DETERMINING THE STRONGLY AND WEAKLY MOST CONGESTED FIRMS IN DATA ENVELOPMENT ANALYSIS

(Only complete the Author list after the review process has been completed)
MOHAMMAD KHOVEYNI
Department of Applied Mathematics, Yadegar-e-Imam Khomeini (RAH) Shahre Rey Branch, Islamic Azad University, Tehran, Iran
mohammadkhoveyni@gmail.com (Corresponding)

ROBABEH ESLAMI
Department of Mathematics, Faculty of Technology and Engineering, South Tehran Branch, Islamic Azad University, Tehran, Iran
sinhx_2002@yahoo.com

ABSTRACT

Determining the strongly and weakly most congested firms (or decision making units (DMUs)) is a very important issue for decision makers, however, there is no study on determining these DMUs in data envelopment analysis (DEA). Hence, to do this, we first propose a DEA approach to recognize congestion status of DMUs and then, two integrated mixed integer programming (MIP)-DEA models is presented to determine the strongly and weakly most congested DMUs. Finally, a numerical example is provided to illustrate the purpose of this research.

Key words: Data envelopment analysis (DEA); Negative data in DEA; Strong and weak congestion; Slack variables; Mixed integer programming (MIP)

INTRODUCTION

Data envelopment analysis (DEA), which was initially introduced by Charnes et al. (CCR) (1978), is a mathematical tool for evaluating the relative efficiency of a set of decision making units (DMUs) with multiple inputs and multiple outputs.
One of the important issues in the DEA literature is recognizing congestion that so far there are many studies in this context.

For instance, a common concept of congestion was presented by Färe and Svensson (1980), then Färe and Grosskopf (1983) extended the topic of congestion such that they considered some inputs of the target DMU as constant values. Brockett *et al.* (1998) indicated that this assumption is not important and congestion of the DMU under evaluation can be removed by decreasing all its inputs. Also, Cooper *et al.* (1996) discussed about the treatment of congestion. In this vein, Copper et al. (2000) introduced a unified additive model approach to identify congestion.

Furthermore, Tone and Sahoo (2004) proposed a unified DEA approach to specify strong and weak congestion. However, their proposed congestion approach has two problems: (1) it is not able to recognize congestion status in the presence of alternative optimal solutions and (2) it is considered that all inputs and output of DMUs are positive. To tackle these problems, Khoveyni *et al.* (2013) presented a slack-based DEA approach for detecting congestion status in the presence of non-negative data. In addition, since there are some DMUs including desirable and undesirable data, Wu *et al.* (2013) and Fang (2015) introduced some DEA approaches to identify congestion of DMUs in this case. Furthermore, Khoveyni et al. (2017) provided a DEA approach to detect congestion status of DMUs in the presence of both negative and non-negative data.

Also, one of the important issues in the DEA literature is to determine the strongly and weakly most congested DMUs that there are no studies in this context. S*trongly most congested DMU* is a strongly congested DMU which has most total of values of the additional inputs; and similarly, *weakly most congested DMU* is a weakly congested DMU which has most total of values of the additional inputs. Thus, in this research, we present two integrated MIP-DEA models to specify the strongly and weakly most congested DMUs in the presence of negative and non-negative inputs and outputs.

The rest of this paper is organized as follows. Section 2 briefly describes the concepts of congestion. Two integrated MIP-DEA models are proposed to determine the strongly and weakly most congested DMUs in Section 3. In Section 4, a numerical example is provided to illustrate the proposed models. Finally, in last section, some concluding comments are presented.

2. Implications of congestion

Consider a set of n firms (or DMUs) such that $DMU_j \ (j=1,...,n)$ uses an input vector of $\mathbf{x}_j=\left(x_{1j},x_{2j},\ldots,x_{mj}\right)^T \underset{\neq}{\geq} 0$ to produce output vector of $\mathbf{y}_j=\left(y_{1j},y_{2j},\ldots,y_{sj}\right)^T \underset{\neq}{\geq} 0$. In the absence of input strong disposability assumption, the production possibility set (PPS) is as below:

$$P_{Convex}=\left\{\begin{pmatrix}\mathbf{x}\\ \mathbf{y}\end{pmatrix}\middle|\sum_{j=1}^{n}\lambda_j\mathbf{x}_j=\mathbf{x},\ \sum_{j=1}^{n}\lambda_j\mathbf{y}_j\geq\mathbf{y},\ \sum_{j=1}^{n}\lambda_j=1,\ \lambda_j\geq 0;\ j=1,2,\ldots,n\right\}. \quad (1)$$

Definition 1. DMU_k has strong congestion if an increase (decrease) in all of its inputs causes a decrease (increase) in all of its outputs.
Definition 2. DMU_k has weak congestion if an increase (decrease) in some of its inputs causes a decrease (increase) in some of its outputs.

