Accepted Manuscript

Congestion measurement in nonparametric analysis under the weakly disposable technology

Lei Fang

PII: S0377-2217(15)00194-0
DOI: 10.1016/j.ejor.2015.03.001
Reference: EOR 12822

To appear in: *European Journal of Operational Research*

Received date: 29 April 2014
Revised date: 26 December 2014
Accepted date: 2 March 2015

Please cite this article as: Lei Fang , Congestion measurement in nonparametric analysis under the weakly disposable technology, *European Journal of Operational Research* (2015), doi: 10.1016/j.ejor.2015.03.001

Highlights

► An approach for measuring congestion in the presence of desirable and undesirable outputs is developed.

► The proposed approach can discriminate between the congested DMUs and the truly efficient DMUs.

► An empirical example is used to illustrate the proposed approach.

Congestion measurement in nonparametric analysis under the weakly disposable technology

Lei Fang

Business School, Nankai University, People's Republic of China

Abstract: Congestion is a widely observed economic phenomenon where outputs are reduced due to excessive amount of inputs. The previous approaches to identify congestion in nonparametric analysis only consider desirable outputs. In the production process, undesirable outputs are usually jointly produced with desirable outputs. In this paper, we propose an approach for measuring congestion in the presence of desirable and undesirable outputs simultaneously. The proposed approach can discriminate between the congested DMUs and the truly efficient DMUs, which are all efficient according to the scores calculated by the directional distance function. Finally, an empirical example is used to illustrate the approach.

Keywords: Data Envelopment Analysis; Undesirable Output; Congestion

---

*Corresponding author. Tel.: +862223505660; fax: +862223505660.

E-mail address: fanglei@nankai.edu.cn.

1. Introduction

The concept of congestion, first introduced by Färe and Svensson (1980), is a widely phenomenon where excessive amounts of the input cause a reduction of the output. Subsequently, it was extended and developed by Färe et al. (1985) and Cooper et al. (1996, 2000) in the context of DEA (data envelopment analysis). Since then, the treatment of congestion within the DEA framework has received considerable attention and several approaches have been proposed to identify congestion (Brockett et al., 1998; äre and Grosskopf, 2000; Cooper et al. ,2001a; Cherchye et al., 2001;Tone and Sahoo, 2004; Sueyoshi and Sekitani 2009; Kao 2010; Khoveyni et al., 2013).

Färe et al. (1985) proposed a radial-model approach in which congestion is measured as the difference between technologies under weak and strong disposability inputs. Cooper et al. (1996) proposed a slack-based approach, where the congestion effect is measured as the difference between the observed amounts and the expected amounts. Cooper et al. (2001a) compared the above approaches and claimed that the approach by Färe et al. (1985) can fail to identify congestion in some situations. See some debates on the subject of congestion (Färe and Grosskopf, 2000; Cooper et al. ,2001b; Cherchye et al. ,2001).

Further, Tone and Sahoo (2004) provided a theoretical linkage between congestion and returns to scale (RTS). Moreover, their approach can detect the strong and weak congestion status. However, Tone and Sahoo (2004) implicitly assume a unique optimal solution in the investigation on DEA-based congestion. In the presence of multiple solutions in the congestion measurement, the economic implications of congestion obtained by Tone and Sahoo (2004) are all problematic from both theoretical and practical perspectives. To deal with the issue, Sueyoshi and Sekitani (2009) proposed an analytical approach to handle an occurrence of multiple solutions and measure the degree of wide congestion.

However, all the previous approaches on congestion only consider desirable outputs. In the production process, undesirable outputs are usually jointly produced with desirable outputs. Therefore, a new framework for measuring congestion should be developed in the presence of desirable and undesirable outputs simultaneously. A pioneering paper by Färe et al. (1989) considers undesirable outputs to be weakly disposable, which means that a reduction in the good outputs should result in an equiproportionate reduction of the undesirable outputs (Chung et al. 1997; Weber and Domazlicky 2001; Färe and Grosskopf 2003, 2004, 2009; Kuosmanen 2005; Kuosmanen and Kortelainen,2005; Zhou et al. 2008; Kuosmanen and Podinovski 2009; Kuosmanen and Matin 2011; Picazo-Tadeo et al. 2012). Treating undesirable outputs in their original forms with the assumption of weak-disposability is consistent with the physical laws and the standard axioms of production theory (Färe and Grosskopf, 2003, 2004, 2009; Sahoo et al. 2011). Based on the weak disposable technology, many empirical studies utilized the directional distance function model (Chung et al. 1997), which expands the desirable outputs and contracts inputs and undesirable outputs along the direction vector path to assess the efficiency.

In this paper, based on the directional distance function, we develop an approach to identify the occurrence of congestion (strong and weak) in the presence of desirable and undesirable outputs simultaneously. Different from the traditional circumstance with desirable outputs only, we find that even if a DMU is efficient by the directional distance function, it maybe evidences congestion when considering both desirable outputs and undesirable outputs. Through our proposed approach, we can discriminate between the congested DMUs and the truly efficient DMUs, which are all efficient according to the scores calculated by the directional distance function.

