Scale Efficiency Measurement in Data Envelopment Analysis with Interval Data: A Two-Level Programming Approach

Chiang Kao *
National Cheng Kung University, Taiwan, Republic of China

Shiang-Tai Liu
Vanung University, Taiwan, Republic of China

Abstract

Conventional data envelopment analysis (DEA) for measuring the relative efficiency of a set of decision-making units (DMUs) requires the observations to have precise values. When observations are imprecise and represented by interval values, the efficiencies are also expected to reflect interval values. Several methods exist to calculate the interval overall and technical efficiencies, but such methods are unable to calculate the interval scale efficiency. The focus of this paper is the application of a two-level programming technique to formulate the problem of determining the bounds of the interval scale efficiency. The associated models are essentially nonlinear programs with only bound constraints for variables in a sophisticated form. Hence, one can modify the conventional quasi-Newton method for unconstrained nonlinear programming problems to solve the two-level programs. Two examples with interval data, one hypothetical and one real, aid in explaining the proposed method and the properties of the results.

Keywords: DEA, efficiency, interval data, uncertainty, two-level mathematical programming, nonlinear programming

JEL Classification codes: C61, C71, C72, M11

Data envelopment analysis (DEA) is a nonparametric approach to measuring the relative efficiency of a set of decision-making units (DMUs) that uses multiple inputs to produce multiple outputs. Due to the solid theoretical basis and wide application of DEA to real-world problems, many researchers have exerted effort in this area since the pioneering work of Charnes, Cooper, and Rhodes (1978). See, for example, the review articles of Cook and Seiford (2009) and Seiford (1997).

As evident in its name, the basis of DEA is observed data. All data, in their original settings, are assumed to be exactly known cardinal values. Cook, Kress, and Seiford (1993, 1996) extended the data type from cardinal to ordinal and qualitative. In the case of ordinal data, the data only have ordinal relationships, while qualitative data involve a verbal description, such as *excellent*, *good*, *fair*, *acceptable*, and *poor*. Cook and Zhu (2006) developed a general framework for modeling and handling qualitative data and provided a unified structure for embedding rank order data into the DEA framework. The values of these two types of data are exact, although unknown, and are determined through the associated program. The efficiencies obtained are precise values, considered as certainty cases.

When uncertainty is involved, the data usually reflect a range of values, as does the calculated efficiency (Mostafaee & Saljooghi, 2010). Cooper, Park, and Yu (1999, 2001) introduced the term *imprecise data* to mean interval and ordinal data, with the associated method being *imprecise DEA* (IDEA). Cooper et al. only calculated the most favorable (upper bound) efficiency, although the efficiency associated with interval data should also lie in ranges.

Several researchers have developed various solution techniques to calculate the upper bound efficiency (Dyson & Shale, 2010; Park, 2010; Zhu, 2004). Other researchers devised different solution methods to calculate both the upper and lower bounds of efficiency (Despotis & Smirlis, 2002; Kao, 2006; Park, 2007; Zhu, 2003). The efficiency obtained is thus an interval measure. Kao and Liu (2004) used predicted data, as interval measures, to predict efficiency, and Kao and Liu (2000) and Smirlis, Maragos, and Despotis (2006) employed interval data to represent missing values to calculate efficiency when some data are missing. In summation, intervals are a common method of representing unknown data.

In all of the studies concerning interval data, researchers limit the discussion to either the overall efficiency, measured through the CCR model (Charnes et al., 1978), or the technical efficiency, measured through the BCC model (Banker, Charnes, & Cooper, 1984), where a linear programming technique may result in a solution. No researchers have investigated the scale efficiency, defined as the ratio of the CCR efficiency to the BCC efficiency. A graphical example in this paper shows that the straightforward calculation of the ratio of CCR efficiency to BCC efficiency is complicated when the data have interval values. A two-level programming technique to formulate the problem is necessary to determine the interval scale efficiency. General two-level programming programs are difficult to solve (Amouzegar, 1999). However, careful examination indicates that the two-level models are essentially nonlinear programs with only bound constraints for variables in a sophisticated form. Therefore, one can modify the existing efficient and effective solution methods for unconstrained nonlinear programs for a solution.

An interpretation of the two-level program is apparent in the Stackelberg game (Simaan & Cruz, 1973). One could also interpret the formulation of the problem for calculating the lower and upper bound efficiencies using the leader-follower concept in game theory. Liang, Wu, Cook, and Zhu (2008) used the game approach to discuss the cross-efficiency among all DMUs and the Nash equilibrium, while Liang, Cook, and Zhu (2008) employed the game approach to model the efficiency decomposition in two-process series production systems. Such ideas are inspiring to researchers in this area.

The following section includes a simple example with only one interval value in a two-dimensional space to illustrate the difficulty involved in calculating the scale efficiency when the data have interval values. An introduction to the two-level programming technique for modeling the problem of calculating the interval scale efficiency appears next before development of the solution methods for the two-level programs. Two examples, one hypothetical and one real case, are apparent to illustrate the idea.

Graphical Illustration

Let X_{ij} and Y_{rj} denote the ith input, $i = 1,\ldots, m$, and the rth output, $r = 1,\ldots, s$, respectively, of the jth DMU, $j = 1,\ldots, n$. The CCR model for calculating the efficiency of DMU k under the assumption of constant returns-to-scale is

$$E_k^{CCR} = \text{Max} \sum_{r=1}^{s} u_r Y_{rk}$$

s.t.

$$\sum_{i=1}^{m} v_i X_{ik} = 1 \tag{1}$$

$$\sum_{r=1}^{s} u_r Y_{rj} - \sum_{i=1}^{m} v_i X_{ij} \leq 0, \ \ j = 1,\ldots, n,$$

$$u_r, v_i \geq \varepsilon, r = 1,\ldots, s, \quad i = 1,\ldots, m,$$

where u_r and v_i are virtual multipliers, and ε is a small non-Archimedean number imposed to avoid ignoring any factor (Charnes & Cooper, 1984; Charnes, Cooper, & Rhodes, 1979). The CCR efficiency, E_k^{CCR}, is usually called the overall efficiency.

