Available online at http://ijim.srbiau.ac.ir

Int. J. Industrial Mathematics Vol. 1, No. 3 (2009) 219-226

Ranking Efficient DMUs with Stochastic Data by Considering Inefficient Frontier

M. H. Behzadi[a]*, N. Nematollahi [b] and M. Mirbolouki[c]

(a) *Department of Statistics, Science and Research Branch, Islamic Azad University, Tehran, Iran.*

(b) *Department of Statistics, Faculty of Economics, Allameh Tabataba'i University, Tehran, Iran.*

(c) *Department of Mathematics, Science and Research Branch, Islamic Azad University, Tehran, Iran.*

Abstract

Data Envelopment Analysis (DEA) models which evaluate the efficiency of a set of decision making units (DMUs) are unable to discriminate between efficient DMUs. The problem of discriminating between these efficient DMUs is an interesting subject. A large number of methods for fully ranking both efficient and inefficient DMUs have been proposed. Through real world applications, analysis may encounter data that are not deterministic or on have a stochastic essence but whose distribution can be defined by collecting data in successive periods and by statistical methods. In this paper, a method for ranking stochastic efficient DMUs is suggested which is based on the full inefficient frontier method. Using a numerical example, we will demonstrate how to use the result.

Keywords: Data envelopment analysis, Quadratic programming; Ranking, Standard normal distribution.

1 Introduction

Data Envelopment Analysis (DEA) was originated by Charnes et al. [2] and then it extended to an approach for evaluating the relative efficiency of DMUs. In real application, we know that usually plural DMUs are efficient. The problem of discriminating between these efficient DMUs is an interesting subject [5]. Sexton et al. [7] were pioneers in the ranking field. They introduced a ranking method based on cross-efficiency. Then, the ranking of DEA-efficient DMUs based on benchmarking, was an approach initially developed by Torgersen et al. [10]. In this method, a DMU is highly ranked if it is chosen as a reference by many other inefficient DMUs. The most popular research stream in ranking DMUs is called super-efficiency. This stream was first developed by Andersen

*Corresponding author.E-mail address: behzadi@srbiau.ac.ir

and Petersen [1]. Thrall [8] pointed out that the model developed by Andersen and Petersen may result in instability when some inputs are close to zero. Then, to avoid this problem, MAJ [6] and SBM [9] models were proposed. All of these methods rank DMUs by comparing DMUs with the efficient frontier. One of the disadvantages of DEA is assessing DMUs in the best conditions. Jahanshahloo and Afzalinejad [4] have introduced the full inefficient frontier and they proposed a method for assessing DMUs in the worst conditions and a ranking method by using this factor. All of the proposed ranking models consider different data such as: deterministic, interval, fuzzy, etc.data In different real world applications, analysis may encounter stochastic data. In this paper, on the basis of Cooper's method [3], the stochastic efficient DMUs have been distinguished and, with the contribution of full inefficient stochastic frontier, a model for ranking DMUs with the stochastic data has been presented. The paper is organized as follows: First, the inefficient frontier and then the stochastic DEA models are introduced. After that, ranking DMUs with the stochastic inefficient frontier is discussed. Using a numerical example, we will demonstrate how to use the result.

2 Inefficient frontier

Jahanshahloo and Afzalinejad [4] have defined the full inefficient frontier. By the contribution of the full inefficient frontier, they have also proposed a model for identifying the worst score of efficiency.

DMU_j is full inefficient if it can not be dominated by other dummy DMUs. That is, DMU_j is full inefficient if it belongs to $F(S)$ which is defined as follows:

$$F(S) = \left\{ (x,y)|\ \forall (x',y') \in R^{m+s}((-x',y') \underset{\neq}{\leq} (-x,y) \Rightarrow (x',y') \notin S) \right\} \subseteq S,$$

where S is the convex hull of observed DMUs. Thus the full inefficient frontier in radial input orientation is defined as $F_I(S)$ where

$$F_I(S) = \{(x,y)|(x,y) \in S \quad \& \quad \forall \psi(\psi > 1 \Rightarrow (\psi x, y) \notin S)\}.$$

