



Nested Cost Efficiency Intervals in the Presence of Interval Data

KEYWORDS

Data Envelopment Analysis, Interval Data, Cost efficiency

M. Venkata Subba Reddy

Department Statistics, S.V. Degree College, Kadapa

C. Subba Rami Reddy

Department Statistics, S.V. University, Tirupati

ABSTRACT *Data Envelopment Analysis is linear programming based procedure to assess efficiency of decision making units. It needs to specify input and output values. If data uncertainty prevails, where inputs and outputs are assumed to lie in intervals, then efficiencies also belong to intervals whose bounds are deduced solving suitably formulated linear programming problems. In the presence of interval data this paper formulates two pairs of cost efficiency problems under weak and strong optimistic and pessimistic view points. The cost efficiency intervals are shown nested and a numerical problem is solved to verify the same.*

INTRODUCTION:

The boundary of the production possibility set plays an important role in efficiency measurement, the production possibility set is constructed by the input and output vectors of observed firms under a set of assumptions. A very widely used production frontier in empirical research is the boundary of Data Envelopment Analysis (DEA) technology set built on the axioms of inclusion, free disposability, convexity and minimum extrapolation (CHARNES et.al, 1978; BANKER et.al, 1984). To assess input and / or output losses, the input and output vector of inefficient production unit are to be projected onto the boundary of technology set. Such a projection point provides a reference firm to the interior firm. The coordinates of the reference point provides targets to the inefficient decision making unit. The frontier targets are identified by implementing distance functions. Choice of distance function is an important issue both to the production manager and Policy maker (FREI and HARKER, 1999; BAEK and LEE, 2009). Reaching frontier by an inefficient firm is not as simple as solving a programming problem. The interior firm shall strive hard to improve managerial skills, human resource ability, input mix and / or output mix.

EX ANTE AND EX POST PRODUCTION – CHIOCE OF DISTANCE FUNCTION:

In ex ante production input substitution or output transformation is possible, for choice of technique the firm management pursue movements along input or output isoquant. With a knowledge of input or output prices cost minimizing or revenue maximizing or profit maximization bench marks can be located on the frontier of technology set. These targets refer to long run, which cannot be reached implementing oriented distance functions. The distance functions provide not only the distance between inefficient production plan and frontier bench marks, but also the efficient targets.

In ex post production neither factor substitution nor output transformation is possible. The targets assigned by the distance functions refer to short run where fixed inputs cannot be varied, but variable inputs can be contracted. Policy maker is interested in ex ante production possibilities, but production manager concentrates on ex post production possibilities.

outputs are varied along a ray that assumes same input mix / output mix through out the movement, the technique remains to be the same. This observation suggests radial distance functions can be used in ex post comparisons, in particular in very short run.

If input prices are known in orientation approach one finds the coordinates of frontier point at which factor cost is minimized. The success of the search suggests the policy maker and the entrepreneur of the firm to look for input substitution which is possible by change of technique.

INTERVAL DATA – DATA ENVELOPMENT ANALYSIS:

A policy maker or a firm manager usually makes his decision in a state of indeterminacy, based on imprecise information or data. The basic DEA models and their modifications assume that the inputs and outputs are measured by exact values on a ratio scale. Imprecise data refers to input and output data whose true values belong to bounded intervals. If the DEA inputs and outputs are assumed to lie in intervals whose upper and lower bounds are known, then efficiencies also belong to intervals whose bounds can be deduced solving suitably formulated linear programming problems. For assessing interval efficiency Desposits et.al (2002) developed a linear programming problem.

$$\left. \begin{aligned}
 \text{Max } H_0^U &= \sum_{r=1}^s \mu_r y_{r0}^U \\
 \text{s.t } \sum_{i=1}^m v_i x_{i0}^L &= 1 \\
 \sum_{r=1}^s \mu_r y_{r0}^U - \sum_{i=1}^m v_i x_{i0}^L &\leq 0 \\
 \sum_{r=1}^s \mu_r y_{rj}^L - \sum_{i=1}^m v_i x_{ij}^U &\leq 0, \quad j \in N, j \neq 0 \\
 \mu_r, v_i &\geq \epsilon, \quad \forall i, r
 \end{aligned} \right\} \dots(1)$$

