}) and , used to achieve the z_0^*, which we denote as Y_{rj}^*, X_{ij}^* and μ^*, ω^*. From these optimal solutions, we further retrieve the data y_{rj}^*, x_{ij}^* via using the equations in (5). Note that all the multiplier values μ^*, ω^* are positive in (6) so that there is no trouble in the data retrieval via (5).

We now use the retrieved data y_{rj}^*, x_{ij}^* $\forall r,i,j$ in the second stage to achieve the individual slacks as well as the referent DMUs and return to scale of DMU_0. To make this more concrete, we write

$$
\begin{aligned}
\max\ & \sum_{r=1}^{s} s_r^+ + \sum_{i=1}^{m} s_i^- \\
\text{s.t.}\ & \sum_{j=1}^{n} x_{ij}^* \lambda_j + s_i^- = x_{i0}^*, \quad i = 1,\dots,m \\
& \sum_{j=1}^{n} y_{rj}^* \lambda_j - s_r^+ = y_{r0}^*, \quad r = 1,\dots,s \\
& s_r^+, s_i^-, \lambda_j \geq 0, \quad \forall r,i,j.
\end{aligned} \tag{8}
$$

Define $z_0'^* = \max(\sum_{r=1}^{s} s_r^+ + \sum_{i=1}^{m} s_i^-)$. It is then clear that $z_0^* = -z_0'^*$ in models (6) and (8) by dual theory of linear programming. Note that the values of data variables in (6) are already fixed as yrj*, xij* $\forall$r,i,j to obtain the optimal objective value z_0* and then the same values fixed are used in (8) to achieve It is then clear that z0*.

Therefore, the total sum of inefficencies z_0^* obtained from model (6) is separated into individual slacks s_r^{+*}, s_i^{-*} $\forall r,i$ which can be obtained from solving model (8). We may then utilize a projection formula as in ordinary additive models *-viz.*, $\hat{y}_{rj}^* = y_{rj}^* + s_r^{+*}$; $\hat{x}_{ij}^* = x_{ij}^* - s_i^{-*}$ - which render DMUs efficient. We can also have peer groups and scale size from obtaining λ_j^* in model (8).

It should be noted that we have achieved only a matrix of exact data$[(y_{rj}^*)^T\ \forall r$, $(x_{ij}^*)^T\ \forall i]$, for each of DMU_0, among various possible matrices of exact data $[(y_{rj})^T\ \forall r$, $(x_{ij})^T\ \forall i]$ satisfying the data constraints as in (3). This matrix of exact data achieved is then used in (8) to obtain s_r^{+*}, s_i^{-*} $\forall r,i$ and λ_j^* $\forall j$. So we should interpret the resulting individual slacks as well as peer groups and scale size for only the data fixed when the efficiency of DMU_0 is maximized in (6). In other words, it may be not true that DMU_0 has the same slacks, peer groups and scale size as obtained above for all possible data.

On the other hand, we want to note that there are other ways besides the present approach to the identification of inefficiency in IDEA problems. As noted in Zhu (2000b),[5] this is done by determining exact data $[(y_{rj}^*)^T\ \forall r,\ (x_{ij}^*)^T\ \forall i]$ *directly* from the data constraints, without utilizing instruments such as model (6) we used. The chosen exact data are then used as in (8), so the inefficiencies are identified in a simpler manner.

However, there exist limitations to potential uses of the approach developed in Zhu (2000b). As mentioned in Park (2000), this approach works only for simple forms of imprecise data that are limited to bounded data and ordinal data known in ratio (but not interval) scale.[6] In contrast, we assume, without loss of generality, that imprecise data are known in arbitrary linear forms. So we have developed and used model (6) to determine exact data from imprecisely known data as well as to measure total sum of inefficiencies.

5. Numerical examples

We here provide two examples to demonstrate how model (6) is used to determine exact data from imprecise data and then the determined data are used in model (8) to identify inefficiencies. One is an example shown in Cooper et al. (1999). The other is the modified version of the first example which includes more complicated forms of imprecise data.

Reference to Table 1, it shows an example involving five DMUs. As indicated in the column headings, the data are to be dealt with in ordinal and bounded forms as well as in the customary exact forms represented by the conditions $\mathbf{y}_r \in D_r^+,\ \mathbf{x}_i \in D_i^-$ in (3).