The remaining structure of this research is organized as follows: In Section2, the concepts of strong and weak congestions in the presence of desirable and undesirable outputs are defined. Section 3 proposes an approach to identify the occurrence of strong and weak congestions. Section 4 compares the proposed approach with the existing three representative approaches and applies the proposed approach to analyze an empirical dataset consisting of 20 power plants. Section 5 concludes the paper.

2. Preliminaries

Assume that there are K units and each unit uses a vector of inputs $x \in \mathbf{R}_{+}^{N}$ to produce a vector of good outputs $y \in \mathbf{R}_{+}^{M}$ and bad outputs $b \in \mathbf{R}_{+}^{I}$. The production technology consisting of all feasible (x, y, b) can be defined by:

$$\Omega = \{(x, y, b) | x \text{ can produce } (y, b)\} \tag{1}$$

Given that we consider K observed DMUs, the production technology set can be formulated as follows:

$$\Omega = \left\{ \begin{array}{ll} (x, y, b) : \theta \sum_{k=1}^{K} z_k y_{km} \geq y_m & m = 1, \cdots, M \\ \theta \sum_{k=1}^{K} z_k b_{ki} = b_i & i = 1, \cdots I \\ \sum_{k=1}^{K} z_k x_{kn} \leq x_n & n = 1, \cdots N \end{array} \right\} \tag{2}$$

By considering a sample of K observed DMUs, inefficiency for unit k_o exhibiting constant returns to scale and weak disposability can be computed by the following direction distance function (Chung et al. 1997):

$$IE(k_o) = \max \ \delta \tag{3}$$

$$\begin{array}{ll} \sum_{k=1}^{K} z_k y_{km} \geq y_{k_o m} + \delta y_{k_o m} & m = 1, \cdots, M \\ \sum_{k=1}^{K} z_k b_{ki} = b_{k_o i} - \delta b_{k_o i} & i = 1, \cdots I \\ \sum_{k=1}^{K} z_k x_{kn} \leq x_{k_o n} - \delta x_{k_o n} & n = 1, \cdots N \\ z_k \geq 0 & k = 1, \cdots, K \end{array}$$

where $z = (z_1, \cdots, z_K)$ are referred to the intensity variables and $g = (x_{k_o n}, y_{k_o m}, b_{k_o i})$ is the direction vector. The less $IE(k_o)$ is, the more efficient k_o is. If $IE(k_o) = 0$, unit k_o is efficient. Otherwise, it is inefficient. If unit k_o is inefficient, make a projection in the following manner:

$$x'_{k_o} = x_{k_0} - \delta^* x_{k_0}, \ y'_{k_o} = y_{k_o} + \delta^* y_{k_o} \ \text{ and } \ b'_{k_o} = b_{k_o} - \delta^* b_{k_o} . \tag{4}$$

The projected point $(x'_{k_0}, y'_{k_0}, b'_{k_0})$ is efficient with respect to Ω.

In the following, we use a simple example to illustrate the drawback of model (3). Table 1 shows the data set of five DMUs with two inputs (x_1 and x_2), two desirable outputs (y_1 and y_2) and two undesirable outputs (b_1 and b_2).

Table 1 The data set for illustration

	x_1	x_2	y_1	y_2	b_1	b_2	inefficiency
DMU1	1	3	2	8	3	2	0
DMU2	1	4	3	7	2.5	4	0
DMU3	2	6	1	6	5	35	0
DMU4	3	6	1	5	10	10	0
DMU5	1.5	3	1.5	8	3	3	0

The inefficiency scores by model (3) are listed in the last column of Table 1. According to the inefficiency scores in the last column, all DMUs are efficient. However, from DMU2 to DMU3, a phenomenon of congestion has occurred because the desirable output decreases and both undesirable outputs increase as the input increases.

Remark. According to Brockett et al. (2004), congestion is often referred to as a ''particularly severe form of inefficiency'' in terms of economics. A DMU evidences congestion if and only if it is not weakly efficient by DEA models when considering desirable outputs only (Wei and Yan, 2004). However, when considering both desirable and undesirable outputs, even if a DMU is efficient, it maybe evidences congestion.

In the real world, undesirable outputs such as smoke pollution or waste are unavoidably generated along with desirable outputs. Thus, in the above scenario, the outputs are divided into two categories, desirable and undesirable. For desirable outputs, the more the value is, the better the performance is while for undesirable outputs, the less the value is, the better performance is. Therefore, similar to Tone and Sahoo (2004), we first define the concepts of "strong congestion" in the presence of desirable and undesirable outputs:

Definition 1 A DMU $_k$ (x_k, y_k, b_k) is "strongly congested" if it is efficient and there exists an activity $(\tilde{x}, \tilde{y}, \tilde{b}) \in \Omega$ such that $\tilde{x} = \alpha x_k$ (with $0 < \alpha < 1$), $\tilde{y} = \beta y_k$ (with $\beta > 1$) and $\tilde{b} = \gamma b_k$ (with $0 < \gamma < 1$).