In oriented distance function estimation since inputs or

$$\left. \begin{aligned}
 &Max H_0^L = \sum_{r=1}^s \mu_r y_{r0}^L \\
 &s.t \sum_{i=1}^m v_i x_{i0}^U = 1 \\
 &\sum_{r=1}^s \mu_r y_{r0}^L - \sum_{i=1}^m v_i x_{i0}^U \leq 0 \\
 &\sum_{r=1}^s \mu_r y_{rj}^U - \sum_{i=1}^m v_i x_{ij}^L \leq 0, j \in N, j \neq 0 \\
 &\mu_r, v_i \geq \epsilon, \forall i, r
 \end{aligned} \right\} \dots\dots(2)$$

where L and U respectively refer to lower and upper bound of an interval

Problems (1) and (2) provide the input projections falling on optimistic and pessimistic frontiers.

Wang et.al (2005) formulated DEA multiplier form models, implementing interval arithmetic to assess interval efficiency for inefficient decision making units:

$$\left. \begin{aligned}
 &\theta_0^U = Max \theta = \sum_{r=1}^s \mu_r y_{r0}^U \\
 &s.t \sum_{i=1}^m v_i x_{i0}^L = 1 \\
 &\sum_{r=1}^s \mu_r y_{rj}^U - \sum_{i=1}^m v_i x_{ij}^L \leq 0, j \in N \\
 &\mu_r, v_i \geq \epsilon, \forall i,
 \end{aligned} \right\} \dots\dots(3)$$

$$\left. \begin{aligned}
 &\theta_0^L = Max \theta = \sum_{r=1}^s \mu_r y_{r0}^L \\
 &s.t \sum_{i=1}^m v_i x_{i0}^U = 1 \\
 &\sum_{r=1}^s \mu_r y_{rj}^U - \sum_{i=1}^m v_i x_{ij}^L \leq 1, j \in N
 \end{aligned} \right\} \dots\dots(4)$$

Models (3) and (4) differ from models (1) and (2). Consequently, the efficiency bounds yielded by them also differ.

For both pairs of models it can be shown that, $\theta_0^L \leq H_0^L$ and

$$\theta_0^U \leq H_0^U$$

. These bounds do not form nested intervals.

In the present study the factor minimal cost function (LOT-Fl et.al, 2007; MOSTAFAEE and SALJOOGHI, 2010) is confronted with interval data under the hypothesis of weak optimistic and pessimistic view, and strong optimistic and pessimistic view, new factor minimal cost functions are proposed and the cost efficiency intervals are derived that are shown nested.

Factor minimal cost can be evaluated solving the following linear programming problem:

Cost efficiency is the ratio of factor minimal cost to observed cost

$$\left. \begin{aligned}
 &C(y_0, p) = Min_{x, \lambda} \sum_{i=1}^m p_i x_i \\
 &s.t \sum_{j=1}^n \lambda_j x_{ij} \leq x_i, i = 1, 2, \dots, m \\
 &\sum_{j=1}^n \lambda_j y_{rj} \geq y_{r0}, r = 1, 2, \dots, s \\
 &\lambda_j \geq 0, j = 1, 2, \dots, n
 \end{aligned} \right\} \dots\dots(5)$$

$$\text{Cost efficiency} = \frac{C(u, p)}{\sum_{i=1}^m p_i x_{i0}}$$

$$0 \leq \frac{C(u, p)}{\sum_{i=1}^m p_i x_{i0}} \leq 1.$$

The cost efficiency problem can alternatively be expressed as follows:

$$\left. \begin{aligned}
 &Min \sum_{j=1}^n a_j \lambda_j \\
 &s.t \sum_{j=1}^n \lambda_j y_{rj} \geq y_{r0}, r = 1, 2, \dots, s \\
 &\lambda_j \geq 0, j = 1, 2, \dots, n
 \end{aligned} \right\} \dots\dots(6)$$