A version of the BCC model for measuring the efficiency of DMU k under the assumption of variable returns-to-scale is

$$E_k^{BCC} = \max \sum_{r=1}^{s} u_r Y_{rk}$$

s.t.

$$v_0 + \sum_{i=1}^{m} v_i X_{ik} = 1,$$

$$\sum_{r=1}^{s} u_r Y_{rj} - (v_0 + \sum_{i=1}^{m} v_i X_{ij}) \leq 0, \ j = 1, \ldots, n, \tag{2}$$

$$u_r, v_i \geq \varepsilon, \ r = 1, \ldots, s, i = 1, \ldots, m$$

v_0 unrestricted in sign.

The BCC efficiency, E_k^{BCC}, is termed the technical efficiency. The ratio of the CCR efficiency to the BCC efficiency is the scale efficiency, E_k^{Scale} (Banker et al., 1984):

$$E_k^{Scale} = E_k^{CCR} / E_k^{BCC}. \tag{3}$$

Consider the simple example of four DMUs, A, B, C, and D, applying 8, 12, 16, and 20 units of input X to produce 6, [7, 10], 8, and 12 units of output Y, respectively (see Figure 1).
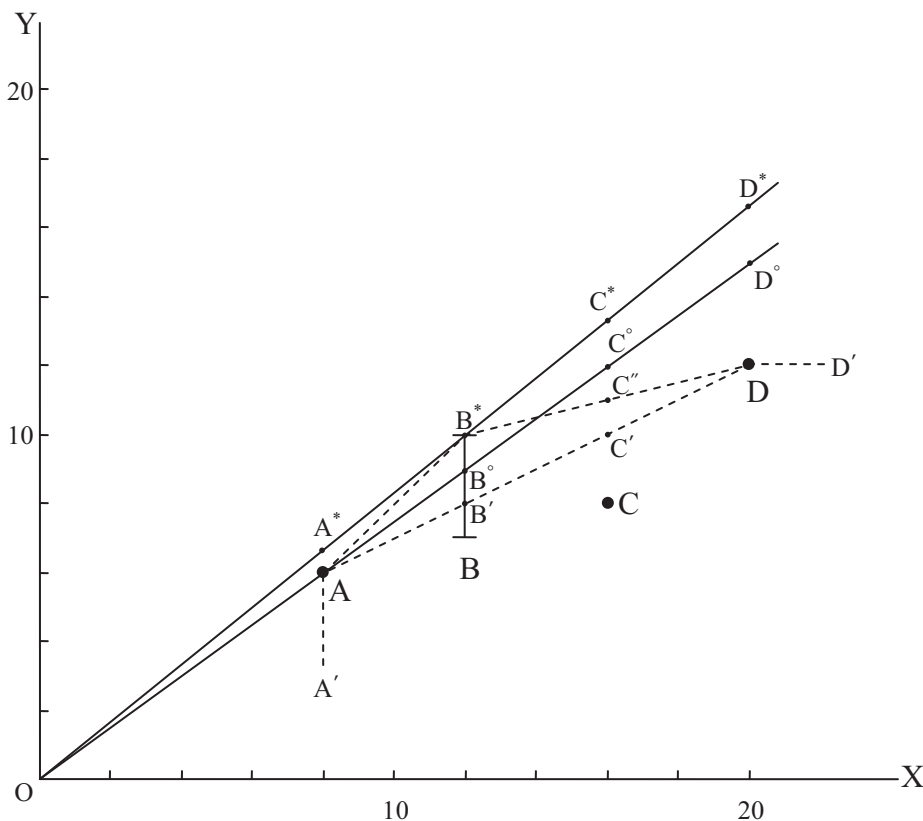

Figure 1. Production frontiers constructed by the CCR and BCC models with one interval data Y_B.

Note that the output of DMU B, Y_B, is imprecise, with values lying in the range of [7, 10]. Discussion of three cases ($7 \leq Y_B \leq 8$, $8 \leq Y_B \leq 9$, and $9 \leq Y_B \leq 10$) is important in measuring the relative efficiency of the four DMUs.

Range $7 \leq Y_B \leq 8$

For Y_B in the range of 7 and 8, the production frontiers constructed by the CCR model (Equation 1) and the BCC model (Equation 2) are Ray OA and the piecewise line segments A'ADD', respectively. Let B°, C°, and D° denote the target points of B, C, and D on the CCR frontier, and let B' and C' denote the target points of B and C on the BCC frontier, respectively. Then, the CCR efficiencies are 1, Y_B/Y_{B° (=$Y_B/9$), Y_C/Y_{C° (=2/3), and Y_D/Y_{D° (=4/5), and the BCC efficiencies are 1, $Y_B/Y_{B'}$ (= $Y_B/8$), $Y_C/Y_{C'}$ (=4/5), and 1, for DMUs A, B, C, and D, respectively. The CCR and BCC efficiencies for DMUs A, C, and D are constant, but for DMU B, they vary in the interval of [7/9, 8/9] and [7/8, 1], respectively, because Y_B lies in the interval of [7, 8].