Table 1. An example shown in Cooper et al. (1999)

	Outputs		**Inputs**	
	Exact	**Ordinal**	**Exact**	**Bound**
DMU j	Revenue y_{1j}	Satisfaction y_{2j}^{1}	Cost x_{1j}	Judgement x_{2j}^{2}
1	2000	4	100	[0.6, 0.7]
2	1000	2	150	[0.8, 0.9]
3	1200	5	150	1
4	900	1	200	[0.7, 0.8]
5	600	3	200	1

주) 1. Weak ranking such that 5 highest rank, ..., 1 lowest rank (i.e., $y_{23} \geq y_{21} \geq \ldots \geq y_{24}$).
2. Ratio bound based on the reference DMUs 3 or 5 (e.g., $0.6 \leq x_{21} \leq 0.7$ with $x_{23} = 1$).

5) See also Zhu (2000c) and Chen et al. (2000).

6) However, it is more convenient that we use the approach of Zhu (2000b-c), if we have imprecise data that can be treated within the method developed in Zhu (2000b-c).

Now using Theorem 1 (or transformation rules as in (7)), we achieve the following linear programming to evaluate DMU$_1$:

$$
\begin{aligned}
\max\ & z_0 = Y_{11} + Y_{21} - X_{11} - X_{21} \\
\text{s.t.}\ & Y_{1j} + Y_{2j} - X_{1j} - X_{2j} \le 0, \quad j = 1,\ldots,5 \\
& B_1^+ : \{Y_{11} = 2000\mu_1;\ \ldots;\ Y_{15} = 600\mu_1\} \\
& B_2^+ : \{Y_{24} \le Y_{22} \le Y_{25} \le Y_{21} \le Y_{23}\} \\
& B_1^- : \{X_{11} = 100\omega_1;\ \ldots;\ X_{15} = 200\omega_1\} \\
& B_2^- : \{0.6\omega_2 \le X_{21} \le 0.7\omega_2;\ \ldots;\ X_{25} = \omega_2\} \qquad (9) \\
& \mu_1, \mu_2, \omega_1, \omega_2 \ge 1.
\end{aligned}
$$

We then solve problem (9) and find that DMU$_1$ is efficient because $z_0^* = 0$, as shown in the cell for row 1 column 1 of Table 2. Carrying out these same operations on the other DMU$_j$, j = 2,3,4,5 also produces the four additional evaluations exhibited in the first column of Table 2. DMU$_3$ is found to be efficient, too, but the other DMUs are not to be efficient. These results are consistent with those in Cooper et al. (1999) as shown in the last column of Table 2. It should be noted that Cooper et al. (1999) used CCR IDEA model to obtain *efficiency* whereas we are using additive IDEA model as in (9) to obtain inefficiency. So efficiency at unity is identical to inefficiency at zero.

Table 2. Evaluation results

DMUs	**z_0^* obtained from model (9)[1]**	**z_0^* in Cooper et al. (1999)[2]**
1	0	1
2	-1321.429	0.87499
3	0	1
4	-1200	0.99999
5	-2314.286	0.69999

주) 1. LINDO was used for the calculations.
2. We copied here the z_0^* values in Cooper et al. (1999), obtained from CCR IDEA model with $\varepsilon=10^{-6}$.

When z_0^* is obtained for each DMU in model (9), we also have the optimal solutions Y_{rj}^*, X_{ij}^*, μ_r^*, μ_i^* for r = 1,2; i = 1,2; j = 1,⋯,5. We then retrieve exact data y_{rj}^*, x_{ij}^* from these optimal solutions via equations in (5). As a result, Table 3 shows the retrieved data when DMU$_1$ is evaluated in model (9). Table 3 also shows that the same data are retrieved for the other four DMUs.

Table 3. The exact data retrieved from model (9) when z_0* is obtained[1]

DMUs	Outputs		Inputs	
	y_{1j}^*	y_{2j}^*	x_{1j}^*	x_{2j}^*
1	2000	0	100	0.7
2	1000	0	150	0.8
3	1200	1664.2858	150	1
4	900	0	200	0.7
5	600	0	200	1

주) 1. The same data matrix is obtained for all the five DMUs.

