E-ISSN: 2717-2937 | P-ISSN: 2783-1337

International Journal of Research in Industrial Engineering

www.riejournal.com

Int. J. Res. Ind. Eng. Vol. 12, No. 1 (2023) 53–72.

Paper Type: Research Paper

# Identification and Evaluation of Congestion in Two-Stage Network Data Envelopment Analysis

**Sarvar Sadat Kassaei[1], Farhad Hosseinzadeh Lotfi[1,*], Alireza Amirteimoori[2], Mohsen Rostamy Malkhalifeh[1], Bijan Rahmani[3]**

[1] Department of Mathematics, Science and Research Branch, Islamic Azad University, Tehran, Iran; sarvar.kassaei@yahoo.com; hosseinzadeh_lotfi@yahoo.com; mohsen_rostamy@yahoo.com.

[2] Department of Applied Mathematics, Islamic Azad University, Rasht, Iran; teimoori@guilan.ac.ir.

[3] Department of Mathematics, Central Tehran Branch, Islamic Azad University, Tehran, Iran; bijanrah40@gmail.com.

**Citation:**

Kassaei, S. S., Hosseinzadeh Lotfi, F., Amirteimoori, A., Rostamy Malkhalifeh, M., & Rahmani, B. (2023). Identification and evaluation of congestion in two-stage network data envelopment analysis. *International journal of research in industrial engineering, 12*(1), 53-72.

Received: 28/01/2022 Reviewed: 01/03/2022 Revised: 09/04/2022 Accepted: 18/05/2022

## Abstract

Congestion is one of the important concepts in data envelopment analysis that occurs when excessive inputs reduce the maximally possible outputs. Identification and elimination of congestion have a significant impact on reducing inputs along with increasing outputs. Hence, various studies have been conducted to detect and evaluate congestion. However, in today's world, no organization can achieve its final output with just one process of input. In other words, today's organizations have a network structure that consists of several subsections. Ignoring the existing influences among the subsections processes may lead to inadequate or even incorrect results for evaluating the congestion. While all of the existing methods only evaluate the congestion of each subsection or the whole system independently. Therefore, in this paper, the concept of congestion is developed for a specific and so practical case of network structure called "two-stage network structure". This case of network structure consists of two series stage such that Stage 1 consume some primary inputs to produce some intermediate outputs. In the following, the intermediate outputs are used as the inputs of Stage 2 to produce the final output. Here, the concept of congestion is defined for systems with a two-stage structure. Then, to examine the congestion of each stage as well as the congestion of the whole system, a single linear programming model is proposed. The validity of the proposed model is investigated using several theorems and it is shown that the new definition is a generalization of the previous definitions of the congestion for the black-box systems. Finally, the proposed model is applied to a case study including 24 non-life insurance companies in Taiwan.

**Keywords:** Network data envelopment analysis, Two-stage network structure, Intermediate output, Congestion.

## 1 | Introduction



Congestion is an important concept in data envelopment analysis that can be effective in improving the efficiency of decision-making units. Evaluating the efficiency of homogenous decision-making units is the primary goal of data envelopment analysis models. The efficiency of these units is affected by the number of resources they consume. On the other hand, an increase in input is often accompanied by an increase in output. But this is not the case in all production technologies. In other words, sometimes increasing the input sources leads to decreasing the output, and that is the concept of congestion. For example, increasing the number of workers in a mine causes them to collide with each other, thus reducing output.

Input congestion is a special case of inefficiency in which reducing at least one of the input components (without increasing the other input components) can increase at least one of the output

Corresponding Author: hosseinzadeh_lotfi@yahoo.com

https://doi.org/10.22105/riej.2023.326879.1294

components (without decreasing the other output components). It should be noted that this type of inefficiency or technical inefficiency is already known and different methods and models in data envelopment analysis can be used to distinguish it from other forms of inefficiency for example [1] Identifying, evaluating, and eliminating congestion can lead to useful results in real-world applications, such as cost reduction and increased output.