In the DEA literature, the congested DMUs are classified to two groups as "strongly congested DMUs" and "weakly congested DMUs". In this research, we specify "strongly most congested DMU" and "weakly most congested DMU" by presenting two integrated MIP-DEA models in next section.

3. Determining the strongly and weakly most congested DMUs

Consider a set of n DMUs as $DMU_j=\left(\mathbf{x}_j^T,\mathbf{y}_j^T\right)\ (j=1,...,n)$ such that DMU_j uses input vector of $\mathbf{x}_j\in\mathbb{R}^m$ to produce output vector of $\mathbf{y}_j\in\mathbb{R}^s$. In this paper, we first use the Khoveyni *et al.* (2017) approach to recognize the strongly and weakly congested DMUs in the presence of both negative and non-negative inputs and outputs. Then, we define two sets as below:

$$S=\left\{j\,\middle|\,DMU_j\ has\ strong\ congestion\right\}, \quad (2)$$
$$W=\left\{j\,\middle|\,DMU_j\ has\ weak\ congestion\right\}. \quad (3)$$

Definition 3. DMU_k is called "strongly most congested DMU" if and only if the following two conditions are satisfied:

(a) DMU_k has strong congestion, i.e. $k\in S$,
(b) DMU_k has most total of values of the additional inputs.

Definition 4. DMU_k is called "weakly most congested DMU" if and only if the following two conditions are satisfied:

(c) DMU_k has weak congestion, i.e. $k \in W$,
(d) DMU_k has most total of values of the additional inputs.

3.1. The strongly most congested DMU

In the case that $S \neq \varnothing$, we have the following two cases:

Case (1). $Card(S) = 1$. In this case, there is only a single strongly congested DMU, thus this DMU is the strongly most congested DMU.

Case (2). $Card(S) \geq 2$. In this case, we first introduce a two-stage process by using the following two DEA models for each $k \in S$. In the first stage (model (4)), sum of the input and output slacks is maximized and we get a projection point as $\left(\hat{\mathbf{x}}_k^T, \hat{\mathbf{y}}_k^T\right) = \left(\mathbf{x}_k^T - \mathrm{s}^{-*}, \mathbf{y}_k^T + \mathrm{s}^{+*}\right)$; then in the second stage (model (5)), the optimal input slacks in evaluating the projection point are obtained.

$$\begin{aligned} Max \quad & \sum_{i=1}^{m} s_i^- + \sum_{r=1}^{s} s_r^+ \\ s.t. \quad & \sum_{j=1}^{n} \lambda_j x_{ij} + s_i^- = x_{ik}, \quad i = 1,\ldots,m, \\ & \sum_{j=1}^{n} \lambda_j y_{rj} - s_r^+ = y_{rk}, \quad r = 1,\ldots,s, \\ & \sum_{j=1}^{n} \lambda_j = 1, \qquad \lambda_j \geq 0, \quad j = 1,\ldots,n, \\ & s_i^- \geq 0, \quad i = 1,\ldots,m, \qquad s_r^+ \geq \varepsilon, \quad r = 1,\ldots,s, \end{aligned} \tag{4}$$

and

$$\begin{aligned} Max \quad & \sum_{i=1}^{m} t_i \\ s.t. \quad & \sum_{j=1}^{n} \lambda_j x_{ij} - t_i = x_{ik} - s_i^{-^*} = \hat{x}_{ik}, \quad i = 1,\ldots,m, \\ & \sum_{j=1}^{n} \lambda_j y_{rj} = y_{rk} + s_r^{+^*} = \hat{y}_{rk}, \quad r = 1,\ldots,s, \\ & \sum_{j=1}^{n} \lambda_j = 1, \qquad \lambda_j \geq 0, \quad j = 1,\ldots,n, \\ & t_i \leq s_i^{-^*}, \quad i = 1,\ldots,m, \end{aligned} \tag{5}$$

where, in model (4), $s_i^- \ (i = 1,\ldots,m)$ and $s_r^+ \ (r = 1,\ldots,s)$ are respectively the input and output slacks. Moreover, in model (5), $s_i^{-^*} \ (i = 1,\ldots,m)$ and $s_r^{+^*} \ (r = 1,\ldots,s)$ are the optimal input and output slacks obtained from model (4).