The exact data retrieved are now used in (8) to obtain individual slacks and lambdas. For instance, the following model is to evaluate DMU1:

$$
\begin{aligned}
z_0'^* = \max\ & s_1^+ + s_2^+ + s_1^- + s_2^- \quad \text{s.t.} \\
& 2000\lambda_1 + 1000\lambda_2 + 1200\lambda_3 + 900\lambda_4 + 600\lambda_5 - s_1^+ = 2000 \\
& 1664.2858\lambda_3 - s_2^+ = 0 \\
& 100\lambda_1 + 150\lambda_2 + 150\lambda_3 + 200\lambda_4 + 200\lambda_5 + s_1^- = 100 \\
& 0.7\lambda_1 + 0.8\lambda_2 + \lambda_3 + 0.7\lambda_4 + \lambda_5 + s_2^- = 0.7 \\
& s_1^+, s_2^+, s_1^-, s_2^-, \lambda_j \geq 0, \quad j = 1,\dots,5.
\end{aligned} \tag{10}
$$

As shown in the first row of Table 4, DMU_1 has no (positive) slack which result is natural since DMU_1 is to be efficient under model (9). Carrying out these same operations on the other four DMUs also produces the four additional evaluations exhibited in the subsequent rows of Table 4. Reference to the second row, it shows that the total sum of inefficiencies for DMU_2, $z_0'^* = 1321.429$ which is to be determined in advance as shown in the cell for row 2 column 1 of Table 2, is now characterized as the two positive slacks (i.e., $s_1^{+*} = 1285.714$, $s_1^{-*} = 35.714$). We also know that, from $\lambda_1^* = 1.143$, these inefficiencies are to be identified by the referent DMU_1. A similar interpretation can apply to other DMUs.

Table 4. Optimal solutions to model (10)

DMUs	z_0^*	Slacks (>0)	Lambdas (>0)
1	0	none	$\lambda_1^* = 1$
2	1321.429	$s_1^{+*} = 1285.714$ $s_1^{-*} = 35.714$	$\lambda_1^* = 1.143$
3	0	none	$\lambda_3^* = 1$
4	1200	$s_1^{+*} = 1100$ $s_1^{-*} = 100$	$\lambda_1^* = 1$
5	2314.286	$s_1^{+*} = 2257.143$ $s_1^{-*} = 57.143$	$\lambda_1^* = 1.429$

In the above example, the same input-output data are determined in evaluating all five DMUs as shown in Table 3. However, it is not necessary for some other IDEA problems. To evidence, let us replace the ordinal data for output 2 in the second column of Table 1 with the new set of constraints,

$$D_2'^{+} = \{(y_{2j}) : 10 \le y_{24};\ 2y_{24} \le y_{22};\ y_{22} \le y_{25};\ y_{24} + y_{22} + y_{25} \le y_{21};\ 50 \le y_{23} - y_{21};\ y_{23} \le 100\} \quad (11)$$

but remain the data for the other output and inputs without change, in order to use them as a new example. Note that the new set in (11) is constructed in a way that the data both in the second column of Table 1 and (11) are consistent in terms of ordinal relations but the data in (11) are more complicated.

In this new example, applying Theorem 1 we then obtain the following linear programming problem to evaluate DMU_1:

$$\begin{aligned} \max\ & z_0 = Y_{11} + Y_{21} - X_{11} - X_{21} \\ \text{s.t.}\ & Y_{1j} + Y_{2j} - X_{1j} - X_{2j} \le 0,\quad j = 1,\dots,5 \\ & B_1^+ : \{Y_{11} = 2000\mu_1;\ \dots;\ Y_{15} = 600\mu_1\} \\ & B_2'^{+} : \{10\mu_2 \le Y_{24};\ 2Y_{24} \le Y_{22};\ \dots;\ Y_{23} \le 100\mu_2\} \\ & B_1^- : \{X_{11} = 100\omega_1;\ \dots;\ X_{15} = 200\omega_1\} \\ & B_2^- : \{0.6\omega_2 \le X_{21} \le 0.7\omega_2;\ \dots;\ X_{25} = \omega_2\} \\ & \mu_1, \mu_2, \omega_1, \omega_2 \ge 1. \end{aligned} \quad (12)$$

Note that only difference between models (9) and (12) is replacing B_2^+ by $B_2'^{+}$.