The concept of congestion was first investigated by Färe and Svensson [2], and then, its more complete version was presented by Färe and Grosskopf [3] and Färe et al. [4]. Their model calculates the congestion impact as a ratio of the observed amounts to the expected amounts. However, it only detects the presence or absence of congestion and may sometimes reach incorrect results. In another study, Cooper et al. [5] proposed a slack-based approach to calculate the congestion impact as the difference between the observed amounts and the expected amounts. This approach has some strong points to the previous method. It determines the congested inputs and provides the amount of congestion in each input. In the following, Cooper et al. [6] expanded a unified additive model using additive models for determining congestion. Later, Cooper et al. [1] proposed a single model to assess congestion without solving two models similar to previous methods. Brockett et al. [7] developed a model known as the "BCSW" based on the proposed method by Cooper et al. [6]. By determining input congestion, Jahanshahloo and Khodabakhshi [8] introduced a suitable combination of inputs for improving outputs. Moreover, other studies by Wei and Yan [9] simultaneously studied congestion and different types of returns to scale using the DEA output-oriented models. They also recognize the necessary and sufficient conditions for the evidence of congestion and different kinds of returns to scale. Sueyoshi and Sekitani [10] proposed a method that can detect congestion in the presence of multiple solutions.

In general, studies on congestion in the DEA framework are important in two aspects. These two aspects are the identification and measurement of congestion. Congestion identification methods have been in development for nearly 4 decades. The first study in this field was conducted by Färe and Svensson [2] in which they explained that congestion occurs when increasing some of the inputs may decrease the outputs. In the following, Färe and Grosskopf [3] pointed out that the reason for the neglect of congestion may be attributed to the assumption that isoquants have no positive slope anywhere. Accordingly, Färe et al. [4] broke this assumption by creating a positive slope isoquants assuming weak disposability of inputs. They proposed the FGL method for congestion identification. Cooper et al. [11] provided a numerical example to show that the FGL method cannot identify congestion well. Due to the shortcomings of the FGL method, Cooper et al. [5] proposed the CTT method based on the strong disposability assumption of inputs. Then, Wei and Yan [12] and Tone and Sahoo [13] proposed the WY-TS method to identify congestion. This method constructs a Production Possibility Set (PPS) for identifying congestion by removing the disposability assumption of inputs. Noura et al. [14] proposed a new method to reduce the computational effort required for calculating congestion. Sueyoshi and Sekitani [15] pointed out that the WY-TS method did not consider the existence of multiple optimal solutions, and this issue may reduce the stability of the WY-TS method in identifying congestion. Mehdiloozad et al. [16] proposed the MZS method to improve this shortcoming of the WY-TS method. On the other hand, from a measurement point of view, it should be noted that congestion refers to the reduction in outputs caused by excessive inputs. Therefore, both the excessive inputs and reduction in outputs can reflect the measure of congestion. For example, the FGL method proposed by Färe et al. [4] uses the amount of excessive input as an indicator to measure the degree of congestion; the WY-TS method proposed by Wei and Yan [12] and Tone and Sahoo [13] uses the output reduction caused by congestion to measure the degree of congestion.

In recent years, various studies have been conducted to develop the concept of congestion in data envelopment analysis. For example, Karimi et al. [17] proposed a method that evaluates the congestion in integer-valued DEA. Khoveyni et al. [18] investigated the recognition of congestion in the presence of negative data and specified the least and the most congested DMUs. Shabanpour et al. [19] proposed a novel DEA model to show that an increase in congested inputs may lead to higher outputs/efficiency. Moreover, they used the concept of input congestion as a tool for ranking decision-making units. Adimi et al. [20] introduced the concept of congestion hyperplane without considering the efficiency value. Salehi et al. [21] presented a new method to identify congestion based on the definition of congestion. Xian-ton Ren et al. [22] innovatively tried to eliminate congestion by increasing inputs in the case of R&D activities of Chinese universities. They also analyzed the relationship between congestion and overinvestment. Navidi et al. [23] represented the method that measures congestion without solving a model. Yang et al. [24] proposed the concept of directional congestion in the framework of data envelopment analysis. Velázquez and Benita [25] investigated the patterns and dynamics of efficiency, productivity, and technological change of the automotive sector in Mexico. Cho and Yang [26] developed a new method for congestion analysis based on the slacks-based measure approach that keeps a close link between undesirable outputs, desirable outputs, and inputs. According to the S-shape form of the production function and concerning the geometric features of the anchor point, Shadab et al. [27] developed an algorithm by the connection between the anchor points and congestion definition. Khoshroo et al. [28] used the Bounded Adjusted Measure (BAM) to improve the efficiency of tomato production as well as decrease the carbon footprint. Fallahnejad et al. [29] considered the effect of input congestion on cost inefficiency and presented a new decomposition of cost efficiency and observed cost versus optimal cost.