Thus, DMU_o is located on the full inefficient frontier with variable returns to scale, if in the following model we have $\psi_o^* = 1$, and it is not located on the full inefficient frontier if $\psi_o^* > 1$.

$$\begin{aligned} \psi_o^* = \max \quad & \psi \\ s.t. \quad & \sum_{j=1}^{n} \lambda_j x_{ij} \geq \psi x_{io}, \quad i = 1, ..., m, \\ & \sum_{j=1}^{n} \lambda_j y_{rj} \leq y_{ro}, \quad r = 1, ..., s, \\ & \sum_{j=1}^{n} \lambda_j = 1, \\ & \lambda_j \geq 0, \quad j = 1, ..., n. \end{aligned} \tag{2.1}$$

ψ_o^* indicates the distance of DMU_o from the inefficient frontier. Therefore, the greater ψ_o^* the better the ranking score. In Fig. 1, the convex hull of DMUs is schematically portrayed. The piecewise linear frontier AGFE is the efficient frontier and ABCD is the inefficient frontier.

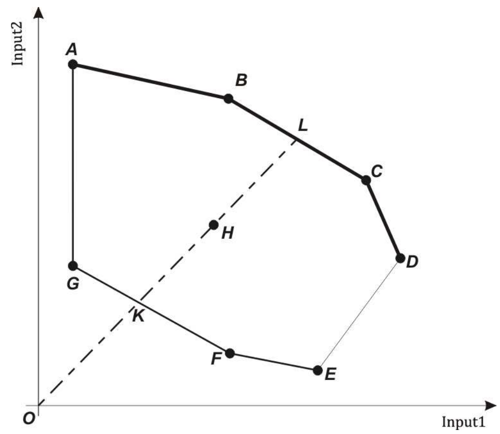

Fig. 1. Convex hull of DMUs with two input and one equal output.

3 Stochastic DEA models

Consider n DMUs with $\tilde{X}_j = (\tilde{x}_{1j},...,\tilde{x}_{mj})$ and $\tilde{Y}_j = (\tilde{y}_{1j},...,\tilde{y}_{sj})$ as random input and output vectors of DMU_j, $j = 1\ldots,n$, respectively. Assume that $X_j = (x_{1j},...,x_{mj})$ and $Y_j = (y_{1j},...,y_{sj})$ stand for the corresponding vectors of expected values of input and output for every DMU_j. All input and output components have been considered to be normally distributed. The chance constrained version of the input-oriented stochastic BCC model is as follows:

$$\begin{aligned}
\min \quad & \theta \\
s.t. \quad & p\{\sum_{j=1}^{n} \tilde{y}_{rj}\lambda_j \geq \tilde{y}_{ro}\} \geq 1-\alpha, \quad r = 1,...,s, \\
& p\{\sum_{j=1}^{n} \tilde{x}_{ij}\lambda_j \leq \theta\tilde{x}_{io}\} \geq 1-\alpha, \quad i = 1,...,m, \\
& \sum_{j=1}^{n} \lambda_j = 1, \\
& \lambda_j \geq 0, \qquad j = 1,...,n.
\end{aligned} \tag{3.2}$$

Model (3.2) can be converted into the following two-stage model with equality constraints:

$$\begin{aligned}
\min \quad & \theta - \varepsilon(\sum_{r=1}^{s} s_r^+ + \sum_{i=1}^{m} s_i^-) \\
s.t. \quad & p\{\sum_{j=1}^{n} \tilde{y}_{rj}\lambda_j - s_r^+ \geq \tilde{y}_{ro}\} = 1-\alpha, \quad r = 1,...,s, \\
& p\{\sum_{j=1}^{n} \tilde{x}_{ij}\lambda_j + s_i^- \leq \theta\tilde{x}_{io}\} = 1-\alpha, \quad i = 1,...,m, \\
& \sum_{j=1}^{n} \lambda_j = 1, \\
& s_i^- \geq 0, \quad s_r^+ \geq 0, \quad i = 1,...,m, \quad r = 1,...,s. \\
& \lambda_j \geq 0, \quad j = 1,...,n,
\end{aligned} \tag{3.3}$$