We then solve problem (12) for each of the five DMUs in ways to change the objective function in accordance with DMU_j, j = 1,·,5 to be evaluated but not to change the constraints. The resulting inefficiencies are listed in the first column of Table 5. In comparing with the first column of Table 4, DMUs 1 and 3 are again found to be efficient. But the slacks for inefficient DMUs are increased which is because we use the new constraints on the data for output 2.

We then retrieve exact data y_{rj}^*, x_{ij}^* from the optimal solutions to (12) via equations in (5). As a result, Table 6 shows the retrieved data for each of the five DMUs to be evaluated. As shown in Table 6, we know that the same data for output 2 (as well as output 1 and input 1) are used in (12) for all five DMUs while the slightly different data for input 2 are used across DMUs to be evaluated.

In any case, we can use the retrieved data in the second stage to obtain individual slacks and lambdas. These results are summarized in the last two columns of Table 5. We thus know that the increased inefficiencies as compared with the first column of Table 4 are specified as the positive slacks for output 2 as shown in the third column of Table 5. All the values of lambda are equivalent to those in Table 4.

Table 5. Total inefficiencies and individual slacks in the second example

DMUs	z_0^* from (12)	Optimal solutions obtained from the second stage		
		z_0^*	Slacks (>0)	Lambdas (>0)
1	0	0	none	$\lambda 1^* = 1$
2	-1378.571	1378.571	s_1^{+*}= 1285.714, s_2^{+*}= 37.143 s_1^{-*}= 35.714	$\lambda 1^* = 1.143$
3	0	0	none	$\lambda 3^* = 1$
4	-1240	1240	s_1^{+*}= 1100, s_2^{+*}= 40 s_1^{-*}= 100	$\lambda 1^* = 1$
5	-2365.714	2365.714	s_1^{+*}= 2257.143, s_2^{+*}= 51.428 s_1^{-*}= 57.143	$\lambda 1^* = 1.429$

Table 6. The exact data retrieved from model (12) when z_0* is obtained[1]

DMUs	DMU0 1		DMU0 2		DMU0 3		DMU0 4		DMU0 5	
	y_{2j}^*	x_{2j}^*	y_{2j}^*	x_{2j}^*	Y_{2j}^*	x_{2j}^*	y_{2j}^*	x_{2j}^*	y_{2j}^*	x_{2j}^*
1	50	**0.6**	50	**0.7**	50	0.7	50	0.7	50	0.7
2	20	0.8	20	0.8	20	0.8	20	0.8	20	0.8
3	100	1	100	1	100	1	100	1	100	1
4	10	0.8	10	0.8	10	**0.8**	10	**0.7**	10	0.7
5	20	1	20	1	20	1	20	1	20	1

주) 1. We omitted here the data y_{1j}^*, x_{1j}^* because these are the same as those in Table 3.

6. Conclusions

We have provided methods to deal with imprecise data in DEA evaluations of performance. This has been done in the manner that arbitrary linear imprecise data are encountered in an additive DEA model. To achieve the inefficiency of DMUs, we have taken into consideration a two-stage way. In the first stage, we obtained an aggregated measure of inefficiency together with the optimal solutions to represent exact data. The aggregated measure of inefficiency is then separated into (or specified by) individual inefficiencies in terms of slacks. We have thus obtained specific inefficiencies as well as peer groups and scale sizes in an IDEA problem.

Our developments may also apply to other DEA models, such as CCR and BCC models, in which the data are known imprecisely. As mentioned in the body of this paper, in doing so there may exist a difficulty which is associated with the problem of non-Archimedean element underlying CCR and BCC models. This implies that we should achieve the technical efficiency without specifying the positive value of ε. As noted in DEA studies, attempts to employ small numbers in place of can lead to problematic results for technical efficiency (see Ali and Seiford, 1993). Otherwise the efficiency measured with positive would be regarded as an AR efficiency because the conditions μ, $\omega \geq \varepsilon$ can be viewed as examples of AR bounds (see Thompson et al., 1995).

To discuss the difficulty in detail, let us try to obtain the radial efficiency (so-called θ^*) in the first stage. This can be done in a way that we set ε=0 in the model. Assume that we then have the optimal solutions for the variables Yrj, Xij and μ, ω, used to achieve the θ^*. Next, we may try to retrieve exact data via the equations, $y_{rj}^* = Y_{rj}^* / \mu_r^*$ and $x_{ij}^* = X_{ij}^* / \omega_i^*$, for use them in the second stage to obtain slacks. However, it is possible that some μ_r^* and ω_i^* are to be zero. So we may fail to retrieve exact data without assumptions like those y_{rj}^*, x_{ij}^* are zero when μ_r^*, ω_i^* are zero. We thus invite developments of methods which can deal with technical efficiency analysis in CCR or BCC IDEA models.