The above definition means that a DMU $_k$ (x_k, y_k, b_k) is in the status of strong congestion requires that a proportionate reduction in all inputs can give rise to an increase in all desirable outputs and a decrease in all desirable outputs. From this viewpoint, definition 1 is too restrictive in some cases. In the following, we define the concept of "weak congestion" by relaxing such stringent requirements.

Definition 2 A DMU $_k$ (x_k, y_k, b_k) is "weakly congested" if it is efficient and there exists an activity that uses less resources in one or more inputs to produce more products in one or more desirable outputs and less undesirable outputs in one or more undesirable outputs.

Note that strong congestion implies weak congestion but not vice versa. In a single input, a single desirable output and an undesirable output case, there is no distinction between strong and weak congestions.

3. Proposed Approach

In this section, we proposed an approach to identify the occurrence of congestion. By making use of the duality theory of linear programming, the dual formulation of the direction distance model (3) is described as follows:

$$\text{Min} \quad b_{k_0}\pi^b + x_{k_0}\pi^x - y_{k_0}\pi^y \tag{5}$$

$$\text{s.t.} \quad y_k\pi^y - b_k\pi^b - x_k\pi^x \le 0 \qquad k = 1, \cdots, K \tag{5.1}$$

$$b_{k_0}\pi^b + x_{k_0}\pi^x + y_{k_0}\pi^y = 1 \tag{5.2}$$

$$\pi^x \ge 0$$

$$\pi^y \ge 0$$

$$\pi^b \quad \text{unconstrained}$$

Let π^{x*}, π^{y*} and π^{b*} be the optimal solution to the model (4).

Theorem 1 A DMU $_{k_0}$ $(x_{k_0}, y_{k_0}, b_{k_0})$ is in the status of strong congestion if and only if for at least one $i \in \{1, \cdots I\}$, π_i^{b*} is negative.

Proof. "only if ":

As we assume that the DMU $_{k_0}$ $(x_{k_0}, y_{k_0}, b_{k_0})$ is efficient, it follows from the duality theorem of linear programming that there exists an optimal solution ($\pi^{x*}, \pi^{y*}, \pi^{b*}$) for model (4) such that

$$b_{k_0}\pi^{*b} + x_{k_0}\pi^{*x} - y_{k_0}\pi^{*y} = 0 \tag{6}$$

Since the DMU $_{k_0}$ $(x_{k_0}, y_{k_0}, b_{k_0})$ is strongly congested, there exists $(\tilde{x}, \tilde{y}, \tilde{b}) \in \Omega$ such that $\tilde{x} = \alpha x_k$ (with $0 < \alpha < 1$), $\tilde{y} = \beta y_k$ (with $\beta > 1$) and $\tilde{b} = \gamma b_k$ (with $0 < \gamma < 1$).

Obviously, $\tilde{b}\pi^{*b} + \tilde{x}\pi^{*x} - \tilde{y}\pi^{*y} \geq 0$ (7)

From equations (6) and (7), we obtain:

$$(\tilde{b} - b_{k_0})\pi^{*b} + (\tilde{x} - x_{k_0})\pi^{*x} + (y_{k_0} - \tilde{y})\pi^{*y} \geq 0 \tag{8}$$

As $\tilde{x} < x_{k_0}$, $y_{k_0} < \tilde{y}$, $\tilde{b} < b_{k_0}$ and $\pi^{*x} \geq 0$, $\pi^{*y} \geq 0$, we have:

$$(\tilde{x} - x_{k_0})\pi^{*x} \leq 0 \tag{9}$$

$$(y_{k_0} - \tilde{y})\pi^{*y} \leq 0 \tag{10}$$

Suppose that $\pi^{*b} \geq 0$, then $(\tilde{b} - b_{k_0})\pi^{*b} \leq 0$. (11)

According to the constraint (5.2), we have:

$$\pi^{*b} \neq 0, \pi^{*x} \neq 0 \text{ and } \pi^{*y} = 0. \tag{12}$$

According to equations (9)-(12), we have $(\tilde{b} - b_{k_0})\pi^{*b} + (\tilde{x} - x_{k_0})\pi^{*x} + (y_{k_0} - \tilde{y})\pi^{*y} < 0$, which contradicts equation (8).

Thus, there exists at least one $i \in \{1, \cdots I\}$ such that π_i^{b*} is negative.

"if ": We prove it by contradiction. Hence we shall show that if DMU $_{k_0}$ $(x_{k_0}, y_{k_0}, b_{k_0})$ is not strongly congested, there exists an optimal solution ($\pi^{x*}, \pi^{y*}, \pi^{b*}$) of model (5) such that $\pi_i^{b*} \geq 0$ for all $i \in \{1, \cdots I\}$.

