J. Fuzzy. Ext. Appl. Vol. 1, No. 1 (2020) 76–84

J. Fuzzy. Ext. Appl

Journal of Fuzzy Extension & Applications

www.journal-fea.com

Efficiency Study with Undesirable Inputs and Outputs in DEA

Abbasali. Monzeli*[*,1], *Behroz Daneshian*[2], *Ghasem. Tohidi*[2], *Masud. Sanei*[2], *Shabnam Razaveian*[2],

[1]*Department of Mathematics, Islamic Azad University of Central Tehran Branch, Iran.*

[2]*Islamic Azad University of Central Tehran Branch, Iran.*

P A P E R   I N F O

Chronicle:
Received: 09 January 2020
Revised: 20 February 2020
Accepted: 01 March 2020

Keywords:
Data Envelopment Analysis.
Undesirable Inputs and Outputs.
Efficiency.
Efficient Boundaries.

A B S T R A C T

Data Envelopment Analysis (DEA) is one of the well-known methods for calculating efficiency, determining efficient boundaries and evaluating efficiency that is used in specific input and output conditions. Traditional models of DEA do not try to reduce undesirable outputs and increase undesirable inputs. Therefore, in this study, in addition to determining the efficiency of Decision-Making Units (DMU) with the presence of some undesirable input and output components, its effect has also been investigated on the efficiency limit. To do this, we first defined the appropriate production possibility set according to the problem assumptions, and then we presented a new method to determine the unfavorable performance of some input and output components in decision-making units. And we determined the impact of unfavorable inputs and outputs on the efficient boundary. We also showed the model result by providing examples for both unfavorable input and output states and solving them and determining the efficiency score and driving them to the efficient boundary by plotting those boundaries.

1. Introduction

Data envelopment analysis is one of the most popular methods for determining efficiency, and the boundary of efficiency based on the concept of condition of defective units. In 1987, Charles et al. [2] identified efficient boundaries using linear programming and used them to determine productivity. In this way, they used both output-axis and input-axis models. Although these two models are not the only ones used, they are still the most popular DEA model. Many researchers use the DEA method to determine the boundary of performance and evaluate performance [4].

Over the past two decades, DEA has established itself as the strongest and most valuable methodology [3]. In many practical issues, some inputs of decision-making units may be such that increasing these inputs increases efficiency and decreasing it reduces efficiency, such as waste recycling operations, scrap metal and glass, etc., where it is necessary to increase the undesirable inputs to improve the level of efficiency, or some of the outputs of decision-making units may be such that increasing these outputs reduces efficiency and decreases it increases efficiency. Consider the

[*] Corresponding author
E-mail address: abbas_monzeli@yahoo.com
DOI: 10.22105/jfea.2020.248018.1005

waste of a factory or the deaths of patients in hospitals and the dismissals of doctors and nurses in training centers, which should be reduced as an undesirable output to increase efficiency. Undesirable outputs are generally desirable products and therefore can only be reduced by reducing them. There are methods for importing undesirable outputs into the DEA that can be divided into two categories:

- Direct methods.
- Indirect methods.

In indirect methods, undesirable inputs and outputs in each unit are converted into desired inputs and outputs by a uniform descending function and then the performance of the units is evaluated using DEA standard models. Direct ones are methods that use assumptions in production possibility set so that they are used in evaluating desirable input and output.

Conventional data envelopment analysis models in most studies treated the units under evaluation as a black box, producing only a series of primary inputs and using them to produce a series of final outputs; However, with the research of the last two decades, they came to the conclusion that the obtained efficiency is not accurate without considering their internal structure, and there are ambiguities in the analysis of its efficiency, and the role of undesirable factors should also be examined.

In the real world, it is not possible to match all inputs and outputs of inefficient units based on DEA results, which Chiang Kao showed with a new model design, which is possible. Unfavorable outputs are generally desirable products and can therefore be reduced only by a concomitant reduction in the second product. To understand this concept, the price of an undesirable output shadow must be negative and the opposite for a positive output. Based on these conditions, Kao et al. [6] in a paper presented a data envelopment analysis model that allows the production units under evaluation to determine the shadow price for both favorable and unfavorable outputs to maximize the measured performance score. The proposed model satisfies the assumption of poor usability of outputs. It is also shown that there is a directional function model in a group that has been widely used in modeling adverse outputs. However, unlike conventional directional distance measures, the proposed model is able to provide performance in the range of zero and one for easy comparison between inefficiently produced units.