where, p denotes "probability" and α is a predetermined number between 0 and 1. On the basis of normal distribution characteristics, the deterministic model for (3.3) can be attained as follows:

$$\begin{aligned} \min \quad & \theta - \varepsilon(\sum_{r=1}^{s} s_r^+ + \sum_{i=1}^{m} s_i^-) \\ s.t. \quad & \sum_{j=1}^{n} y_{rj}\lambda_j - s_r^+ + \Phi^{-1}(\alpha)\sigma_r^o(\lambda) = y_{ro}, \quad r = 1, ..., s, \\ & \sum_{j=1}^{n} x_{ij}\lambda_j + s_i^- - \Phi^{-1}(\alpha)\sigma_i^I(\lambda, \theta) = \theta x_{io}, \quad i = 1, ..., m, \\ & \sum_{j=1}^{n} \lambda_j = 1, \\ & s_i^- \geq 0, \quad s_r^+ \geq 0, \quad i = 1, ..., m, \quad r = 1, ..., s, \\ & \lambda_j \geq 0, \quad j = 1, ..., n, \end{aligned} \tag{3.4}$$

where

$$(\sigma_r^o(\lambda))^2 = \sum_{j \neq o} \sum_{k \neq o} \lambda_j \lambda_k cov(\tilde{y}_{rj}, \tilde{y}_{rk}) + 2(\lambda_o - 1) \sum_{j \neq o} \lambda_j cov(\tilde{y}_{rj}, \tilde{y}_{ro}) + (\lambda_o - 1)^2 var(\tilde{y}_{ro}),$$

and

$$(\sigma_i^I(\lambda, \theta))^2 = \sum_{j \neq o} \sum_{k \neq o} \lambda_j \lambda_k cov(\tilde{x}_{ij}, \tilde{x}_{ik}) + 2(\lambda_o - \theta) \sum_{j \neq o} \lambda_j cov(\tilde{x}_{ij}, \tilde{x}_{io}) + (\lambda_o - \theta)^2 var(\tilde{x}_{io}).$$

Here, Φ is the cumulative distribution function of the standard normal distribution and $\Phi^{-1}(\alpha)$, is its inverse at the level of α. The above model is a nonlinear programming model which can be converted into a quadratic programming model.

Definition 3.1. *DMU_o is stochastic efficient if and only if in the optimal solution of model (3.4) the following conditions are satisfied:*
(i) $\theta^* = 1$
(ii) Slack values are all zeros.

4 Stochastic inefficient frontier

Consider n DMUs with stochastic data as defined in section 3. The stochastic version of model (2.1) for evaluating the stochastic inefficiency score of DMU_o, $o \in \{1, ..., n\}$, is as follows:

$$
\begin{aligned}
\max \quad & \psi \\
s.t. \quad & p\{\sum_{j=1}^{n} \tilde{x}_{ij}\lambda_j \geq \psi\tilde{x}_{io}\} \geq 1-\alpha, \quad i=1,...,m, \\
& p\{\sum_{j=1}^{n} \tilde{y}_{rj}\lambda_j \leq \tilde{y}_{ro}\} \geq 1-\alpha, \qquad r=1,...,s, \\
& \sum_{j=1}^{n} \lambda_j = 1, \\
& \lambda_j \geq 0, \qquad\qquad j=1,...,n,
\end{aligned} \tag{4.5}
$$

The deterministic form of model (4.5) can be attained as follows:

$$
\begin{aligned}
\widehat{\psi}_o = \max \quad & \psi \\
s.t. \quad & \sum_{j=1}^{n} x_{ij}\lambda_j + \Phi^{-1}(\alpha)\sigma_i^I(\lambda,\psi) \geq \psi x_{io}, \quad i=1,...,m, \\
& \sum_{j=1}^{n} y_{rj}\lambda_j - \Phi^{-1}(\alpha)\sigma_r^o(\lambda) \leq y_{ro}, \qquad r=1,...,s, \\
& \sum_{j=1}^{n} \lambda_j = 1, \\
& \lambda_j \geq 0, \quad j=1,...,n.
\end{aligned} \tag{4.6}
$$