Since DMU $_{k_0}$ $(x_{k_0}, y_{k_0}, b_{k_0})$ is not strongly congested, the following linear system

$$\sum_{k=1}^{K} z_k y_{km} > y_{k_o m} \qquad m = 1, \cdots, M$$

$$\sum_{k=1}^{K} z_k b_{ki} < b_{k_o i} \qquad i = 1, \cdots I$$

$$\sum_{k=1}^{K} z_k x_{kn} < x_{k_o n} \qquad n = 1, \cdots N$$

has no solution.

Thus, the optimal value of the following model (13) is zero.

$$\max \ \varphi \tag{13}$$

$$\sum_{k=1}^{K} z_k y_{km} \geq y_{k_o m} + \varphi y_{k_o m} \qquad m = 1, \cdots, M$$

$$\sum_{k=1}^{K} z_k b_{ki} \leq b_{k_o i} - \varphi b_{k_o i} \qquad i = 1, \cdots I$$

$$\sum_{k=1}^{K} z_k x_{kn} \leq x_{k_o n} - \varphi x_{k_o n} \qquad n = 1, \cdots N$$

The corresponding dual program to model (13) is given by:

$$\text{Min} \quad b_{k_0} \pi^b + x_{k_0} \pi^x - y_{k_0} \pi^y \tag{14}$$

$$\text{s.t.} \quad y_k \pi^y - b_k \pi^b - x_k \pi^x \leq 0 \qquad k = 1, \cdots, K$$

$$b_{k_0} \pi^b + x_{k_0} \pi^x + y_{k_0} \pi^y = 1$$

$$\pi^x \geq 0$$

$$\pi^y \geq 0$$

$$\pi^b \geq 0$$

Let ($\pi^{x*}, \pi^{y*}, \pi^{b*}$) be the optimal solution to model (14). Note that model (14) and model (5) differ only in the fifth set of constraints. Obviously, ($\pi^{x*}, \pi^{y*}, \pi^{b*}$) is the feasible solution to model (5). According to the duality theorem, the optimal value of model (14) is zero, that is, $b_{k_0} \pi^{*b} + x_{k_0} \pi^{*x} - y_{k_0} \pi^{*y} = 0$. Thus, ($\pi^{x*}, \pi^{y*}, \pi^{b*}$) is the optimal solution of model (5), in which $\pi_i^{b*} \geq 0$ for all $i \in \{1, \cdots I\}$.

From **Theorem 1**, a supporting hyperplane for DMU$_{k_0}$ $(x_{k_0}, y_{k_0}, b_{k_0})$ is mathematically specified by $\sum_{m=1}^{M} y_{k_0 m}\pi_m^y - \sum_{i=1}^{I} b_{k_0 i}\pi_i^b - \sum_{n=1}^{N} x_{k_0 n}\pi_n^x = 0$. Note that the marginal product is a differential characteristic of the production frontier. In dealing with multiple undesirable outputs, the following formulation calculates the marginal product of $y_{k_0 m}$, $m=1,\cdots,M$ and $x_{k_0 n}$, $n=1,\cdots N$ with respect to $b_{k_0 i}$:

$$MR_{mi}^{k_0} = \partial y_m \big/ \partial b_i = -\pi_m^{*y} \big/ \pi_i^b$$

$$MR_{ni}^{k_0} = \partial x_n \big/ \partial b_i = \pi_n^{*x} \big/ \pi_i^b$$

Thus, the negative sign of a dual variable related to an undesirable output implies the occurrence of congestion.

According to **Theorem 1,** we propose the following approach for identifying the occurrence of strong congestion:

$$\begin{aligned}
&\text{Max} \quad \alpha && (15)\\
&\sum_{k=1}^{K} z_k y_{km} \geq y_{k_o m} + \delta y_{k_o m} && m=1,\cdots,M\\
&\sum_{k=1}^{K} z_k b_{ki} = b_{k_o i} - \delta b_{k_o i} && i=1,\cdots I\\
&\sum_{k=1}^{K} z_k x_{kn} \leq x_{k_o n} - \delta x_{k_o n} && n=1,\cdots N\\
&b_{k_0}\pi^b + x_{k_0}\pi^x - y_{k_0}\pi^y = \delta &&\\
&y_k\pi^y - b_k\pi^b - x_k\pi^x \leq 0 && k=1,\cdots,K\\
&b_{k_0}\pi^b + x_{k_0}\pi^x + y_{k_0}\pi^y = 1 &&\\
&\pi_i^b - \alpha \geq 0 && i=1,\cdots I
\end{aligned}$$

Let $(\pi^{x*}, \pi^{y*}, \pi^{b*}, \delta^*, \alpha^*)$ be the optimal solution of model (15). If $\alpha^* < 0$, then the projected point ($x_{k_0} - \delta^* x_{k_0}$, $y_{k_o} + \delta^* y_{k_o}$, $b_{k_o} - \delta^* b_{k_o}$) of DMU k_0 is strongly congested. If $\alpha^* > 0$, then the projected point ($x_{k_0} - \delta^* x_{k_0}$, $y_{k_o} + \delta^* y_{k_o}$, $b_{k_o} - \delta^* b_{k_o}$) of DMU k_0 is not strongly congested.