Cross-productivity evaluation methods have long been proposed as an option for ranking decision-making units in data envelopment analysis. Neutral reciprocity performance evaluation methods are developed in a way that is only self-interested and indifferent to other DMUs. Accordingly, in 2019, Shi et al. [19] introduced a new cross-performance evaluation method in which each DMU has a neutral attitude towards its other peer units. This is done by introducing an Ideal Virtual Border (IVF) and a Non-Ideal Virtual Border (AVF). Unlike cross-performance evaluation methods, this cross-performance evaluation method determines the set of input and output weights for each DMU. The most important operation in this study is to introduce an ideal virtual boundary and a non-ideal virtual boundary improving DMU performance by considering IVF and AVF as evaluation criteria, minimizing deviation from IVF and maximizing deviation from AVF. In 2019, Wu et al. [5] studied the environmental efficiency measurement of thermoelectric power plants using an efficient frontier DEA approach with fixed-sum undesirable output.

In 2020, Song et al. [9] Studied accident deaths as undesirable output in the production and safety evaluation in Chinese coal mines. In 2020, Walheer [1] studied the output, input, and undesirable output interconnections in data envelopment analysis: Convexity and returns-to-scale. In 2020, Yu et al. [8] assessed environmental provincial eco-efficiency in China an improved network data envelopment analysis model with undesirable output. In 2020, Gómez-Calvet et al. [7] evaluated European energy efficiency evaluation based on the use of super-efficiency under undesirable outputs in SBM models. In this research, we have presented the possibility of production in accordance with the concept of undesirable inputs and outputs. Then, in the concept of inputs and outputs, we have examined the efficiency of decision-making units and the efficiency boundary diagram with the presence of undesirable inputs and outputs by providing an example.

2. Production Possibility Set

Suppose we have n observations on n DMUs with input and output vectors (x_j, y_j) for j = 1, 2,..., n. Let $x_j = (x_1,...,x_{mj})^T$ and $y_j = (y_{1j},...,y_{sj})$. All $x_j \in R^m$ and $y_j \in R^s$ and $x_j > 0$, $y_j > 0$ *for j = 1, 2,... n.* The input matrix *X* and output matrix *Y* can be represented as $X = [x_1,...,x_j,...,x_n]$, $Y = [y_1,...,y_j,...,y_n]$.

Where *X* is an $(m \times n)$ matrix and *Y* an $(s \times n)$ matrix.

The production possibility set T is generally defined as

$$T = \{(x,y)1 \text{ x can produce y}\}. \quad (1)$$

In DEA, the production possibility set under a Variable Return to Scale (VRS) technology is constructed form the observed data (x_j, y_j) *for j = 1, 2, ..., n* as follows:

$$T = \left\{ (x,y) \middle| x \geq \sum_{j=1}^{n} \lambda_j x_j \ , \ y \leq \sum_{j=1}^{n} \lambda_j y_j \ , \lambda_j \geq 0 \ , \ \sum_{j=1}^{n} \lambda_j = 1, \ j = 1,\dots,n \right\}. \quad (2)$$

In the absence of undesirable factors when a DMU_O , $o \in \{1,2,...,n\}$, is under evaluation, we can use the following BCC model:

$$\begin{aligned} \min \ & \theta \\ s.t \ & \theta x_o - X\lambda \geq 0 \\ & Y\lambda \geq y_o \ , \\ & 1^T \lambda = 1, \\ & \lambda \geq 0. \end{aligned} \quad (3)$$

Corresponding to each output *y*, *L(y)* is defined as the following:

$$L(y_j) = \{x | (x, y_j) \in T\}. \quad (4)$$

In fact, $L(y_j)$ is a function that y_j portrays to a subset of inputs so that inputs can produce y_j.

Now suppose that some inputs are undesirable so input matrix X can be represented

as $X=(X^g,X^b)^T$, where $X^g=(x_{1j}^g,\ldots,x_{m_1j}^g), j=1,\ldots,n$ and $X^b=(x_{1j}^b,\ldots,x_{m_1j}^b)$ $j=1,\ldots,n$ *are* $(m_1\times n)$ *and* $(m_2\times n)$ matrixes that represent desirable (good) and undesirable (bad) inputs, respectively. And similarly, suppose that some outputs are undesirable so outputs. Matrix Y can be represented as $Y=(Y^g,Y^b)^T$, where $Y^g=(y_{1j}^g,\ldots,y_{s_1j}^g), j=1,\ldots,n$ and $Y^b=(y_{1j}^b,\ldots,y_{s_2j}^b)$, $j=1,\ldots,n$ are $(s_1\times n)$ and $(s_2\times n)$ matrixes that represent. Desirable (good) and undesirable (bad) inputs, respectively.