In addition to the aforementioned studies, there are various methods to recognize the evidence of congestion. But the remarkable point in all of these methods is the lack of attention to the internal structure of the decision-making units. In other words, existing methods identify and evaluate congestion without considering the internal structure or the existing communication between the stages. However, nowadays, the complexity of goods or services is such that few organizations or institutions can produce products alone and without cooperation with other organizations. This disregard for the influences between the stages may lead to inadequate or even incorrect results.

Here, a real example is provided for further explanation. Suppose we are looking to evaluate a factory that produces wood products such as tables, chairs, sofas, etc. The factory uses raw materials such as human resources, costs, and raw wood. The factory also has a fixed area for equipment installation and storing of raw materials and production. In the first stage of this factory, the input woods are converted into different types needed for construction in the next stage. This stage includes the cutting and processing work required for the next stage. In the second stage, by using the products of the first stage (e.g. table, chair, sofa) and the specialized human resources, the final products are provided. Now, if this factory has too much manpower or raw materials for wood, this will cause congestion. Because the presence of too much manpower causes confusion and disruption in the work and the presence of too much raw wood occupies the space needed to perform other activities. The same is true for the wood produced in the first stage. Overproduction of processed wood in the first stage may reduce final production by occupying the workspace and disrupting other tasks.

The aforementioned example is a sample of a system with a two-stage network structure in which the outputs of the first stage (called intermediate products) appear in the role of inputs of the second stage. Many practical problems can be modeled in today's world according to the two-stage network structures. For example, Zuo et al. [30] used a two-stage DEA model to construct indicators to measure Chinese provinces Mining Technological Innovation Efficiency (MTIE), Mining Eco-Efficiency (MEE), and Mining Comprehensive Efficiency (MCE). Silva et al. [31] addressed how socioeconomic conditions influence entrepreneurship-based activities in 18 European countries grouped into subregions (North, South, East, and West) during the period 2008–2018. They conducted their empirical study under a two-stage DEA model. Jingxin et al. [32] constructed a two-stage DEA model to measure Urban Construction Land Use (UCLU) efficiency. Henriques et al. [33] analyzed 59 papers and divided them into ten classes that cover various perspectives of two-stage DEA studies, such as the economic context, geographic region of the banking units, methodological characteristics, and type of the models, either internal or external. Chen et al. [34] proposed an extended two-stage network DEA approach for measuring the operating efficiency of 52 Chinese universities. Izadikhah et al. [35] developed a novel fuzzy chance-constrained two-stage data envelopment analysis model. Mozaffari and Ostovan [36] presented a two-stage supply chain with random data and the CRA model with ratio data used to calculate the projection of DMUs. Marzband

Kassaei et al.| Int. J. Res. Ind. Eng. 12(1) (2023) 53-72

[37] investigated the efficiency of supply chains in manufacturing and industrial companies. Nematizadeh and Nematizadeh [38] introduced a two-stage feedback structure including undesirable factors. Then, by applying the assumption of weak disposability for undesirable factors, they provided a method for analyzing the relative performance of such network structures.