Proof: consider,

$$\frac{\sum_{i=1}^m p_{i0} x_i}{\sum_{i=1}^m p_{i0} x_{i0}} = \frac{\sum_{i=1}^m p_{i0} \left(\sum_{j=1}^n \lambda_j x_{ij} \right)}{\sum_{i=1}^m p_{i0} x_{i0}}$$

$$= \sum_{j=1}^n \left(\frac{\sum_{i=1}^m p_{i0} x_{ij}}{\sum_{i=1}^m p_{i0} x_{i0}} \right) \lambda_j$$

$$= \sum_{j=1}^n a_j \lambda_j$$

$$Min_{x, \lambda} \frac{\sum_{i=1}^m p_{i0} x_i}{\sum_{i=1}^m p_{i0} x_{i0}} = Min_{\lambda} \sum_{j=1}^n a_j \lambda_j$$

Weak optimistic and pessimistic view point: Under weak pessimistic view point a decision making unit under evaluation considers itself performing worst but best while it is placed in the reference technology. Under weak optimistic view point the targeted decision making unit considers itself performing best but worst while its inputs and outputs are augmented to the reference set.

Strong optimistic and pessimistic view point: Under strong optimistic view point the DMU in evaluation rates itself performing best and the same is assumed while it is placed in reference technology also. Under the pessimistic view point the DMU under evaluation rates itself performing worst, and the same is assumed while its inputs and outputs are augmented to reference technology.

Under weak optimistic view point, we propose the following linear programming problems:

$$CE_{\bar{u}} = \text{Min}_{\lambda} \sum_{j=1}^n \bar{a}_j \lambda$$

$$\sum_{j=1}^n \lambda_j y_{rj}^L \geq y_{r0}^U, r \in S$$

subject to (7)

$$\lambda_j \geq 0$$

$$\bar{a}_j = \frac{\sum_{i=1}^m p_{i0} x_{ij}^U}{\sum_{i=1}^m p_{i0} x_{i0}^L} \geq 0, j \in N$$

where

$$CE_L = \text{Min}_{\lambda} \sum_{j=1}^n \bar{b}_j \lambda_j$$

$$\sum_{j=1}^n \lambda_j y_{rj}^U \geq y_{r0}^L, r \in S$$

subject to (8)

≥

$$\bar{b}_j = \frac{\sum_{i=1}^m p_{i0} x_{ij}^L}{\sum_{i=1}^m p_{ij} x_{i0}^U}$$

where , j = 1, 2, , n

It can be shown that

$$CE_L \leq CE_U$$

$$\sum_{j=1}^n \lambda_j y_{rj}^U \geq \sum_{j=1}^n \lambda_j y_{rj}^L \geq y_{r0}^U \geq y_{r0}^L$$

Proof:

$$\sum_{j=1}^n \lambda_j y_{rj}^L \geq y_{r0}^U \Rightarrow \sum_{j=1}^n \lambda_j y_{rj}^U \geq y_{r0}^L$$

Every feasible solution of (7) is a feasible solution of (8).
Optimal solution of (7) is a feasible solution of (8),

Let $\hat{\lambda}_j, j = 1, 2, \dots, n$ be optimal solution of (7)

$$\bar{a}_j = \frac{\sum_{i=1}^m p_{i0} x_{ij}^U}{\sum_{i=1}^m p_{i0} x_{i0}^L}$$

$$\bar{b}_j = \frac{\sum_{i=1}^m p_{i0} x_{ij}^L}{\sum_{i=1}^m p_{ij} x_{i0}^U}$$

$$\bar{b}_j \leq \bar{a}_j, \forall j$$

$$\sum_{j=1}^n \hat{\lambda}_j \bar{b}_j \leq \sum_{j=1}^n \hat{\lambda}_j \bar{a}_j$$

$$\sum_{j=1}^n \hat{\lambda}_j \bar{b}_j \leq CE_U$$

$$\text{Min}_{\lambda} \sum_{j=1}^n \lambda_j \bar{b}_j \leq \sum_{j=1}^n \hat{\lambda}_j \bar{b}_j \leq CE_U$$

$$CE_L \leq CE_U$$

Under strong pessimistic and optimistic view points we formulate the following linear programming problems:

$$\overline{CE}_U = \text{Min}_{\lambda} \sum_{j=1}^n \overline{\bar{a}}_j \lambda_j$$

$$\sum_{j=1}^n \lambda_j y_{rj}^L + \lambda_0 y_{r0}^U \geq y_{r0}^U$$

..... (9)