Definition 1. Let DMU of $(x_1^g,x_1^b,y_1^g,y_1^b)$ is dominant to DMU of $(x_2^g,x_2^b,y_2^g,y_2^b)$ if $(x_1^g\le x_2^g,x_1^b\ge x_2^b,y_1^g\ge y_2^g)$ and $y_1^b\le y_2^b$ the unequal be strict at least in a component. So that,

$$\begin{pmatrix}-x_1^g\\ x_1^b\\ y_1^g\\ -y_1^b\end{pmatrix}\ge\begin{pmatrix}-x_2^g\\ x_2^b\\ y_2^g\\ -y_2^b\end{pmatrix}.$$

Definition 2. DMU_O is efficient if in T there is no DMU to be dominant over it.

We consider the properties of the Production Possibility Set as the following:

- T is convex.
- T is closed.
- The monotony property of desirable inputs and outputs. So that, $\forall u\in R_+^{m_1},v\in R_+^{s_1},(x^g,x^b,y^g,y^b)\in T\Rightarrow(x^g+u,x^b,y^g-v,y^b)\in T$

This is not necessarily established for undesirable factors, because in this case, T has no efficient DMU.

We can define the production possibility set T satisfying *Eq. (1)* through *Eq. (3)* by

$$T=\left\{(x^g,x^b,y^b,y^g)\left|\begin{array}{c}x^g\ge\sum_{j=1}^n\lambda_jx_j^g,x^b=\sum_{j=1}^n\lambda_jx_j^b,y^b=\sum_{j=}^n\lambda_jy_j^b,y^g\le\sum_{j=1}^n\lambda_jy_j^g\\ \sum_{j=1}^n\lambda_j=1,\lambda_j\ge0,j=1,\ldots,n\end{array}\right.\right\}\quad(5)$$

3. Measures of Efficiency Using Undesirable Factors

In input oriented data, the efficiency of the DMU under evaluation is obtained by decreasing and increasing the desirable and undesirable input, respectively. And similarly, in output oriented data, we increase desirable output and decrease the undesirable output.

3.1. Nature of the Input

Suppose $DMU_o = (x_o^g, x_o^b, y_o^g, y_o^b)$ be unit under evaluation, corresponding to the output $y_o = (y_o^g, y_o^b)$ and using *Eq. (2)* $L(y_o^g, y_o^b)$ in defined as follows:

$$L(y_o^g, y_o^b) = \left\{(x^g, x^b) \middle| (x^g, x^b, y_o^g, y_o^b) \in T\right\}. \tag{6}$$

And we consider the subset of $L(y_o^g, y_o^b)$ as:

$$\partial^p L(y_o^g, y_o^b) = \left\{(x^g, x^b) \middle| \forall (u,v) \geq 0, (u,v) \neq 0 \quad \Rightarrow \quad (x^g - u, x^b + v) \notin L(y_o^g, y_o^b)\right\}. \tag{7}$$

That $\partial^s L(y_o^g, y_o^b)$ includes all inputs of the efficient DMUs which can produce (y_o^g, y_o^b).

The model to evaluate the efficiency of DMUo with the most decrease of x_o^g and the most increase of x_o^b is as follows:

$$d_o{}^g = x_o{}^g,$$

$$d_o{}^b = x_o{}^b - x_{max}^b.$$

So that

$$(x_{max}^b)_i = Max_j \left\{x_{ij}^b\right\}.$$

Therefore, according to the definition of inefficiency we have:

$$\begin{aligned}
&\theta_o^* = Max \quad \theta_o, \\
&st. \\
&\quad \sum_{j=1}^{n} \lambda_j x_j^g + s^- = x_o^g - \theta_o d_o{}^g. \\
&\quad \sum_{j=1}^{n} \lambda_J x_j^b = x_o^b - \theta d_o{}^b. \\
&\quad \sum_{j=1}^{n} \lambda_j y_j^g - s^+ = y_o^g. \\
&\quad \sum_{j=1}^{n} \lambda_j y_j^b = y_o^b. \\
&\quad \sum_{j=1}^{n} \lambda_j = 1. \\
&\quad \lambda_j \geq 0 \qquad \text{for} \quad \text{all} \quad j = 1, \ldots, n.
\end{aligned} \tag{8}$$

According to the definition of production possibility set, *model (1)* is possible in this set.