So, the congestion value for i^{th} input of DMU_k $\left(where\ k \in S\right)$ is identified as below:

$$c_{ik} = s_i^{-*} - t_i^*, \qquad (i = 1,\ldots,m), \tag{6}$$

where t_i^* $(i = 1,\ldots,m)$ are the optimal values obtained from model (5).

Then, we introduce the following DEA model for each $k \in S$:

$$Max \quad \sum_{i=1}^{m} \alpha_i \tag{7}$$

$$s.t. \quad \sum_{j=1}^{n} \lambda_j x_{ij} + \alpha_i = x_{ik}, \quad i = 1,\ldots,m, \tag{7.1}$$

$$\sum_{j=1}^{n} \lambda_j y_{rj} = \hat{y}_{rk}, \quad r = 1,\ldots,s, \tag{7.2}$$

$$\sum_{j=1}^{n} \lambda_j = 1, \qquad \lambda_j \geq 0, \quad j = 1,\ldots,n, \tag{7.3}$$

$$\alpha_i \leq c_{ik}, \quad i = 1,\ldots,m, \tag{7.4}$$

$$\alpha_i \geq 0, \quad i = 1,\ldots,m.$$

In model (7), α_i $(i = 1,\ldots,m)$ is the additional value of i^{th} input of the target DMU, i.e. $\left(\mathbf{x}_k^T, \hat{\mathbf{y}}_k^T\right)$.

Therefore, $\sum_{i=1}^{m} \alpha_i$ indicates total of values of additional inputs in assessing $\left(\mathbf{x}_k^T, \hat{\mathbf{y}}_k^T\right)$.

The dual form of model (7) is as below:

$$\begin{aligned} Min \quad & \delta_k = \sum_{i=1}^{m} v_i x_{ik} + \sum_{r=1}^{s} u_r \hat{y}_{rk} + \sum_{i=1}^{m} w_i c_{ik} + t \\ s.t. \quad & \sum_{i=1}^{m} v_i x_{ij} + \sum_{r=1}^{s} u_r y_{rj} + t \geq 0, \quad j = 1,\ldots,n, \\ & v_i + w_i \geq 1, \quad i = 1,\ldots,m, \\ & w_i \geq 0, \quad i = 1,\ldots,m, \\ & v_i : free, \quad i = 1,\ldots,m, \\ & u_r : free, \quad r = 1,\ldots,s, \\ & t : free, \end{aligned} \tag{8}$$

where v_i $(i = 1,\ldots,m)$, u_r $(r = 1,\ldots,s)$, t, and w_i $(i = 1,\ldots,m)$ are respectively the dual variables corresponding to the constraints (7.1), (7.2), (7.3), and (7.4).

Here, we propose an integrated MIP-DEA model to specify the strongly most congested DMU as follows:

$$
\begin{aligned}
Min \quad & \delta_{max} \\
s.t. \quad & \delta_{max} - \sum_{i=1}^{m} v_i x_{ik} - \sum_{r=1}^{s} u_r \hat{y}_{rk} - \sum_{i=1}^{m} w_i c_{ik} - t - \beta_k = 0, \quad k \in S, \\
& \sum_{i=1}^{m} v_i x_{ij} + \sum_{r=1}^{s} u_r y_{rj} + t \geq 0, \quad j = 1,\ldots,n, \\
& \sum_{k \in S} \theta_k = Card(S) - 1, \\
& v_i + w_i \geq 1, \quad i = 1,\ldots,m, \\
& \beta_k \leq M\theta_k, \quad k \in S, \\
& \theta_k \leq N\beta_k, \quad k \in S, \\
& \theta_k \in \{0,1\}, \quad k \in S, \\
& \beta_k \geq 0, \quad k \in S, \qquad w_i \geq 0, \quad i = 1,\ldots,m, \\
& v_i : free, \quad i = 1,\ldots,m, \\
& u_r : free, \quad r = 1,\ldots,s, \qquad t : free,
\end{aligned} \tag{9}
$$

where M and N are large enough numbers. In model (9), we have:

- If $\theta_k = 1$, then $\beta_k > 0$. Because, in this case, the constraint $\beta_k \leq M\theta_k$ is redundant and the constraint $\theta_k \leq N\beta_k$ forces β_k to take a positive number.
- If $\theta_k = 0$, then $\beta_k = 0$. Because, in this case, the constraint $\theta_k \leq N\beta_k$ is redundant and the constraints $\beta_k \leq M\theta_k$ and $\beta_k \geq 0$ make β_k to be zero.