where

$$(\sigma_r^o(\lambda))^2 = \sum_{j\neq o}\sum_{k\neq o} \lambda_j\lambda_k cov(\tilde{y}_{rj},\tilde{y}_{rk}) + 2(\lambda_o-1)\sum_{j\neq o}\lambda_j cov(\tilde{y}_{rj},\tilde{y}_{ro}) + (\lambda_o-1)^2 var(\tilde{y}_{ro}),$$

and

$$(\sigma_i^I(\lambda,\psi))^2 = \sum_{j\neq o}\sum_{k\neq o} \lambda_j\lambda_k cov(\tilde{x}_{ij},\tilde{x}_{ik}) + 2(\lambda_o-\psi)\sum_{j\neq o}\lambda_j cov(\tilde{x}_{ij},\tilde{x}_{io}) + (\lambda_o-\psi)^2 var(\tilde{x}_{io}).$$

The above model can also be converted into a quadratic programming model. To simplify this model, we can assume that outputs and inputs for different DMUs are independent. This assumption implies that $cov(\tilde{x}_{ij},\tilde{x}_{ik}) = 0$ and also $cov(\tilde{y}_{rj},\tilde{y}_{rk}) = 0$ for every $j \neq k$.

Theorem 4.1. *Model (3.3) is feasible for any α level.*

Proof. Let $\theta = 1$, $\lambda_o = 1$, $\lambda_j = 0$ for all $j \neq o$. This is a feasible solution for any α level. □

Theorem 4.2. *Model (4.6) is feasible for any α level.*

Proof. Let $\psi = 1$, $\lambda_o = 1$, $\lambda_j = 0$ for all $j \neq o$. This is a feasible solution for any α level. □

Theorem 4.3. *The stochastic inefficiency score of DMU_o is greater than or equal to 1, $\widehat{\psi}_o \geq 1$.*

Proof. Let $\psi = 1$, $\lambda_o = 1$, $\lambda_j = 0$ for all $j \neq o$. Then $\sigma_i^I(\lambda, \psi) = 0$, $\sigma_r^o(\lambda) = 0$ and all constraints of model (4.6) will be satisfied by this solution. Model (4.6) is a maximization model, so the proof is completed. □

Definition 4.1. *DMU_o is located on the inefficient stochastic frontier if and only if in the optimal solution of model (4.6), $\widehat{\psi}_o = 1$.*

5 An application

In this section, we consider 10 branches of an Iranian bank with two stochastic inputs and two stochastic outputs and run the proposed model in order to fully rank the stochastic efficient units. In this model, “payable benefit” and “delayed requisitions” are inputs and “amount of deposits ” and “received benefit” are outputs. These data, which are obtained from an observation of ten successive months, have normal distribution and their scaled parameters are presented in Table 1. We intend to assess the total performance of these units. In this example, these DMUs have been assessed with two different α levels by model (3.4) and stochastic efficient DMUs have been ranked by their inefficiency score and applying model (4.6). We consider $\alpha = 0.05$, i.e., at least 95% confidence in the results which is shown in Table 2.