If $\alpha^* = 0$, for the projected point ($x_{k_0} - \delta^* x_{k_0}$, $y_{k_o} + \delta^* y_{k_o}$, $b_{k_o} - \delta^* b_{k_o}$) of DMU k_0, we solve the following programming:

$$\text{Max} \quad \beta \qquad (16)$$

$$\sum_{k=1}^{K} z_k y_{km} \geq y_{k_o m} + \delta^* y_{k_o m} + t_m^+ \qquad m = 1, \cdots, M$$

$$\sum_{k=1}^{K} z_k b_{ki} \leq b_{k_o i} - \delta^* b_{k_o i} \qquad i = 1, \cdots I$$

$$\sum_{k=1}^{K} z_k x_{kn} \leq x_{k_o n} - \delta^* x_{k_o n} - t_n^- \qquad n = 1, \cdots N$$

$$\sum_{m=1}^{M} t_m^+ - \beta \geq 0$$

$$\sum_{n=1}^{N} t_n^- - \beta \geq 0$$

$$z_k \geq 0 \qquad k = 1, \cdots, K$$

where δ^* is the optimal solution to model (14). If $\beta > 0$, then the projected point ($x_{k_0} - \delta^* x_{k_0}, y_{k_o} + \delta^* y_{k_o}, b_{k_o} - \delta^* b_{k_o}$) of DMU k_0 has weak congestion. Otherwise, it is not congested.

In the following, we demonstrate the proposed approach with the simple example in Table 1. From Table 1, we can see that all the DMUs are efficient. We applied model (15) to the data set and the results are displayed in Table 2. From α, we identified DMU3 and DMU 4 as strongly congested. DMU1 and DMU2 have no congestion. For DMU 5, we further solve model (16) and $\beta^* = 0.5$. Hence, DMU5 is weakly congested.

Table 2 The data set for illustration

	x_1	x_2	y_1	y_2	b_1	b_2	inefficiency	α	Congestion
DMU1	1	3	2	8	3	2	0	0.1	No
DMU2	1	4	3	7	2.5	4	0	0.077	No
DMU3	2	6	1	6	5	35	0	-0.021	Strong
DMU4	3	6	1	5	10	10	0	-0.216	Strong
DMU5	1.5	3	1.5	8	3	3	0	0	Weak

Remark. From the inefficiency score by the direction distance function (3), all the DMUs are efficient. However, it is obvious that DMU 3 is dominated by DMU2 in inputs and outputs, which means that DMU 3 is misclassified as an efficient DMU. The contribution of our paper is that our proposed approach can further discriminate between the congested DMUs and the truly efficient DMUs. This finding shows that it is cautious for the decision-maker to use the directional distance function to evaluate the efficiency in the presence of desirable and undesirable outputs.

4. Illustrative examples

In this section, we first use a numerical example to make comparisons between our proposed approach and the existing approaches in the traditional scenario. According to the review in the introduction section, the congestion approaches for the traditional scenario only consider the desirable outputs. For comparison, we extend Färe et al. (1985)'s approach to include undesirable outputs by considering the following two models:

$$
\begin{aligned}
& h^* = \max \; h && (17)\\
& \sum_{k=1}^{K} z_k y_{km} \ge y_{k_o m} + h y_{k_o m} && m = 1,\cdots,M\\
& \sum_{k=1}^{K} z_k b_{ki} = b_{k_o i} - h b_{k_o i} && i = 1,\cdots I\\
& \sum_{k=1}^{K} z_k x_{kn} = x_{k_o n} && n = 1,\cdots N\\
& z_k \ge 0 && k = 1,\cdots,K
\end{aligned}
$$

and

$$
\begin{aligned}
& \tau^* = \max \; \tau && (18)\\
& \sum_{k=1}^{K} z_k y_{km} \ge y_{k_o m} + \tau y_{k_o m} && m = 1,\cdots,M\\
& \sum_{k=1}^{K} z_k b_{ki} = b_{k_o i} - \tau b_{k_o i} && i = 1,\cdots I\\
& \sum_{k=1}^{K} z_k x_{kn} \le x_{k_o n} && n = 1,\cdots N\\
& z_k \ge 0 && k = 1,\cdots,K
\end{aligned}
$$

Basically, model (17) and (18) differ only in that the constraints for inputs of model (17) are under weak disposability, but those of model (18) are under strong disposability. Similar to Färe et al. (1985), the congestion is defined as the ratio of $(1+\tau^*)/(1+h^*)$.