The scale efficiencies, which are the ratios of the CCR and BCC efficiencies, for the four DMUs are 1, 8/9, 5/6, and 4/5, respectively. Interestingly, the scale efficiency for DMU B is a constant of 8/9, even though Y_B has an interval value. Note also that for each specific value of Y_B, the scale efficiency is the ratio of the CCR efficiency to the BCC efficiency, which is a constant of 8/9 in this case. However, for the whole range of Y_B (i.e., [7, 8]), the CCR and BCC efficiencies of DMU B have interval values of [7/9, 8/9] and [7/8, 1], respectively, where their ratio cannot be defined. Table 1, under the heading $7 \leq Y_B \leq 8$, includes a summary of the results.

Table 1
Overall, Technical, and Scale Efficiencies of the Graphical Example

DMU	Overall efficiency		Technical efficiency		Scale efficiency	
	Value	Range	Value	Range	Value	Range
			$7 \leq Y_B \leq 8$			
A	1	[1, 1]	1	[1, 1]	1	[1, 1]
B	$Y_B/9$	[7/9, 8/9]	$Y_B/8$	[7/8, 1]	8/9	[8/9, 8/9]
C	2/3	[2/3, 2/3]	4/5	[4/5, 4/5]	5/6	[5/6, 5/6]
D	4/5	[4/5, 4/5]	1	[1, 1]	4/5	[4/5, 4/5]
			$8 \leq Y_B \leq 9$			
A	1	[1, 1]	1	[1, 1]	1	[1, 1]
B	$Y_B/9$	[8/9, 1]	1	[1, 1]	$Y_B/9$	[8/9, 1]
C	2/3	[2/3, 2/3]	$16/(Y_B+12)$	[16/21, 4/5]	$(Y_B+12)/24$	[5/6, 7/8]
D	4/5	[4/5, 4/5]	1	[1, 1]	4/5	[4/5, 4/5]
			$9 \leq Y_B \leq 10$			
A	$9/Y_B$	[9/10, 1]	1	[1, 1]	$9/Y_B$	[9/10, 1]
B	1	[1, 1]	1	[1, 1]	1	[1, 1]
C	$6/Y_B$	[3/5, 2/3]	$16/(Y_B+12)$	[8/11, 16/21]	$3(Y_B+12)/(8Y_B)$	[33/40, 7/8]
D	$7.2/Y_B$	[18/25, 4/5]	1	[1, 1]	$7.2/Y_B$	[18/25, 4/5]
			Summary			
A		[9/10, 1]		1		[9/10, 1]
B		[7/9, 1]		[7/8, 1]		[8/9, 1]
C		[3/5, 2/3]		[8/11, 4/5]		[33/40, 7/8]
D		[18/25, 4/5]		1		[18/25, 4/5]

Range $8 \leq Y_B \leq 9$

When Y_B lies in the interval of 8 and 9 (i.e., DMU B appears between B' and B° in Figure 1), the production frontier constructed by the CCR model is still the ray of OA. Hence, the CCR efficiencies of the four DMUs are the same as those of the first case: 1, $Y_B/9$, 2/3, and 4/5. The frontier constructed by the BCC model is the line segments A'ABDD', where B is a point between B' and B°. In this case, DMUs A, B, and D are on the frontier with a perfect BCC efficiency of 1, and their scale efficiencies are the same as their CCR efficiencies. For DMU C, its target point on the BCC frontier A'ABDD' has an output value of $Y_B + (X_C - X_B)(Y_D - Y_B)/(X_D - X_B) = (Y_B + 12)/2$. Therefore, its BCC efficiency is $8/[(Y_B + 12)/2] = 16/(Y_B + 12)$, and its scale efficiency is $(8/12)/[16/(Y_B + 12)] = (Y_B + 12)/24$.

Substituting the possible values of Y_B into the expression of the three types of efficiency for the DMUs produces CCR efficiencies of 1, [8/9, 1], 2/3, and 4/5; BCC efficiencies of 1, 1, [16/21, 4/5], and 1; and scale efficiencies of 1, [8/9, 1], [5/6, 7/8], and 4/5 for A, B, C, and D, respectively. Table 1, under the heading $8 \leq Y_B \leq 9$, includes a summary of the results. In this case, although only DMU B has interval data, DMU C also has interval efficiency measures.

Range $9 \leq Y_B \leq 10$

For DMU B with an output value in the range of 9 and 10, the frontier constructed by the BCC model has the same form as that of the second case: A'ABDD'. Hence, DMUs A, B, C, and D have the same BCC efficiencies as those of the second case: 1, 1, $16/(Y_B + 12)$, and 1, respectively. The frontier constructed by the CCR model, in contrast, has been raised to the ray of OB, where B is a point between B^o and B^* in Figure 1. In this case, only DMU B has a perfect CCR efficiency of 1. Ray OB has a slope of $Y_B/12$.

Consequently, the target points for DMUs A, C, and D have output values of $X_A \times Y_B/12$ (=$2Y_B/3$), $X_C \times Y_B/12$ (=$4Y_B/3$), and $X_D \times Y_B/12$ (=$5Y_B/3$), which produce the CCR efficiencies of $9/Y_B$, $6/Y_B$, and $7.2/Y_B$, respectively. The scale efficiencies of the four DMUs, thus, become $9/Y_B$, 1, $3(Y_B + 12)/(8Y_B)$, and $7.2/Y_B$, respectively. By substituting all possible values of Y_B into the efficiency expressions of the four DMUs, one can calculate the intervals of the three types of efficiency, evident in Table 1 under the heading $9 \leq Y_B \leq 10$. Interestingly, all DMUs, except for B, have interval efficiencies, while only B has interval data.