Theorem 1. The DMUo in *model (8)* is efficient if and only if

- $\theta_o^* = 1$.
- All slacks are zero for all optimal solutions.

Theorem 2. If all optimal solution of *model (8)* be (θ^*, s^{-*}), then

$$(x^g - \theta^* d^g - s^{-*}, x^b - \theta^* d^b) \in \partial^p L(y_o^b, y_o^g).$$

s^- is one of optimal answers.

4. Numerical Example 1

As an example, consider seven DMUs with one desirable input, one undesirable input to produce a desirable output normalized at level 1. These DMUs were explained in *Table 1*.

Regarding *Table 1* and *Fig. 1*, it can be seen that DMUs D, E, and F are efficient and they are on the $\partial^s L(y_G^g)$. On the other hand, efficiency of other DMUs have been examined through their image on $\partial^s L(y_G^g)$. (Efficient Frontiers).

Table 1. *The inputs and outputs data for 7 DMUs.*

DMU's	x^g	x^b	y^g	$1-\theta^*$
A	3	1	1	0.33
B	2	2	1	0.5
C	1	3	1	1
D	1	5	1	1
E	2	6	1	1
F	3	7	1	1
G	4	4	1	0.43

Similar discussion can be presented for the output oriented.

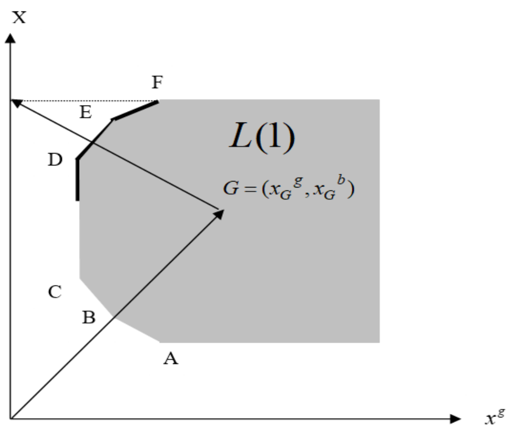

Fig. 1. *The graph of the* $L(y_G)$.

3.1. Nature of the Output

Suppose $DMU_o = (x_o^g, x_o^b, y_o^g, y_o^b)$ be unit under evaluation, corresponding to the output $x_o = (x_o^g, x_o^b)$ and using *Eq. (2)* $p(x_o^g, x_o^b)$ is defined as follows:

$$p(x_o^g, x_o^b) = \{(y^g, y^b) | (x_o^g, x_o^b, y^g, y^b) \in T\}.$$

And we consider the subset of $p(x_o^g, x_o^b)$ as:

$$\partial^p p(x_o^g, x_o^b) = \{(y^g, y^b) | \forall (u,v) \geq 0, (u,v) \neq 0 \quad \Rightarrow \quad (y^g + u, y^b - v) \notin p(x_o^g, x_o^b)\} \qquad (9)$$

That $\partial^s L(y_o^g, y_o^b)$ includes all inputs of the efficient DMUs which can produce (y_o^g, y_o^b).

The model to evaluate the efficiency of DMUo with the most decrease of y_o^g and the most increase of y_o^b is as follows:

$$NE^d(x_o, y_o) = \sup\{\beta | y_o + \beta d \in p(x_o)\}.$$

where $d = (d^g, d^b)$ indicate the direction of unit under evaluation such that $d^g \in R_+^{s_1}$ and $d \in R_-^{m_2}$ leads to increase the corresponding outputs and decreasing the unconfirmed outputs.

In this research, we direct the desired outputs to the efficient boundary in a radial direction. Thus: $d^g = y_o^g$.

We also reduce the undesirable outputs in the radial direction, i.e.

$$d^I = -y_o^b.$$

Therefore, according to the definition we have:

$$\begin{aligned}
&\beta_o^* = \text{Max} \quad \beta_o, \\
&\text{st.} \\
&\quad \sum_{j=1}^{n} \lambda_j x_j^g + s^- = x_o^g. \\
&\quad \sum_{j=1}^{n} \lambda_J x_j^b = x_o^b. \\
&\quad \sum_{j=1}^{n} \lambda_j y_j^g - s^+ = y_o^g + \beta_o d_o^g. \\
&\quad \sum_{j=1}^{n} \lambda_j y_j^b = y_o^b + \beta_o d_o^b. \\
&\quad \sum_{j=1}^{n} \lambda_j = 1. \\
&\quad \lambda_j \geq 0 \qquad \text{for} \quad \text{all} \quad j = 1, \ldots, n.
\end{aligned}$$

Theorem 3. The DMUo in *model (10)* is efficient if and only if

- $\beta_o^* = 1$.
- All slacks are zero for all optimal solutions.