$$\bar{a}_j = \frac{\sum_{i=1}^m p_{i0} x_{ij}^U}{\sum_{i=1}^m p_{i0} x_{i0}^L}, \quad j \neq 0$$

where

$$\bar{a}_0 = \frac{\sum_{i=1}^m p_{i0} x_{ij}^L}{\sum_{i=1}^m p_{i0} x_{i0}^L}, \quad j = 0$$

$$\overline{CE}_L = \text{Min}_{\lambda} \sum_{j=1}^n \bar{b}_j \lambda_j$$

$$\sum_{j \neq 0} \lambda_j y_{rj}^U + \lambda_0 y_{r0}^L \geq y_{r0}^L$$

subject to (10)

$$\lambda_j \geq 0$$

It can be shown that $\overline{CE}_u \leq CE_u$

$$\sum_{j \neq 0} \lambda_j y_{rj}^L + \lambda_0 y_{r0}^U \geq \sum_{j=1}^n \lambda_j y_{rj}^L \geq y_{r0}^U$$

Proof:

$$\sum_{j=1}^n \lambda_j y_{rj}^L \geq y_{r0}^U \Rightarrow \sum_{j \neq 0} \lambda_j y_{rj}^L + \lambda_0 y_{r0} \geq y_{r0}$$

Every feasible solution of (7) is a feasible solution of (9).
Optimal solution (7) is a feasible solution of (9).

Let $\hat{\lambda}_j$ be optimal solution of (7)

$$\bar{a}_j = \frac{\sum_{i=1}^m p_{i0} x_{ij}^U}{\sum_{i=1}^m p_{i0} x_{i0}^L}, \quad j \neq 0$$

$$\bar{a}_0 = \frac{\sum_{i=1}^m p_{i0} x_{i0}^L}{\sum_{i=1}^m p_{i0} x_{i0}^L} = 1$$

, j = 0

$$\bar{a}_j = \frac{\sum_{i=1}^m p_{i0} x_{ij}^U}{\sum_{i=1}^m p_{i0} x_{i0}^L}, \quad j \neq 0$$

$$\bar{a}_0 = \frac{\sum_{i=1}^m p_{i0} x_{i0}^L}{\sum_{i=1}^m p_{ij} x_{i0}^U}, \quad j = 0$$

$$\bar{a}_j \leq \bar{a}_j$$

, $\forall j$

$$\sum_{j=1}^n \bar{a}_j \hat{\lambda}_j \leq \sum_{j=1}^n \bar{a}_j \hat{\lambda}_j$$

$$\sum_{j=1}^n \bar{a}_j \hat{\lambda}_j \leq CE_u$$

$$\text{Min}_{\lambda} \sum_{j=1}^n \bar{a}_j \lambda_j \leq \sum_{j=1}^n \bar{a}_j \hat{\lambda}_j \leq CE_u$$

$$\overline{CE}_u \leq CE_u$$

It can be shown that $CE_L \leq \overline{CE}_L$

$$\sum_{j \neq 0} \lambda_j y_{rj}^U \geq \sum_{j \neq 0} \lambda_j y_{rj}^U + \lambda_0 y_{r0}^L \geq y_{r0}^L$$

Proof:

$$\sum_{j \neq 0} \lambda_j y_{rj}^U + \lambda_0 y_{r0}^L \geq y_{r0}^L \Rightarrow \sum_{j=0}^n \lambda_j y_{rj}^U \geq y_{r0}^L$$

Every feasible solution of (10) is feasible solution of (8).

Optimal solution of (10) is feasible solution of (8).

Let $\hat{\lambda}_j$ be optimal for (8)

$$\bar{b}_j = \frac{\sum_{i=1}^m p_{i0}x_{ij}^L}{\sum_{i=1}^m p_{i0}x_{ij}^U}, j \neq 0$$

where

$$\bar{b}_0 = \frac{\sum_{i=1}^m p_{i0}x_{i0}^U}{\sum_{i=1}^m p_{i0}x_{i0}^L} = 1, j = 0$$

$$\bar{b}_j = \frac{\sum_{i=1}^m p_{i0}x_{ij}^L}{\sum_{i=1}^m p_{i0}x_{ij}^U}, j \neq 0$$