Theorem 1. *If $Card(S) \geq 2$, then a single strongly most congested DMU is only obtained by applying model (9) that δ^*_{max} is the total of values of its additional inputs with a common set of optimal weights.*

Proof. Assume that $\left(\delta^*_{max}, \mathbf{v}^*, \mathbf{u}^*, \mathbf{w}^*, t^*, \boldsymbol{\beta}^*, \boldsymbol{\theta}^*\right)$ is an optimal solution obtained from model (9). Also, suppose that $\theta^*_p = 0$ and $\theta^*_k = 1$ for each $k \in S - \{p\}$. So, $\beta^*_p = 0$ and $\beta^*_k > 0$ for each $k \in S - \{p\}$. Hence, we attain:

$$\delta^*_{max} = \mathbf{v}^{*^T} \mathbf{x}_p + \mathbf{u}^{*^T} \hat{\mathbf{y}}_p + \mathbf{w}^{*^T} \mathbf{c}_p + t^*. \tag{10}$$

Therefore, there exists at least one strongly most congested DMU, i.e. DMU_p, by using the common set of optimal weights $\left(\mathbf{v}^*, \mathbf{u}^*, \mathbf{w}^*\right)$.

Conversely, on the contrary, suppose that DMU_q $\left(where\ q \in S - \{p\}\right)$ is another strongly most congested DMU by using the common set of optimal weights $\left(\mathbf{v}^*, \mathbf{u}^*, \mathbf{w}^*\right)$. So, we deduce that total of values of the additional inputs of DMU_q is also as below:

$$\delta^*_{max} = \mathbf{v}^{*^T} \mathbf{x}_q + \mathbf{u}^{*^T} \hat{\mathbf{y}}_q + \mathbf{w}^{*^T} \mathbf{c}_q + t^*. \tag{11}$$

Thus, according to (10) and (11), we get $\beta_p^* = \beta_q^* = 0$ and so $\theta_p^* = \theta_q^* = 0$. Hence, we find that $\sum_{k \in S} \theta_k^* = Card(S) - 2$ that is a contradiction. So, DMU_p and DMU_q cannot be strongly most congested DMU by using the common set of optimal weights $(\mathbf{v}^*, \mathbf{u}^*, \mathbf{w}^*)$, simultaneously. Thus, the proof is complete. **Q.E.D.**

3.2. The weakly most congested DMU

In the case that $W \neq \varnothing$, we have the following two cases:

Case (3). $Card(W) = 1$. In this case, there is only a single weakly congested DMU hence, this DMU is the weakly most congested DMU.

Case (4). $Card(W) \geq 2$. Models (4), (5), (7), and (8) and relation (6) are used for each $k \in W$ instead of $k \in S$ and then, the an integrated MIP-DEA model is similarly presented to specify the weakly most congested DMU as below:

$$
\begin{aligned}
Min \quad & \eta_{max} \\
s.t. \; & \eta_{max} - \sum_{i=1}^{m} v_i x_{ik} - \sum_{r=1}^{s} u_r \hat{y}_{rk} - \sum_{i=1}^{m} w_i c_{ik} - t - \gamma_k = 0, \quad k \in W, \\
& \sum_{i=1}^{m} v_i x_{ij} + \sum_{r=1}^{s} u_r y_{rj} + t \geq 0, \quad j = 1, \ldots, n, \\
& \sum_{k \in W} \varphi_k = Card(W) - 1, \\
& v_i + w_i \geq 1, \quad i = 1, \ldots, m, \\
& \gamma_k \leq M \varphi_k, \quad k \in W, \\
& \varphi_k \leq N \gamma_k, \quad k \in W, \\
& \varphi_k \in \{0,1\}, \quad k \in W, \\
& \gamma_k \geq 0, \quad k \in W, \qquad w_i \geq 0, \quad i = 1, \ldots, m, \\
& v_i : free, \quad i = 1, \ldots, m, \qquad u_r : free, \quad r = 1, \ldots, s, \qquad t : free,
\end{aligned}
\tag{12}
$$

where M and N are large enough numbers. In model (12):

- If $\varphi_k = 1$, then $\gamma_k > 0$. Because, in this case, $\gamma_k \leq M \varphi_k$ is a redundant constraint and the constraint $\varphi_k \leq N \gamma_k$ forces γ_k to take a positive number.
- If $\varphi_k = 0$, then $\gamma_k = 0$. Because, in this case, $\varphi_k \leq N \gamma_k$ is a redundant constraint and the constraints $\gamma_k \leq M \varphi_k$ and $\gamma_k \geq 0$ make γ_k to be zero.