Combining the three cases, Table 1 illustrates the results for Y_B lying in the range of 7 and 10 under the heading *Summary*. For the four DMUs, the overall efficiencies are [9/10, 1], [7/9, 1], [3/5, 2/3], and [18/25, 4/5]; the technical efficiencies are 1, [7/8, 1], [8/11, 4/5], and 1; and the scale efficiencies are [9/10, 1], [8/9, 1], [33/40, 7/8], and [18/25, 4/5], respectively. Because the technical efficiencies of DMUs A and D are a constant of 1, their interval scale efficiencies are the same as their respective interval overall efficiencies. For DMU B, the lower and upper bounds of its scale efficiencies are the respective ratios of those of its overall and technical efficiencies. DMU C is more complicated. Its lower bound scale efficiency occurs at the ratio of those of the overall and technical efficiencies. Its upper bound efficiency, however, occurs at the overall efficiency of 2/3 and technical efficiency of 16/21, where 16/21 is not a bound value of the technical efficiency. Figure 2 depicts the overall, technical, and scale efficiencies of DMU C at different values of Y_B.
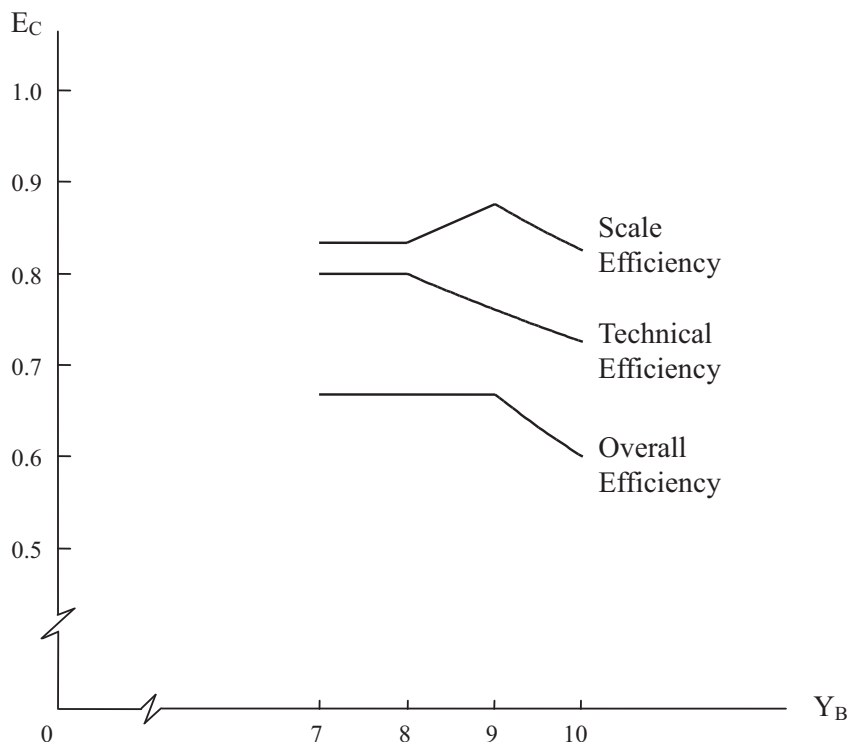

Figure 2. Overall, technical, and scale efficiencies of DMU C changing with the value of Y_B in the graphical example.

Three observations are necessary. First, although only DMU *B* has interval data, all four DMUs have different types of interval efficiency. Specifically, apart from the fact that the technical efficiencies of DMUs *A* and *D* are a constant of 1, all three types of efficiency for all four DMUs have interval values. Second, while the point scale efficiency is defined as the ratio of the CCR to the BCC efficiencies for each specific value of Y_B, one cannot define the interval scale efficiency directly as the ratio of the interval overall efficiency to the interval technical efficiency. The interval scale efficiency of DMU *C* is an example. Third, and most important, the bound values of the interval scale efficiency may not occur at the bound values of the interval data.

Note that in calculating the upper bound of the CCR or BCC efficiency of any DMU, the most favorable condition for this DMU is chosen, while in calculating the lower bound, the most unfavorable condition is used (Kao, 2006). For example, in calculating the upper (lower) bound of the interval efficiency of DMU *A*, *C*, or *D*, one must set the output of *B*, Y_B, to the lower (upper) bound, and in calculating the upper (lower) bound of the interval efficiency of DMU *B*, one must set Y_B to the upper (lower) bound. For scale efficiency, however, this is not the case. For example, the largest scale efficiency of DMU *C*, 7/8, occurs at the interior value of 9 for Y_B (see Figure 2). This property of an interior-point solution causes the major difficulty in devising a solution method for calculating the interval scale efficiency.

The Two-Level Programming Model

When all the data have exact values, conventional DEA models for calculating the relative efficiency are useful. In the case of some data being inexact, represented by intervals, the measured efficiencies also have interval values. Without loss of generality, suppose all the data have interval values. Note that one can consider exact values as degenerated interval values with only one value in the interval. Let [X_{ij}^L , X_{ij}^U], [Y_{rj}^L , Y_{rj}^U], and [E_k^L , E_k^U] denote the ranges of X_{ij}, Y_{rj}, and E_k, respectively. To determine the interval in which the efficiencies lie, find the lower bound E_k^L and upper bound E_k^U of the efficiencies.

In calculating the efficiency of DMU *k*, the idea is to assign a set of X_{ij} and Y_{rj} values from their respective ranges to all *n* DMUs, then apply Equations 1, 2, or 3, depending on whether the type of efficiency required is overall, technical, or scale, respectively. Different sets of X_{ij} and Y_{rj} values produce different efficiency values. One can formulate the problem of finding the sets of X_{ij} and Y_{rj} values, which produce E_k^L and E_k^U, as a two-level mathematical program, which has a specific interpretation from the viewpoint of the Stackelberg game.