Similarly, we extend Sueyoshi and Sekitani (2009)'s approach to the undesirable output case by considering the following model:

$$\text{Max} \quad \varepsilon \tag{19}$$

$$\text{s.t.} \quad \sum_{m=1}^{M} y_{km}\pi_m^y - \sum_{i=1}^{I} b_{ki}\pi_i^b - \sum_{n=1}^{N} x_{kn}\pi_n^x + \sigma \le 0 \qquad k=1,\cdots,K$$

$$\sum_{m=1}^{M} y_{k_0m}\pi_m^y - \sum_{i=1}^{I} b_{k_0i}\pi_i^b - \sum_{n=1}^{N} x_{k_0n}\pi_n^x + \sigma = 0$$

$$\sum_{m=1}^{M} y_{k_0m}\pi_m^y + \sum_{i=1}^{I} b_{k_0i}\pi_i^b = 1$$

$$x_{k_0n}\pi_n^x - \varepsilon \ge 0 \qquad n=1,\cdots N$$

$$\pi_m^y \ge 0 \qquad m=1,\cdots,M$$

$$\pi_n^x \quad \text{unconstrained} \qquad n=1,\cdots N$$

$$\pi_i^b \quad \text{unconstrained} \qquad i=1,\cdots I$$

$$\sigma \quad \text{unconstrained}$$

According to Sueyoshi and Sekitani (2009), if the optimal objective value of model (19) is negative, then the k_0th DMU suffers from wide congestion.

In addition, we compare our approach with the following slacks-based approach that is widely used among the DEA models with undesirable outputs (Fukuyama and Weber,2010; Barros et al., 2012).

$$\text{Max} \quad \frac{1}{3}\left(\frac{1}{M}\sum_{m=1}^{M}\frac{s_m^y}{y_{k_0m}} + \frac{1}{I}\sum_{i=1}^{I}\frac{s_i^b}{b_{k_0i}} + \frac{1}{N}\sum_{n=1}^{N}\frac{s_n^x}{x_{k_0n}}\right) \tag{20}$$

$$\sum_{k=1}^{K} z_k y_{km} = y_{k_o m} + s_m^y \qquad m = 1, \cdots, M$$

$$\sum_{k=1}^{K} z_k b_{ki} = b_{k_o i} - s_i^b \qquad i = 1, \cdots, I$$

$$\sum_{k=1}^{K} z_k x_{kn} = x_{k_o n} - s_n^x \qquad n = 1, \cdots, N$$

$$s_m^y \geq 0 \qquad m = 1, \cdots, M$$

$$s_i^b \geq 0 \qquad i = 1, \cdots, I$$

$$s_n^x \geq 0 \qquad n = 1, \cdots, N$$

$$z_k \geq 0 \qquad k = 1, \cdots, K$$

4.1 Numerical example

In this subsection, we use the example in Table 1 to compare our proposed approach with the above modified three approaches. Table 3 summarizes the status of congestion measured by four different approaches.

Table 3 Congestion by four different approaches

DMU	Färe et al. (1985) (Modified)	Sueyoshi and Sekitani (Modified)	Slacks-based approach	Proposed approach
DMU1	No	No	No	No
DMU2	No	No	No	No
DMU3	No	No	Weak	Strong
DMU4	Congestion	No	Weak	Strong
DMU5	No	No	Weak	Weak

From Table 3, we observe that the modified Sueyoshi and Sekitani (2009)'s approach (19) can't identify the congestion. The reason is that the shadow prices for the inputs and undesirable outputs are unconstrained, which indicates that the modified version of Sueyoshi and Sekitani (2009)'s approach is be inappropriate for the case with undesirable outputs. The modified version of Färe et al. (1985)'s approach produces a different result regarding congestion from the other two approaches. For example, the modified Färe et al. (1985)'s approach only identify the congestion on the fourth DMU. In contrast, the other two approaches have the same results that

DMU 3, DMU 4 and DMU 5 are under congestion, which indicates that the slacks-based approach and our proposed approach are more sensitive.

According to the inefficiency scores in the last column in Table 1, all DMUs are efficient by the direction distance function (3). Both the slacks-based approach and our proposed approach can discriminate between the congested DMUs and the truly efficient DMUs. However, the approach proposed in this study has the advantage over the slacks-based approach that has the separation capability between strong and weak congestion. The slacks-based approach always identifies the DMU as the weak congestion if the DMU is under congestion. The reason is that at least one s_n^x ($n = 1, \cdots, N$) is zero.

4.2 Empirical example

In this section, to illustrate the use of the proposed approach, we analyse an empirical dataset consisting of 20 power plants from Sueyoshi and Goto (2012a). The data set consists of two inputs (nameplate capacity and fuel consumption), one desirable output (net generation) and three undesirable outputs (SO2, NOx and CO2). Table 4 lists the input and output dataset for 20 power plants.