Finally, we want to note that there are advantages to the approach developed in Zhu (2000b-c). This enables us to obtain the necessary information on DEA evaluations such as slacks as well as peer groups and scale sizes in a simpler manner. Specifically, we can exclude computational efforts which are encountered in the first stage developed in this paper. Thus, extensions and generalizations of Zhus approach will also be opportunities for further research toward broadening the use of DEA involving imprecise data problems.

References

1. Ali AI and Seiford LM. Computational accuracy and infinitesimals in data envelopment analysis. *INFOR* 1993; **31**: 290-297.
2. Allen R, Athanassopoulos A, Dyson RG and Thanassoulis E. Weight restrictions and value judgments in data envelopment analysis: Evolution, development and future directions. *Ann. Opns. Res.* 1997; **73**: 13-34.
3. Banker RD, Charnes A and Cooper WW. Some models for estimating technical and scale inefficiencies in data envelopment analysis. *Management Science* 1984; **30/9**: 1078-1092.
4. Brockett PL, Charnes A, Cooper WW, Huang ZM and Sun DB. Data transformations in DEA cone ratio envelopment approaches for monitoring bank performances. *Eur. J. Opl. Res.* 1997; **98**: 250-268.
5. Charnes A, Cooper WW, Golany B, Seiford LM and Stutz J. Foundations of data envelopment analysis for Pareto-Koopmans efficient empirical production functions. *J. Econometrics* 1985; **30**: 91-107.

6. Charnes A, Cooper WW, Huang ZM and Sun DB. Polyhedral cone-ratio DEA models with an illustrative application to large commercial banks. *J. Econometrics* 1990; **46**: 73-91.
7. Charnes A, Cooper WW and Rhodes E. Measuring the efficiency of decision making units. Euro. *J. Oprl. Res.* 1978; **2**: 429-444.
8. Chen Y, Seiford LM and Zhu J. Imprecise data envelopment analysis. Submitted to *Management Science* (2000).
9. Cooper WW, Park KS and Pastor JT. RAM: A range adjusted measure of efficiency for use with additive models, and relations to other models and measures in DEA. *J. Productivity Analysis* 1999; **11**: 5-42.
10. Cooper WW, Park KS and Yu G. IDEA and AR-IDEA: Models for dealing with imprecise data in DEA. *Management Science* 1999; **45**: 597-607.
11. Cooper WW, Park KS and Yu G. An illustrative application of IDEA (Imprecise Data Envelopment Analysis) to a Korean mobile telecommunication company. Accepted for publication in *Operations Research* (2000a).
12. Cooper WW, Park KS and Yu G. IDEA (Imprecise Data Envelopment Analysis) with CMDs (Column Maximum Decision Making Units) absent. Accepted for publication in *J. Oprl. Res. Soc.* (2000b).
13. Kim SH, Park CG, and Park KS. An application of data envelopment analysis in telephone offices evaluation with partial data. *Computers and Operations Research* 1999; **26**: 59-72.
14. Park KS. Revisit to IDEA and AR-IDEA: Models for dealing with imprecise data in DEA. Submitted to *J. Oprl. Res.* Soc. 2000.
15. Thompson RG, Dharmapala PS and Thrall RM. Linked-cone DEA profit ratios and technical efficiency with application to Illinois coal mines. *Int. J. Production Economics* 1995; **39**: 99-115.
16. Thompson RG, Langemeier LN, Lee CT, Lee E and Thrall RM. The role of multiplier bounds in efficiency analysis with applications to Kansas farming. *J. Econometrics* 1990; **46**: 93-108.
17. Zhu J. Models for dealing with imprecise data in DEA: Simplified and improved. Submitted to *Management Science* (2000a).
18. Zhu J. An application of data envelopment analysis in telephone offices evaluation with partial data: Revisited. Submitted to *Computers and Operations Research* (2000b).
19. Zhu J. Imprecise DEA via standard linear DEA models with a revisit to the Korean mobile telecommunication company. Submitted to *Operations Research* (2000c).