Consider two players, a supplier and producer, labeled *k* in a set of *n* producers. The supplier determines the amount of resources X_{ij} and Y_{rj} allocated to the *n* producers. Based on the resources acquired, producer *k* applies the most appropriate technology, through the technology parameters u_r and v_i, to convert the resources into a product called *efficiency*.

In the conventional Stackelberg game, one of the two players, the leader, knows the profit function of the second player, the follower, who may or may not know the profit function of the leader. However, the follower knows the strategy of the leader and considers this in determining his or her own strategy. In addition, the leader fully understands the reactions of the follower and can optimize his or her own choice of strategy. In the current case, the supplier is the leader, whose job is to allocate resources to the *n* producers, while DMU *k* is the follower, who uses the allocated resources to maximize efficiency. Unable to obtain any profit (i.e., efficiency), the leader is jealous of the profit earned by DMU *k* and will make the least favorable allocation for DMU *k* to generate the smallest value of efficiency. The associated mathematical model for CCR efficiency follows:

$$E_k^L = \operatorname*{Min}_{\substack{X_{ij}^L \le X_{ij} \le X_{ij}^U \\ Y_{rj}^L \le Y_{rj} \le Y_{rj}^U \\ \forall i,r,j}} \operatorname*{Max}_{\substack{u_r, v_i \\ \forall r,i}} \sum_{r=1}^{s} u_r Y_{rk}$$

s.t.

$$\sum_{i=1}^{m} v_i X_{ik} = 1 \tag{4}$$

$$\sum_{r=1}^{s} u_r Y_{rj} - \sum_{i=1}^{m} v_i X_{ij} \le 0, \ j = 1, \dots, n,$$

$$u_r, v_i \ge \varepsilon, \ r = 1, \dots, s, \ i = 1, \dots, m.$$

In the first-level, or outer, program, the supplier determines the amount of various resources allocated to all n producers, while in the second-level, or inner, program, producer k selects the most appropriate technology for production.

One can treat the problem of finding the upper bound efficiency value E_k^U as a cooperative game. Knowing that the supplier is jealous of the profit that producer k can earn and is making the least favorable allocation of resources, producer k decides to cooperate with the supplier by sharing some of the profit. To maximize the gain, the supplier will help producer k to generate the largest amount of profit from the resources under its control. Because the supplier is aware of the production mechanism of producer k, it will supply the specific amount of resources to enable producer k to generate the largest amount of profit. One can formulate the problem as follows:

$$E_k^U = \underset{\substack{X_{ij}^L \le X_{ij} \le X_{ij}^U \\ Y_{rj}^L \le Y_{rj} \le Y_{rj}^U \\ \forall i,r,j}}{\text{Max}} \underset{\substack{u_r,v_i \\ \forall r,i}}{\text{Max}} \sum_{r=1}^{s} u_r Y_{rk}$$

s.t.

$$\sum_{i=1}^{m} v_i X_{ik} = 1 \qquad (5)$$

$$\sum_{r=1}^{s} u_r Y_{rj} - \sum_{i=1}^{m} v_i X_{ij} \le 0, \ j = 1,\ldots, n$$

$u_r, v_i \ge \varepsilon, \ r = 1,\ldots, s, \ i = 1,\ldots, m.$

After calculating the lower bound E_k^L and the upper bound E_k^U from Equations 4 and 5, respectively, one obtains the interval of the efficiency as $[E_k^L, E_k^U]$. If the product is technical efficiency, replace the second-level programs of Equations 4 and 5 with the BCC model (Equation 2). Similarly, if the product is scale efficiency, replace the inner program with Equation 3.

Solution Methods

Equations 4 and 5 are special cases of the two-level programming model of the following form (Bialas & Karwan, 1984):

$$\begin{aligned} &\underset{\mathbf{x}}{\text{Max}} \ f(\mathbf{x}, \mathbf{y}) & &\text{where } \mathbf{y} \text{ solves} \\ &\underset{\mathbf{y}}{\text{Max}} \ g(\mathbf{x}, \mathbf{y}) & &\text{s.t. } (\mathbf{x}, \mathbf{y}) \in S \end{aligned} \qquad (6)$$

where S denotes the feasible region. In the current case, f and g are the same. When the product is overall efficiency (Equation 1) or technical efficiency (Equation 2), it falls in the category of linear two-level programming because all the associated functions are linear. In contrast, if the product is scale efficiency (Equation 3), it falls in the category of nonlinear two-level programming because the objective function is nonlinear.

Several authors have proposed methods for solving linear two-level programs (e.g., see the review of Vicente & Calamai, 1994). However, due to the special structure of Equations 4 and 5, easier solution methods can be devised (Cook et al., 1996; Kao, 2006; Zhu, 2003, 2004). The focus of this paper will be the calculation of scale efficiency where the associated two-level models are nonlinear.