Table 4 Data set of U.S. power plants

Plant	Nameplate capacity	Fuel consumption	Net generation	SO_2	NO_x	CO_2
	MW	1000 MMBtu	GWh	Ton	Ton	1000 Ton
1	2842	122482	14087	37109.314	8440.868	11424.864
2	138	3855	261	3925.846	795.335	365.722
3	1417	52698	5207	5027.653	5373.364	5695.994
4	1969	51648	4688	12861.358	3526.372	5278.828
5	721	51927	4465	11720.507	4589.779	5265.576
6	538	32907	3026	5644.99	3193.01	3496.782
7	110	6510	592	2679.396	1273.507	766.309
8	257	20258	1836	1308.302	3607.039	2164.309
9	349	27868	2656	1948.292	4355.391	2892.501

10	1129	74881	7236	7955.926	9211.297	8441.262
11	1207	13523	1521	2.933	103.511	580.963
12	85	61	4	0.093	0.91	3.162
13	559	7929	688	642.152	1010.558	580.239
14	2390	146042	14620	7156.835	5903.727	15588.088
15	1429	70867	7097	20576.961	4401.534	7687.792
16	574	5589	722	1.65	30.376	326.848
17	1010	75739	6840	18012.943	14346.412	8509.161
18	138	64	2	0.02	19.904	3.849
19	183	676	45	0.199	233.665	39.33
20	752	1559	132	0.425	191.427	84.14

We use the model (15) to evaluate the efficiency and identify the congestion of 20 plants. The results of the inefficiency scores (δ^*) are reported in the second column of Table 5. As observed from Table 5, 15 plants are efficient. As presented in Table 5, congestion can be identified by α , which are shown in the third column of Table 5 . The results show that only five of the 15 efficient plants are truly efficient, and the other 10 plants have strong congestion. For the five inefficient plants, we make a projection according to equation (4), and all the projected points are identified as strongly congested.

The above case demonstrates that it calls for a careful inspection to use the directional distance function in the actual efficiency analysis in the presence of undesirable outputs. Through our proposed approach we can further discriminate the congested DMUs from the truly efficient DMUs.

Table 5 Inefficiency scores and congestion status

Plant	δ^*	α	Congestion
1	0	0.2362745E-05	No
2	0	-0.1060751E-03	Strong
3	0.0037	-0.5917792E-04	Strong

4	0.078	-0.9468108E-04	Strong
5	0	-0.4584385E-05	Strong
6	0.0315	-0.4205287E-04	Strong
7	0	-0.3299549E-04	Strong
8	0	-0.1266256E-04	Strong
9	0	0.1957027E-04	No
10	0	-0.1894897E-04	Strong
11	0	0.7273711E-03	No
12	0.2948	-0.1614217E-01	Strong
13	0.0292	-0.3471911E-03	Strong
14	0	0.1454874E-04	No
15	0	-0.6217529E-05	Strong
16	0	0.1214732E-02	No
17	0	-0.2381698E-05	Strong
18	0	-5.451708	Strong
19	0	-0.2034579E-02	Strong
20	0	-4.479254	Strong

5. Conclusion

Congestion is a widely observed economic phenomenon. Many previous approaches to identify congestion in nonparametric analysis only consider desirable outputs. In the actual production process, undesirable outputs are usually jointly produced with desirable outputs. Chung et al. (1997) developed the directional distance function model, which expands the desirable outputs and contracts inputs and undesirable outputs along the direction vector path.

However, different from the traditional circumstance with desirable outputs only, we find that even if a DMU is efficient by the directional distance function, it maybe evidences congestion when considering both desirable outputs and undesirable outputs.This paper makes a novel attempt to propose an approach for measuring congestion in the presence of desirable and undesirable outputs simultaneously. Through the proposed approach, we can discriminate between the congested DMUs and the truly efficient DMUs, which are all efficient according to the scores calculated by the directional distance function. Furthermore, we compare our proposed approach with the existing three representative approaches with a numerical example. The results indicate that the proposed approach is capable of identifying strong and weak congestion in the presence of desirable outputs and undesirable outputs.

For future extensions, how to develop the new model to evaluate their efficiencies for the congested DMUs will be an important research issue. In addition, the congestion and returns to scale are closely connected to each other (Sueyoshi and Sekitani, 2009). To explore the economic scale concepts for dealing with undesirable outputs, Toshiyuki and Goto (2012b, 2013) proposed the concept of damages to scale. Thus, a theoretical linkage between congestion, returns to scale and damages to scale is an interesting topic for future research. We note that all variables are assumed to be continuous and discretionary in this paper. How to dispose the non-discretionary and integer variables is another issue for future study.

References

Barros, C.P., Managi, S., Matousek, R., (2012). The technical efficiency of the Japanese banks: Non-radial directional performance measurement with undesirable output. Omega,40,1-8.