To the best of our knowledge, all of the existing methods only evaluate the congestion of each stage or the whole system independently. GholamAzad and Pourmahmoud [39] have proposed a new method for measuring congestion in stages of the network. However, this method also does not examine the congestion in the whole system. Moreover, they have not theoretically discussed the proposed model. Accordingly, the current paper tries to develop the existing methods to identify and evaluate the congestion in decision-making units with a series two-stage network structure. To this end, a single linear model is presented that examines the possible congestion in each of these stages and the whole system by considering the congestion relationship between the stages and the whole system. The proposed model can be considered an extension of Cooper et al. [1] one-model method. Therefore, the main contribution of this paper is proposing the first valid method to identify and evaluate the congestion in the two-stage network structure.

The rest of the paper is organized as follows: Section 2 contains the required concepts and definitions, along with a brief description of the one-model method proposed by Cooper et al. [1]. In Section 3, the new model is proposed to detect and evaluate the congestion of the DMU with the series two-stage network structure. Section 4 provides a small numerical example to show the efficiency of the proposed model. Moreover, a case study of Taiwanese non-life insurance companies is conducted in this section. Finally, Section 5 concludes this research.

## 2 | Background

In this section, some required concepts and definitions of data envelopment analysis and network data envelopment analysis are presented. The first subsection includes the managerial implications and the second subsection presents the model proposed by Cooper et al. [1] to detect and evaluate the congestion in the black-box view.

### 2.1 | Managerial Implications

DEA is a non-parametric methodology for measuring the efficiency of homogenous DMUs that use multiple inputs to produce multiple outputs. On the other hand, Productivity is one of the basic concepts in management, which includes efficiency and effectiveness. There are various definitions of efficiency, effectiveness, and productivity. For example, Pritchard [40] illustrated three definitions related to productivity as follows:

I. Productivity is output/input or in other words, is a measure of efficiency.
II. Productivity is a composition of effectiveness and efficiency.
III. Productivity is referred to the broader concept that whatever makes the organization has a better function.

What is certain, however, is that productivity can be considered a function of efficiency and effectiveness. Where efficiency is interpreted as “Doing things right” and effectiveness is interpreted as “doing the right things”. The relationship between these three concepts can be seen in *Fig. 1*:

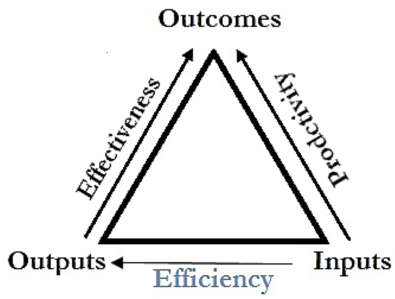


**Fig. 1. The relationship between efficiency, effectiveness, and productivity.**

Indeed, each DMU consumes the inputs to produce the outputs to achieve the outcomes. Meanwhile, the manager of DMU tries to produce maximum outputs by consuming minimum inputs. With these explanations, a DMU is productive if it works efficiently and if its master DMU has been planned effectively [41]. Therefore, the efficiency of the DMUs is evaluated to quantify their contribution to productivity.

To measure the efficiency of DMUs, DEA constructs the PPS by using some postulates. Then, the efficiency values of the DMUs are calculated relative to the frontier of PPS.

**Definition 1.** A DMU is efficient if no other DMU in PPS produces the same or more output by using the same or less input. In this way, technical inefficiency occurs when some input or output components can be improved without worsening other components.

In the meantime, a special type of inefficiency should be given more attention. This type of inefficiency is known as congestion. Identification and elimination of congestion have a significant impact on reducing inputs along with increasing outputs.

**Definition 2.** Input congestion occurs when an increase in one/more input components results in a decrease in one/more output components (without improving other input and output components); or conversely, a decrease in one/more input components results in an increase in one/more output components (without worsening other input and output components) [1].

Remark 1: although congestion can be considered a special case of technical inefficiency, it should be noted that congestion appears and is discussed in the DMUs in which the principle of input possibility is not established. In other words, by increasing the inputs of these DMUs, one cannot find a unit in the set that produces the same output.  On the other hand, some technically inefficient DMUs (which do not meet this requirement and, of course, do not have congestion) may be true in *Definition 2* and be mistaken for congestion. Therefore, to prevent such errors, it is better to modify *Definition 2* For this purpose, it can be stated in *Definition 2* that there should be no DMU with more input/output related to the DMU under evaluation.