Theorem 2. *If $Card(W) \geq 2$, then a single weakly most congested DMU is only obtained by using model (12) that η_{max}^* is the total of values of its additional inputs with a common set of optimal weights.*

Proof. This theorem is proved similar to Theorem 1. **Q.E.D.**

4. Numerical example

Here, the proposed approach is applied to recognize the congestion status of fifteen DMUs with one input and one output in the presence of both negative and non-negative data. Table 3 shows the dataset of these DMUs.

Table 3. Dataset of the DMUs.

DMU	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
Input 1	-1	-3	0	-2	-2	2	4	-2	-2	4	3	-2	3	2	3
Input 2	-3	-1	-2	0	2	-2	-2	4	2	4	-2	3	3	1	3
Output	-1	-1	1	1	1	1	0	0	1	0	0.5	0.5	0.5	-2	-3

First, we use the Khoveyni *et al.* (2017) congestion approach and find that DMUs 10 and 13 have strong congestion, DMUs 7, 8, 11, and 12 have weak congestion, and the rest of DMUs have no congestion. So, we deduce that $S=\{10,13\}$ and $W=\{7,8,11,12\}$. Since $Card(S)=2$ and $Card(W)=4$, we use models (9) and (12) to determine the strongly and weakly most congested DMUs, respectively, (see *Cases (2)* and *(4)). Table 4 shows the obtained results from models (9) and (12).

Table 4. Results of models (9) and (12), ($M=1500$ and $N=1000$).

Strongly congested DMU	Model (9)			Weakly congested DMU	Model (12)		
	δ^*	θ^*	β^*		η^*	φ^*	γ^*
10	$\delta^*_{max}=$ 10	0	0 (Strongly most congested DMU)	7	4	1	0.1429
13	8	1	2	8	$\eta^*_{max}=$ 4.1429	0	0 (Weakly most congested DMU)
				11	3	1	1.1429
				12	3.0714	1	1.0715

According to Table 4, we have $\theta^*_{10}=0$ and $\theta^*_{13}=1$. So, as per *Case (2)*, DMU_{10} is "strongly most congested DMU" with the most total of the additional inputs $\delta^*_{max}=10$. Also, we have $\varphi^*_8=0$ and $\varphi^*_7=\varphi^*_{11}=\varphi^*_{12}=1$ therefore, based on *Case (4)*, DMU_8 is "weakly most congested DMU" with the most total of the additional inputs $\eta^*_{max}=4.1429$.

5. Concluding remarks

So far there are many congestion approaches in the DEA literature however none of them do not discuss about strongly and weakly congested DMUs. Hence, in this study, we propose two integrated MIP-DEA models to determine the strongly and weakly congested DMUs in the presence of both negative and non-negative data. Lastly, a numerical example is provided to illustrate the proposed models.

REFERENCES

Brockett, P.L., Cooper, W.W., Shin, H.C., Wang, Y., (1998(, Inefficiency and congestion in Chinese production before and after the 1978 economic reforms. Socio-Economic Planning Sciences, 32, 1-20.

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

Cooper, W.W., Thompson, R.G., Thrall, R.M., (1996), Extensions and new developments in DEA. Annals of Operations Research, 66, 3-45.

Fang, L., (2015), Congestion measurement in nonparametric analysis under the weakly disposable technology. European Journal of Operational Research, 245, 203-208.

Färe, R., Svensson, L., (1980), Congestion of production factors. Econometrica, 48, 1745-1753.

Färe, R., Grosskopf, S., (1983), Measuring congestion in production. Journal of Economics, 43, 257-271.

Khoveyni, M., Eslami, R., Khodabakhshi, M., Jahanshahloo, G.R., Hosseinzadeh Lotfi, F., (2013), Recognizing strong and weak congestion slack based in data envelopment analysis. Computers & Industrial Engineering, 64, 731-738.

Khoveyni, M., Eslami, R., Yang, G.l., (2017), Negative data in DEA: Recognizing congestion and specifying the least and the most congested decision making units. Computers & Operations Research, 79, 39-48.

Tone, K., Sahoo, B.K., (2004), Degree of scale economies and congestion: A unified DEA approach. European Journal of Operational Research, 158, 755-772.

Wu, J., An, Q., Xiong, B., Chen, Y., (2013), Congestion measurement for regional industries in China: A data envelopment analysis approach with undesirable outputs. Energy Policy, 57, 7-13.