$$\bar{b}_0 = \frac{\sum_{i=1}^m p_{i0}x_{i0}^U}{\sum_{i=1}^m p_{i0}x_{i0}^L}, j = 0$$

$$\sum_{j=1}^n \bar{b}_j \hat{\lambda}_j \leq \sum_{j=1}^n \bar{b}_j \hat{\lambda}_j = \overline{CE_L}$$

$$\sum_{j=1}^n \bar{b}_j \hat{\lambda}_j \leq \overline{CE_L}$$

$$\text{Min}_{\lambda} \sum_{j=1}^n \bar{b}_j \lambda_j \leq \sum_{j=1}^n \bar{b}_j \hat{\lambda}_j \leq \overline{CE_L}$$

$$CE_L \leq \overline{CE_L}$$

It can be shown that $\overline{CE_L} \leq \overline{CE_U}$

Proof :Consider the linear programming problems (9) and (10).

$$\sum_{j \neq 0}^n \lambda_j y_{rj}^L + \lambda_0 y_{r0}^U \geq y_{r0}^U$$

$$\sum_{j \neq 0} \lambda_j \left(\frac{y_{rj}^L}{y_{rj}^U} \right) + \lambda_0 \geq 1$$

$$\sum_{j \neq 0} \lambda_j y_{rj}^U + \lambda_0 y_{r0}^L \geq y_{r0}^L$$

$$\sum_{j \neq 0} \lambda_j \left(\frac{y_{rj}^U}{y_{r0}^L} \right) + \lambda_0 \geq 1$$

$$\sum_{j \neq 0} \lambda_j \left(\frac{y_{rj}^U}{y_{r0}^L} \right) + \lambda_0 \geq \sum_{j \neq 0} \lambda_j \left(\frac{y_{rj}^U}{y_{rj}^L} \right) + \lambda_0 \geq 1$$

$$\sum_{j \neq 0}^n \lambda_j y_{rj}^L + \lambda_0 y_{r0}^U \geq y_{r0}^U \Rightarrow \sum_{j \neq 0} \lambda_j y_{rj}^U + \lambda_0 y_{r0}^L \geq y_{r0}^L$$

Every feasible solution of (9) is a feasible solution of (10).
Optimal solution of (9) is a feasible solution of (10).

Let $\hat{\lambda}_j$ be optimal for (9)

$$\bar{a}_j = \frac{\sum_{i=1}^m p_{i0}x_{ij}^U}{\sum_{i=1}^m p_{i0}x_{i0}^L}, j \neq 0$$

$$\bar{a}_0 = \frac{\sum_{i=1}^m p_{i0}x_{ij}^L}{\sum_{i=1}^m p_{i0}x_{i0}^L} = 1, j \neq 0$$

$$\bar{b}_j = \frac{\sum_{i=1}^m p_{i0}x_{ij}^L}{\sum_{i=1}^m p_{i0}x_{i0}^U}, j \neq 0$$

$$\bar{b}_0 = \frac{\sum_{i=1}^m p_{i0}x_{ij}^U}{\sum_{i=1}^m p_{i0}x_{i0}^U} = 1, j \neq 0$$

$$\bar{b}_j \leq \bar{a}_j, \forall j$$

$$\sum_{j=1}^n \bar{b}_j \hat{\lambda}_j \leq \sum_{j=1}^n \bar{a}_j \hat{\lambda}_j$$

$$\sum_{j=1}^n \bar{b}_j \hat{\lambda}_j \leq \overline{CE_u}$$

where $\hat{\lambda}_j$ optimal solution of $\overline{CE_u}$

$$\text{Min}_j \sum_{j=1}^n \bar{b}_j \lambda_j \leq \sum_{j=1}^n \bar{b}_j \hat{\lambda}_j \leq \overline{CE_u}$$

$$\overline{CE_L} \leq \overline{CE_U}$$

Combining the results of theorems (2), (3), (4), and (5) we obtain,

$$CE_L \leq \overline{CE_L} \leq CE \leq \overline{CE_u} \leq CE_u$$

From the above inequality it follows that, weak optimistic and pessimistic view points provide larger efficiency interval than strong optimistic and pessimistic view points.