Table 1
Predicted inputs and outputs

inputs				outputs			
x_{ij}	$N(\mu,\sigma)$	x_{ij}	$N(\mu,\sigma)$	y_{rj}	$N(\mu,\sigma)$	y_{rj}	$N(\mu,\sigma)$
X1,1	N(18.79,9.41)	X2,1	N(7.28,0.76)	Y1,1	N(49.6,6.93)	Y2,1	N(4.7,0.64)
X1,2	N(44.3,25.3)	X2,2	N(1.11,0.15)	Y1,2	N(73.13,3.62)	Y2,2	N(1.85,0.15)
X1,3	N(19.73,16.63)	X2,3	N(19.2,0.69)	Y1,3	N(108.04,15.02)	Y2,3	N(6.06,0.12)
X1,4	N(17.43,11.06)	X2,4	N(59.47,0.92)	Y1,4	N(44.97,3.71)	Y2,4	N(4.9,1.29)
X1,5	N(10.38,4.59)	X2,5	N(12.23,7.74)	Y1,5	N(31.63,6.24)	Y2,5	N(2.78,0.66)
X1,6	N(16.67,10.42)	X2,6	N(568.63,37.42)	Y1,6	N(71.98,8.37)	Y2,6	N(13.19,3.03)
X1,7	N(25.46,13.67)	X2,7	N(552.85,20.78)	Y1,7	N(78.05,13.99)	Y2,7	N(7.79,1.89)
X1,8	N(123.06,65.3)	X2,8	N(14.78,0.25)	Y1,8	N(219.69,19.38)	Y2,8	N(35.3,3.92)
X1,9	N(36.16,19.59)	X2,9	N(361.88,34.11)	Y1,9	N(86.25,6.95)	Y2,9	N(17.64,1.92)
X1,10	N(46.41,23.06)	X2,10	N(12.81,0.62)	Y1,10	N(194.58,42.15)	Y2,10	N(25.9,3.52)

Table 2
Results

efficient DMU	$\widehat{\psi}$	RANK
DMU3	3.432	2
DMU5	2.091	3
DMU6	7.215	1
DMU10	1.117	4

The results in Table 5.2 show that DMU_6 has the best ranking score because its distance from the inefficient stochastic frontier is more than that of other DMUs, which can be determine by the objective value of model (4.6).

6 Conclusion

In DEA, models have been formulated for evaluating efficiency and ranking DMUs in various fields with different data such as: deterministic, interval, fuzzy, etc.data In real world applications, managers may encounter data which are not deterministic. In such situations, the need to have a model with the ability to rank DMUs has been well recognized. Cooper et al. [3] have proposed a model in which DMUs with stochastic data have been assessed and have thus defined the stochastic efficient DMUs. In the present paper, on the basis of the inefficiency score of efficient DMUs, a model for ranking stochastic efficient DMUs with stochastic data has been presented. In this model, a DMU with greater distance from the inefficient stochastic frontier has a better ranking score. The proposed model is a quadratic programming model and the objective function is a function of α, which is the level of error that should be determined by the managers. It is noteworthy to say that, if α increases, the level of reliability of results will decrease. In addition to the normal distribution, which be used this paper, different other distributions can be considered from this point of view.

References

[1] P. Andersen, N.C. Petersen, A procedure for ranking efficient units in data envelopment analysis, Management Science 39 (10) (1993) 1261-1264.

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

[3] W. W. Cooper, H. Deng, Zhimin Huang, Sus X.Li. Chance constrained programming approaches to congestion in stochastic data envelopment analysis, European Journal of Operational Research, 155 (2004) 487-501.

[4] G. R. Jahanshahloo, M. Afzalinejad, A ranking method based on a full inefficient frontier, Applied Mathematical Modeling. 30 (2006) 248-260.

[5] M. Matric, G. Savic, An application of DEA for comparative analysis and ranking of regions in Serbia with regards to socio-economic development, European Journal of Operational Research 132 (2001) 343-356.

[6] S. Mehrabian, M. R. Alirezaee, G. R. Jahanshahloo, A complete efficiency ranking of decision making units in DEA, Computational Optimization and Applications (COAP) 14 (1999) 261-266.

[7] T. R. Sexton, R. H. Silkman, A.J. Hogan, Data envelopment analysis: critique and extension, in: R.H. Silkman (Ed.), Measuring Efficiency: An Assessment of Data Envelopment Analysis, Jossey-Bass, San Fransisco, CA, (1986) 73-105.

[8] R. M. Thrall, Duality classification and slacks in data envelopment analysis, The Annals of Operations Research 66 (1996) 109-138.

[9] K. Tone, A slack-based measure of super-efficiency in data envelopment analysis, European Journal of Operational Research 143 (2002) 32-41.

[10] A. M. Torgersen, F. R. Forsund, S. A. C. Kittelsen, Slack-adjusted efficiency measures and ranking of efficient units, The Journal of Productivity Analysis 7 (1996) 379-398.