For nonlinear two-level programming problems, the existing solution methods require the objective functions to be convex (Amouzegar, 1999; Bard 1988). Furthermore, as stated in Amouzegar (1999), solving large-scale problems, with up to 30 decision variables, is extremely difficult. The two-level models for calculating the lower and upper bounds of the interval scale efficiencies have the following structure:

$$(E_k^{Scale})^L = \underset{\substack{X_{ij}^L \le X_{ij} \le X_{ij}^U \\ Y_{rj}^L \le Y_{rj} \le Y_{rj}^U \\ \forall i,r,j}}{\text{min}} \ E_k^{CCR} / E_k^{BCC} \qquad (7)$$

$$(E_k^{Scale})^U = \underset{\substack{X_{ij}^L \le X_{ij} \le X_{ij}^U \\ Y_{rj}^L \le Y_{rj} \le Y_{rj}^U \\ \forall i,r,j}}{\text{max}} \ E_k^{CCR} / E_k^{BCC} \qquad (8)$$

where E_k^{CCR} and E_k^{BCC} are defined in Equations 1 and 2, respectively. Denote $f(X, Y) = E_k^{CCR}(X, Y) / E_k^{BCC}(X, Y)$. Equations 7 and 8 are actually nonlinear programming problems with only bound constraints for variables. Although the form of the objective function f is unknown, one can calculate its value once a set of (X, Y) values is evident. Therefore, a numerical method for unconstrained nonlinear programming problems can be devised.

Conceptually, one starts with a trial feasible point $(X^{(0)}, Y^{(0)})$, for example, the center of the feasible region. Next, generate a search direction for making improvement, and conduct a line search along the search direction in the feasible region to reach a new point $(X^{(1)}, Y^{(1)})$. From this new point, generate another search direction and subsequent line search in the feasible region to reach another point $(X^{(2)}, Y^{(2)})$. Repeat the process until the search direction has a length close to zero. To accelerate the convergence, one can incorporate the idea of quasi-Newton modification of the search direction, such as the DFP or BFGS formula, into the algorithm (Fletcher, 1987). Every time one reaches the boundary of the feasible region in the line search, the update of the quasi-Newton formula must start anew.

Equations 7 and 8 are actually the same, except that one searches for the minimum and the other for the maximum. Because the problem is similar to an unconstrained nonlinear program, finding the optimal solution is quite easy. One could also calculate the intervals of the overall and technical efficiencies using this nonlinear programming solution method, but it is not as efficient as the existing methods.

Examples

The section includes two examples to illustrate how to use the method developed in this paper to calculate the interval scale efficiency when some observations have interval values. A discussion of the characteristics of the solutions obtained accompanies the examples. The first example is hypothetical, and the second is a real case.

Example 1

The first example involves four DMUs using one input X to produce one output Y (see Table 2). Of the eight observations, four have precise values, and four have interval values. Table 2 shows the overall (CCR) and technical (BCC) efficiencies of each DMU, solved using the method of Kao (2006). Equations 7 and 8 aided in calculating the scale efficiencies (see Table 2). The nonlinear solution method developed in this paper was first used to calculate the overall and technical efficiency bounds via Equations 4 and 5 to verify that it was capable of finding optimal solutions for the two-level type of problems before it was used to solve Equations 7 and 8.

Table 2
Data and Interval Efficiencies of Example 1

DMU	Input	Output	Overall efficiency	Technical efficiency	Scale efficiency
A	[9, 12]	4	[0.6250, 1]	[1, 1]	[0.6250, 1]
B	30	[12, 16]	[0.8182, 1]	[0.8276, 1]	[0.9421, 1]
C	40	11	[0.5156, 0.6875]	[0.5500, 0.9167]	[0.6750, 0.9996]
D	[45, 55]	[12, 22]	[0.4091, 1]	[0.7498, 1]	[0.4909, 1]

All four DMUs have interval overall efficiencies, and three have interval technical efficiencies. DMU A has a precise technical efficiency of 1; thus, its interval scale efficiency is the same as its interval overall efficiency. One cannot obtain the interval scale efficiencies for the other three DMUs directly as the ratio of the interval overall efficiency to the interval technical efficiency. The upper bounds of the scale efficiencies of DMUs B and D are both 1, which is the ratio of their respective upper bounds of overall and technical efficiencies. Their lower bounds, in contrast, have no relationship with the bounds of their overall and technical efficiencies.

The situation is even more complicated for DMU C. The value obtained using Equation 8 to calculate the upper bound of the scale efficiencies of DMU C is 0.9996, which occurs at $A = (10.5, 4)$, $B = (30, 12)$, $C = (40, 11)$, and $D = (49.22, 19.71)$, where the overall and technical efficiencies are 0.6866 and 0.6869, respectively. Obviously, this problem has multiple solutions. However, careful study will show that none of the extreme

points of $D = \{(X, Y), | X \in [45, 55], Y \in [12, 22]\}$, the set for the interval data D, is the optimal solution, and 0.6866 and 0.6869 are interior points of their interval overall and technical efficiencies, respectively. Similarly, the value obtained using Equation 7 to calculate the lower bound of the scale efficiencies is 0.6750, which occurs at $A = (9, 4)$, $B = (30, 12)$, $C = (40, 11)$, and $D = (54.47, 12)$, where the overall and technical efficiencies are 0.6188 and 0.9167, respectively. Note that two extreme points for $D = (45, 12)$ and $(55, 12)$ are alternative solutions. Nevertheless, both 0.6188 and 0.9167 are still interior points of their respective intervals.

In this example, only four observations have interval values, meaning that there are only four variables in the nonlinear programs (Equations 7 and 8). Hence, finding the solution is very easy. Note also that the overall efficiency is the product of the technical and scale efficiencies, and the range of the interval overall efficiencies is expected to be wider than those of the interval technical and scale efficiencies. This phenomenon is apparent for all of the DMUs, except C, as is clear from the values in Table 2.

Example 2

Chien, Chen, Lo, and Lin (2007) measured the performance of eight thermal power plants in Taiwan. The inputs considered were total installed capacity, total number of employees, and total production cost. Only one output existed: total energy generated. The data spanned six years (1994 to 1999). Because the data are unstable, one assumes that they are interval-valued. The smallest and largest values observed over the six years reflect the bounds of the interval data. Table 3 indicates the results.