Theorem 3. If be optimal solution of *model (10)* in, then

$$(y_o^* + \beta_o^* d_o^g + s^{+^*}, y_o^b + \beta_o^* d_o^b) \in \partial^p p(x_o^g, x_o^b).$$

5. Numerical Example 2

We consider five decision-making units with an optimal input to produce an undesirable output and a desirable output. These decision-making units are described in *Table 2. Fig. 2* shows that the decision-making units D, E and F are efficient. On the other hand, other decision-making units have been examined through their image on the (efficient border) of their efficiency.

Table 2. *The inputs and outputs data for 5 DMUs by model 10.*

$DMU's$	x^g	y^g	y^b	$1-\beta^*$
A	*1*	*4*	*1*	*1*
B	*1*	*5*	*2*	*1*
C	*1*	*5*	*4*	*1*
D	*1*	*4*	*5*	*0.25*
E	*1*	*3*	*3*	*0.5*

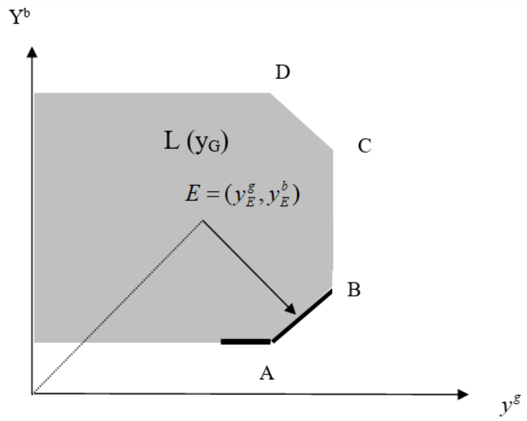

Fig. 2. *The graph of the L.*

6. Conclusion

Our proposed models in this study determine the efficiency of decision-making units, assuming that some of their input and output components may be undesirable. Numerical examples and model diagrams show that these models ensure that the presence of undesirable input and output factors is effective in determining the efficiency boundary of the decision-making units under evaluation and are compared with a unit corresponding to the efficient boundary set. By decreasing undesirable output and increasing undesirable input, the efficiency of decision-making units can improve and push them towards the efficient frontier.

References

[1] Walheer, B. (2020). Output, input, and undesirable output interconnections in data envelopment analysis: convexity and returns-to-scale. *Annals of operations research*, *284*(1), 447-467.

[2] Charnes, A., Cooper, W. W., & Rhodes, E. (1978). Measuring the efficiency of decision making units. *European journal of operational research*, *2*(6), 429-444**.**

[3] Chen, Y., & Ali, A. I. (2002). Output–input ratio analysis and DEA frontier. *European journal of operational research*, *142*(3), 476-479.

[4] Jahanshahloo, G. R., Lotfi, F. H., & Zohrehbandian, M. (2005). Finding the piecewise linear frontier production function in data envelopment analysis. *Applied mathematics and computation*, *163*(1), 483-488.

[5] Wu, J., Xia, P., Zhu, Q., & Chu, J. (2019). Measuring environmental efficiency of thermoelectric power plants: a common equilibrium efficient frontier DEA approach with fixed-sum undesirable output. *Annals of operations research*, *275*(2), 731-749.

[6] Kao, H. Y., Wu, D. J., & Huang, C. H. (2017). Evaluation of cloud service industry with dynamic and network DEA models. *Applied mathematics and computation*, *315*, 188-202.

[7] Gómez-Calvet, R., Conesa, D., Gómez-Calvet, A. R., & Tortosa-Ausina, E. (2020). European energy efficiency evaluation based on the use of super-efficiency under undesirable outputs in SBM models. In *Advances in efficiency and productivity II* (pp. 193-208). Springer, Cham.

[8] Yu, S., Liu, J., & Li, L. (2020). Evaluating provincial eco-efficiency in China: an improved network data envelopment analysis model with undesirable output. *Environmental science and pollution research*, *27*(7), 6886-6903.

[9] Song, M., Wang, J., Zhao, J., Baležentis, T., & Shen, Z. (2020). Production and safety efficiency evaluation in Chinese coal mines: accident deaths as undesirable output. *Annals of operations research*, *291*, 827–845.