**Definition 3.** As mentioned, the concept of input congestion is discussed on the PPSs that lack the principle of input possibility. This set includes all DMUs that can produce the output vector $Y$ by using the input vector $X$, considering the four principles: inclusion of observations, convexity[1], output possibility[2], and the principle of minimum interpolation. Such a production possibility set for $n$ observed decision-making units, i.e. $DMU_j = (X_j, Y_j),\ j = 1,\dots,n$, will be as the *Relation (1):*

---

[1] The convexity principle states that any weighted average (convex combination) of feasible production plans is feasible as well [42].

[2] According to the output possibility principle, if a unit (x, y) belongs to PPS, then any semi-positive $(x, \bar{y})$ with $y \le \bar{y}$ is included in PPS.

$$T=\left\{(X,Y)\mid \sum_{j=1}^{n}\lambda_j X_j = X, \sum_{j=1}^{n}\lambda_j Y_j \ge Y, \sum_{j=1}^{n}\lambda_j = 1, \ \lambda_j \ge 0 \ , j=1,\ldots,n\right\}. \tag{1}$$

**Definition 4.** The decision-making unit with a series two-stage network structure refers to a unit that itself consists of two stages, according to *Fig. 2*. In such a unit that produces the output vector $Y$ using the input vector $X$, there are some intermediate productions called the vector $Z$ that is the output vector of the first stage. In other words, in such units, to generate the final output $Y$, the main input vector $X$ is first provided to the first stage to generate the output $Z$. Then, the output vector of the first stage is given to the second stage as the input vector to generate the final output vector $Y$.

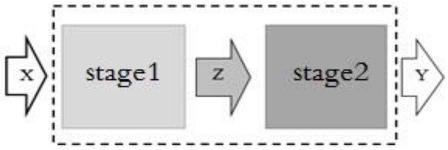

**Fig. 2. DMU with two-stage series network structure.**

The PPS for $n$ observed DMUs, i.e. $DMU_j=(X_j,Z_j,Y_j)$, with the series two-stage network structure, is as the *Relation (2)*:

$$T=\left\{(X,Z,Y)\ \middle|\ \begin{array}{l}\sum_{j=1}^{n}\lambda_j X_j \le X, \sum_{j=1}^{n}\lambda_j Z_j \ge Z, \sum_{j=1}^{n}\mu_j Z_j \le Z, \sum_{j=1}^{n}\mu_j Y_j \ge Y, \\ ,\sum_{j=1}^{n}\lambda_j = 1, \sum_{j=1}^{n}\mu_j = 1, \lambda_j \ge 0, \mu_j \ge 0 \ j=1,\ldots,n\end{array}\right\}. \tag{2}$$

Note that the PPS in *Relation (2)* is constructed by considering the principles of input and output possibility. By removing the principle of input possibility, the PPS is changed as the *Relation (3):*

$$T=\left\{(X,Z,Y)\ \middle|\ \begin{array}{l}\sum_{j=1}^{n}\lambda_j X_j = X, \sum_{j=1}^{n}\lambda_j Z_j \ge Z, \sum_{j=1}^{n}\mu_j Z_j = Z, \sum_{j=1}^{n}\mu_j Y_j \ge Y, \\ ,\sum_{j=1}^{n}\lambda_j = 1, \sum_{j=1}^{n}\mu_j = 1, \lambda_j \ge 0, \mu_j \ge 0 \ j=1,\ldots,n\end{array}\right\}. \tag{3}$$