Table 3
Data of Example 2

Plant	Installed capacity (X_1, kMW)	Employees (X_2, people)	Production cost (X_3, \$10^6 NT)	Energy generation (Y, billion kWh)
A	2000	[416, 471]	[11137, 14623]	[10861, 12844]
B	[785, 1085]	[360, 411]	[3422, 4133]	[3124, 3609]
C	400	[260, 304]	[2604, 2903]	[1627, 2718]
D	[2480, 4680]	[601, 782]	[8525, 19263]	[15625, 33040]
E	[2100, 4625]	[628, 694]	[8775, 10765]	[12807, 14369]
F	2572	[546, 616]	[12038, 16293]	[10582, 13483]
G	[1415, 1628]	[478, 562]	[5888, 9086]	[5284, 7621]
H	[49, 74]	[124, 136]	[384, 832]	[178, 248]

By using the methods developed in this paper, one can calculate the interval, overall, technical, and scale efficiencies, as shown in Table 4.

Table 4
Interval Efficiencies of Example 2

Plant	Overall efficiency	Technical efficiency	Lower bound correspondence	Scale efficiency
A	[0.4195, 1]	[0.4506, 1]	0.6180/0.7532	[0.8205, 1]
B	[0.2162, 1]	[0.2197, 1]	0.5703/0.6968	[0.8184, 1]
C	[0.3054, 1]	[0.3260, 1]	0.4077/0.4693	[0.8689, 1]
D	[0.6765, 1]	[1, 1]	0.6765/1	[0.6765, 1]
E	[0.3357, 1]	[0.3876, 1]	0.7385/0.8696	[0.8492, 1]
F	[0.3125, 1]	[0.3203, 1]	0.5216/0.5720	[0.9118, 1]
G	[0.2436, 1]	[0.2452, 1]	0.3195/0.3671	[0.8703, 1]
H	[0.1804, 0.9332]	[1, 1]	0.1804/1	[0.1804, 0.9332]

Similar to the process followed in the previous case, the method of Kao (2006) was useful in calculating

the bounds of the overall and technical efficiencies, and Equations 7 and 8 aided in determining the scale efficiencies. In addition, the nonlinear solution method was first used to solve Equations 4 and 5 to verify whether it was able to obtain the same solution as that obtained from the method of Kao (2006) before it was used to solve Equations 7 and 8.

Of the eight power plants, *D* and *H* have a precise technical efficiency of 1; therefore, their interval scale efficiencies are the same as their interval overall efficiencies. For the remaining six power plants, the upper bounds of the interval scale efficiencies are the ratios of those of their respective overall to technical efficiencies. However, one cannot obtain the lower bounds in such a straightforward manner. See Table 4 (column 4) for the values of the overall and technical efficiencies corresponding to the lower bounds of the scale efficiencies. All of them are interior points of the corresponding intervals.

One can also use interval values for approximate ranking. In terms of technical efficiency, *D* and *H* have the best performance because they have a perfect value of 1. For the other six power plants, because all have an upper bound value of 1, the one with narrower intervals should be ranked higher. Therefore, the ranks are *A*, *E*, *C*, *F*, *G*, and *B*, in sequence.

A similar phenomenon to Example 1 is that the interval of the overall efficiencies is wider than are those of the technical and scale efficiencies for every power plant. The reason is the overall efficiency is the product of the technical and scale efficiencies, and multiplying two imprecise values should result in a more imprecise value. Finally, this problem has 29 interval observations. Hence, there are 29 variables in the nonlinear programming problem, which, according to Amouzegar (1999), is a large-scale problem. However, using the method presented in this paper makes the solution easy to obtain.

Conclusions

The literature reflects studies focused on measuring the overall (CCR) and technical (BCC) efficiencies in DEA for cases where observations are interval values. The basic idea is to use the least and most favorable conditions to calculate the smallest and largest efficiencies, respectively, to form the efficiency interval for each DMU. For scale efficiency, which is the ratio of the overall efficiency to the technical efficiency, the issue is not so straightforward. This paper involved applying a two-level programming technique to model the problem and to calculate the interval scale efficiency. The study included two key points.

First, the two-level programs for calculating the bounds of scale efficiencies are nonlinear. However, they are essentially unconstrained nonlinear programs (with only bound constraints for variables), which make the solution easy to obtain. Notably, the closed form of the objective function is unknown; one can only obtain the objective value via numerical calculations. Second, because the lower (upper) bound of the scale efficiencies is not simply the ratio of the lower (upper) bound of the overall efficiencies to the upper (lower) bound of the technical efficiencies, the interval of the former is not wider than those of the latter two. Conversely, the interval of the overall efficiencies is generally wider than are those of the other two because the former is conceptually the product of the latter two.

Interval efficiency measures can alert decision makers to the fact that the efficiency is not a fixed value, allowing them to plan subsequent decisions more carefully, considering uncertainty. However, the efficiency values close to the bounds of the interval efficiencies, either lower or upper, are the result of a combination of rare situations. The likelihood of their occurrence is low, so decision makers should not be affected too much by these extreme values. In this regard, determining the distribution of the scale efficiency is desirable, similar to the approach of Kao and Liu (2009) for overall and technical efficiencies. This is possible only when one knows the distributions of the data in addition to the intervals.

References

Amouzegar, M. A. (1999). A global optimization method for nonlinear bilevel programming problems. *IEEE Transactions on Systems, Man, and Cybernetics-B, 29*(6), 771-776.

Banker, R. D., Charnes, A., & Cooper, W. W. (1984). Some models for estimating technical and scale efficiencies in data envelopment analysis. *Management Science, 30*(9), 1078-1092.