Brockett, P., Cooper, W.W., Deng, H., Golden, L., Ruefli, T.W., (2004). Using DEA to identify and manage congestion. Journal Productivity Analysis, 22, 207-226.

Cherchye, L., Kuosmanen, T., Post, T., (2001). Alternative treatments of congestion in DEA—a rejoinder to Cooper, Gu and Li. European Journal of Operational Research, 132, 75–80.

Chung, Y., R. Färe, S. Grosskopf. (1997). Productivity and undesirable outputs: a directional distance function approach. Journal of Environmental Management. 51,229–40.

Cooper, W.W., Gu, B.S., Li, S.L., (2001a). Comparisons and evaluations of alternative approaches to the treatment of congestion in DEA. European Journal of Operational Research132,62-74.

Cooper, W.W., Gu, B.S., Li, S.L., (2001b).Note: Alternative treatments of congestion in DEA- a response to the Cherchye, Kuosmanen and Post critique. European Journal of Operational Research132,81-87.

Cooper, W.W., Seiford, L.M., Zhu, J., (2000). A unified additive model approach for evaluating inefficiency and congestion with associated measures in DEA. Socio-Economic Planning Sciences, 34,1-25.

Cooper, W.W., Thompson, R.G., Thrall, R.M., (1996). Introduction: Extensions and new developments in DEA. Annals of Operations Research ,66, 3-46.

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

Färe, R., Grosskopf, S., Lovell, C.A.K., (1985). The measurement of efficiency of production Norwell, MA: Kluwer-Nijihoff.

Färe, R., Grosskopf, S., Lovell, C.A.K., Pasurka , C.,(1989). Multilateral productivity comparisons when some outputs are undesirable: a nonparametric approach. The Review of Economics and Statistics,71,90-98.

Färe, R., Grosskopf, S., (2000). Slacks and congestion: A response. Socio-Economic Planning Sciences 34, 35–50.

Färe, R., Grosskopf, S., (2003). Non-parametric productivity analysis with undesirable outputs: comment. American Journal of Agricultural Economics, 85 (4) 1070–1074.

Färe, R., Grosskopf, S., (2004). Modeling undesirable factors in efficiency evaluation: comment. European Journal of Operational Research , 157, 242–245.

Färe, R., Grosskopf, S., (2009). A comment on weak disposability in nonparametric production analysis. American Journal of Agricultural Economics, 91(2) 535–538.

Fukuyama,H., Weber, W.L.,(2010). A slacks-based inefficiency measure for a two-stage system with bad outputs. Omega 38,398-409

Kao, C., (2010). Congestion measurement and elimination under the framework of data envelopment analysis. International Journal of production economics, 123, 257–265

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 and Industrial Engineering, 64, 731 – 738.

Kuosmanen, T., (2005). Weak disposability in nonparametric production analysis with undesirable outputs. American Journal of Agricultural Economics,87 (4) 1077–1082.

Kuosmanen, T., Kortelainen, M.,(2005). Measuring Eco-efficiency of Production with Data Envelopment Analysis. Journal of Industrial Ecology, 9, 59-72.

Kuosmanen, T., Podinovski, V., (2009). Weak disposability in nonparametric production analysis: reply to Färe and Grosskopf . American Journal of Agricultural Economics,87 (4), 539–545.

Kuosmanen, T., R.K. Matin. 2011. Duality of weakly disposable technology. Omega. 87(4) 504–512.

Picazo-Tadeo, A. J., Beltrán-Esteve, M., Gómez-Limón, J.A., (2012). Assessing eco-efficiency with directional distance functions. European Journal of Operational Research ,220,798–809.

Sueyoshi, T., Sekitani, K., (2009). Congestion and returns to scale under an occurrence of multiple optimal projections. European Journal of Operational Research, 194, 592–607.

Sueyoshi, T., Goto, M., (2012a). DEA radial measurement for environmental assessment and planning: Desirable procedures to evaluate fossil fuel power plants. Energy Policy, 41, 422-432.

Sueyoshi, T., Goto, M., (2012b). Returns to scale and damages to scale under natural and managerial disposability: Strategy, efficiency and competitiveness of petroleum firms. Energy Economics, 34, 645-622

Sueyoshi, T., Goto, M., (2013). Returns to scale vs. damages to scale in data envelopment analysis: An impact of U.S. clean air act on coal-fired power plants. Omega, 41, 164-175.

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

Wei, Q.L., Yan.H., (2004). Congestion and returns to scale in data envelopment analysis. European Journal of Operational Research,153, 641-660.

Weber, W.L., Domazlicky,B., (2001). Productivity growth and pollution in state manufacturing. Review of Economics and Statistics, 83 , 195–199.

Zhou, P., Ang, B.W., Poh, K.L.,(2008). A survey of data envelopment analysis in energy and environmental studies. European Journal of Operational Research,189 1–18.