EMPIRICAL INVESTIGATION :

For numerical verification of the above inequality the data derived are from Annual Survey of Industries (ASI, 2005-2006, 2008-2009). The Value Added by Fixed Capital and Work Force is treated as output. Fixed Capital and Work Force are inputs. The data refer to two discrete time points, 2005-2006 and 2008-2009.

DMU	CE _L	CE _L Bar	CE _U Bar	CE _U
Andhra Pradesh	0.1859	0.1859	0.7427	0.7425
Chattishgarh	0.2182	0.2182	1	1
Gujarat	0.2521	0.2521	0.5169	0.5169
Haryana	0.2894	0.2894	0.9537	0.9537
Karnataka	0.2053	0.2053	0.8942	0.8941
Madhya Pradesh	0.1527	0.1527	0.7446	0.7446
Maharashtra	0.6694	1	1	1
Orissa	0.1012	0.1012	0.6984	0.6984
Punjab	0.1836	0.1836	0.695	0.695
Rajasthan	0.1948	0.1948	0.9327	0.9327
Tamil Nadu	0.2181	0.2181	0.4393	0.5578
Uttar Pradesh	0.2179	0.2179	0.4546	0.5712
Westbengal	0.1547	0.1547	0.4302	0.5469

The computational values satisfy the theoretical inequalities,

$$CE_L \leq \overline{CE_L} \leq \overline{CE_u} \leq CE_u$$

ties,

SUMMARY AND CONCLUSIONS:

In the presence of data uncertainty, in particular if lower and upper bounds are specified for inputs and outputs cost efficiencies are realized in interval form. In this study under weak and strong optimistic and pessimistic view points the cost efficiency intervals are shown nested. The inequalities are verified for a numerical example referring to data obtained from the Annual Survey of Industries (ASI) bulletins. Value Added is treated as output, Fixed Capital and number of Employees as inputs. Total wages and salaries are divided by number of employees to obtain wage rate, and Value Added minus total wages and salaries are divided by Fixed Capital to arrive at the price of Fixed Capital. Cost efficiency problems are formulated under weak and strong pessimistic and optimistic view points, the resultant cost efficiency inequalities are shown nested. The nested property is verified for the above live problem covering the two discrete time points 2005-2006 and 2008-2009, for the total manufacturing sectors of 13 Indian Major States.

REFERENCE

[1]R. Banker, A. Charnes, W.W. Cooper, E. Rhodes (1984), 'Some models for estimating technical and scale inefficiencies in data envelopment analysis', European Journal of Operations Research, 30(9), 1078-1092. | [2]A. Charnes, W.W. Cooper, E. Rhodes (1978), 'Measuring the efficiency of decision making units', European Journal of Operations Research, 2(6), 422-444. | [3]Chulwoo Baek and Jeong-Dong Lee (2009), 'The Relevance of DEA bench marking information and the Least-Distance Measure', Mathematical and Computer Modeling, 49, 265-275. | [4]D.K. Desopitsis, Y.G. Smirlis (2002), 'Data envelopment analysis with imprecise data', European Journal of Operations Research, 140, 24-36. | [5]Fare, R. Lovell, C.A.K (1978), 'Measuring the technical efficiency of production', Journal of Economic theory, 19(1), 150-162. | [6]Frances X. Frei and Patric T. Harker (1999), 'Projections Onto Efficient Frontiers: Theoretical and Computational Extensions to DEA', Journal of Productivity Analysis, 11, 275-300. | [7]Govt. of INDIA (2005-2006, 2008-2009), Annual Survey of Industries. | [8] F.Hosseinzadh Lotfi, G.R. Jahanshahloo, R. Shahverdi, M. Rostamy-Malkhalifeh (2007), 'Cost Efficiency and Cost Malmquist Productivity Index with Interval Data', International Mathematical Forum, 2, 441-453. | [9]A. Mostafaeae and F.H. Saljooghi (2010), 'Cost Efficiency Measures in Data Envelopment Analysis with Data Uncertainty', European Journal of Operations Research, 202, 595-603 | [10]Y.M. Wang, R. Greatbanks, J.B. Yang (2005), 'Interval Efficiency assessment using data envelopment analysis', Fuzzy Sets and Systems 153, 347-370. |