Bard, J. F. (1988). Convex two-level optimization. *Mathematical Programming, 40*(1), 15-27.

Bialas, W. F., & Karwan, M. H. (1984). Two-level linear programming. *Management Science, 30*(8), 1004-1020.

Charnes, A., & Cooper, W. W. (1984). The non-Archimedean CCR ratio for efficiency analysis: A rejoinder to Boyd and Färe. *European Journal of Operational Research, 15*(3), 333-334.

Charnes, A., Cooper, W. W., & Rhodes, E. (1978). Measuring the efficiency of decision making units. *European Journal of Operational Research, 2*(6), 429-444.

Charnes, A., Cooper, W. W., & Rhodes, E. (1979). Short communication: Measuring the efficiency of decision making units. *European Journal of Operational Research, 3*(4), 339.

Chien, C. F., Chen, W. C., Lo, F. Y., & Lin, Y. C. (2007). A case study to evaluate the productivity change of the thermal power plants of the Taiwan Power Company. *IEEE Transactions on Energy Conversion, 22*(3), 680-688.

Cook, W. D., Kress, M., & Seiford, L. (1993). On the use of ordinal data in data envelopment analysis. *Journal of the Operational Research Society, 44*(2), 133-140.

Cook, W. D., Kress, M., & Seiford, L. (1996). Data envelopment analysis in the presence of both quantitative and qualitative factors. *Journal of the Operational Research Society, 47*(7), 945-953.

Cook, W. D., & Seiford, L. (2009). Data envelopment analysis (DEA)—Thirty years on. *European Journal of Operational Research, 192*(1), 1-17.

Cook, W. D., & Zhu, J. (2006). Rank order data in DEA: A general framework. *European Journal of Operational Research, 174*(2), 1021-1038.

Cooper, W. W., Park, K. S., & Yu, G. (1999). IDEA and AR-IDEA: Models for dealing with imprecise data in DEA. *Management Science, 45*(4), 597-607.

Cooper, W. W., Park, K. S., & Yu, G. (2001). An illustrative application of IDEA (imprecise data envelopment analysis) to a Korean mobile telecommunication company. *Operations Research, 49*(6), 807-820.

Despotis, D. K., & Smirlis, Y. G. (2002). Data envelopment analysis with imprecise data. *European Journal of Operational Research, 140*(1), 24-36.

Dyson, R. G., & Shale, E. A. (2010). Data envelopment analysis, operational research and uncertainty. *Journal of the Operational Research Society, 61*(1), 25-34.

Fletcher, R. (1987). *Practical methods of optimization* (2nd ed.). Chichester, UK: John Wiley & Sons.

Kao, C. (2006). Interval efficiency measures in data envelopment analysis with imprecise data. *European Journal of Operational Research, 174*(2), 1087-1099.

Kao, C., & Liu, S. T. (2000). Data envelopment analysis with missing data: An application to university libraries in Taiwan. *Journal of the Operational Research Society, 51*(8), 897-905.

Kao, C., & Liu, S. T. (2004). Predicting bank performance with financial forecasts: A case of Taiwan commercial banks. *Journal of Banking and Finance, 28*(10), 2353-2368.

Kao, C., & Liu, S. T. (2009). Stochastic data envelopment analysis in measuring the efficiency of Taiwan commercial banks. *European Journal of Operational Research, 196*(1), 312-322.

Liang, L., Wu, J., Cook, W. D., & Zhu, J. (2008). The DEA game cross-efficiency model and its Nash equilibrium. *Operations Research, 56*(5), 1278-1288.

Liang, L., Cook, W. D., & Zhu, J. (2008). DEA models for two-stage processes: Game approach and efficiency decomposition. *Naval Research Logistics, 55*(7), 643-653.

Mostafaee, A., & Saljooghi, F. H. (2010). Cost efficiency measures in data envelopment analysis with data uncertainty. *European Journal of Operational Research, 202*(2), 595-603.

Park, K. S. (2007). Efficiency bounds and efficiency classifications in DEA with imprecise data. *Journal of the Operational Research Society, 58*(4), 533-540.

Park, K. S. (2010). Duality, efficiency computations and interpretations in imprecise DEA. *European Journal of Operational Research, 200*(1), 289-296.

Seiford, L. M. (1997). A bibliography for data envelopment analysis (1978-1996). *Annals of Operations Research, 73*(1), 393-438.

Simaan, M., & Cruz, J. B. (1973). On the Stackelberg strategy in nonzero-sum games. *Journal of Optimization Theory and Applications, 11*(5), 533-555.

Smirlis, Y. G., Maragos, E. K., & Despotis, D. K. (2006). Data envelopment analysis with missing values: An interval DEA approach. *Applied Mathematics and Computation, 177*(1), 1-10.

Vicente, L. N., & Calamai, P. H. (1994). Bilevel and multilevel programming: A bibliography review. *Journal of Global Optimization, 5*(3), 291-306.

Zhu, J. (2003). Imprecise data envelopment analysis (IDEA): A review and improvement with an application. *European Journal of Operational Research, 144*(3), 513-529.

Zhu, J. (2004). Imprecise DEA via standard linear DEA models with a revisit to a Korean mobile telecommunication company. *Operations Research, 52*(2), 323-329.

Authors Note

Chiang Kao, Department of Industrial and Information Management, National Cheng Kung University, Tainan 701, Taiwan, Republic of China.

Shiang-Tai Liu, Graduate School of Business and Management, Vanung University, Chung-Li 320, Taiwan, Republic of China.

Correspondence concerning this article should be addressed to Chiang Kao, Email: ckao@mail.ncku.edu.tw