## 2.2 | Detection and Evaluation of Congestion in Black-Box View

As mentioned in the first section, several methods exist to identify and evaluate congestion. In the following, the proposed method by Cooper et al. [1] is briefly reviewed to detect and evaluate the congestion. They developed a one-model method by using the BCSW method [7] and combining its two models. To better explain, suppose that $X_j=(x_{1j},\ldots,x_{ij},\ldots,x_{mj})^T$ and $Y_j=(y_{1j},\ldots,y_{rj},\ldots,y_{sj})^T$ are the input and output vectors of DMU$_j$ (j=1,…,n), respectively. Then, the proposed model by Cooper et al. [1] to detect and evaluate the congestion of DMU$_O$ is as the *Model (4)*:

$$\begin{aligned}
\text{Max} \quad & \varphi+\varepsilon\left(\sum_{r=1}^{s}s_r^{+}-\varepsilon\sum_{i=1}^{m}s_i^{-c}\right),\\
\text{s.t.} \quad & \sum_{j=1}^{n}\lambda_j x_{ij}+s_i^{-c}=x_{io}, \quad i=1,\ldots,m,\\
& \sum_{j=1}^{n}\lambda_j y_{rj}-s_r^{+}=\varphi y_{ro}, \quad r=1,\ldots,s,\\
& \sum_{j=1}^{n}\lambda_j=1, \ \lambda_j\ge 0, \quad j=1,\ldots,n,\\
& s_i^{-c}, s_r^{+}\ge 0, \quad i=1,\ldots,m, r=1,\ldots,s.
\end{aligned} \tag{4}$$

Where $\varepsilon$ is a small positive non-Archimedean value. Note that the presence of $\varepsilon$ in *Model (4)* indicates the optimizing priority of the variables in the objective function. In fact, in this model, at the first, the variable $\varphi$ is maximized, and then the variables $s_r^+ (r=1,...,s)$ and $s_i^{-c} (i=1,...,m)$ are maximized and minimized, respectively.

**Theorem 1.** $DMU_o$ has input congestion when the optimal solution of *Model (4)*, i.e., $(\varphi^*, \lambda^*, S^{-c*}, S^{+*})$, holds at least one of the following conditions [1]:

- $\varphi^* > 1 \text{ and } \sum_{i=1}^{m} s_i^{-c*} > 0.$
- $\sum_{r=1}^{s} s_r^{+*} > 0 \text{ and } \sum_{i=1}^{m} s_i^{-c*} > 0.$

Accordingly, the amount of congestion in i-th input will be determined by the value of $s_i^{-c*}$.

**Theorem 2.** Suppose that $(\varphi^*, \lambda^*, S^{-c*}, S^{+*})$ is an optimal solution of *Model (4)*. Then, $DMU_o$ is inefficient if at least one of the following conditions occurs [1]:

I. $\varphi^* > 1.$

II. $\sum_{r=1}^{s} s_r^{+*} > 0$ .

III. $\sum_{i=1}^{m} s_i^{-c*} > 0.$

Conversely, if $\varphi^* = 1$, $\sum_{r=1}^{s} s_r^{+*} = 0$ and $\sum_{i=1}^{m} s_i^{-c*} = 0$, then $DMU_o$ is on the (efficient or inefficient) frontier of PPS defined in the *Relation (1)*.

**Lemma 1 (Identification and evaluation of congestion in series two-stage network structure).** Consider n observed decision-making units $DMU_j = (X_j, Z_j, Y_j)\ (j=1,...,n)$, with the series two-stage network structure, where $X_j = (x_{1j},...,x_{ij},...,x_{mj})^T$, $Z_j = (z_{1j},...,z_{tj},...,z_{pj})^T$ and $Y_j = (y_{1j},...,y_{rj},...,y_{sj})^T$. Suppose the aim is to identify and evaluate the congestion of each stage and $DMU_o$ as a whole, according to the proposed approach by Cooper et al. [1]. Therefore, the congestion assessment models for the first stage, the second stage, and the whole unit can be written as *Models (5)*, *(6)*, and *(7)*, respectively:

$$
\begin{aligned}
\text{Max} \quad & \varphi_1 + \varepsilon\left(\sum_{t=1}^{p} d_{1,t}^{+} - \varepsilon\sum_{i=1}^{m} s_{1,i}^{-}\right), \\
\text{s.t.} \quad & \sum_{j=1}^{n} \lambda_j x_{ij} + s_{1,i}^{-} = x_{io}, \quad i=1,...,m, \\
& \sum_{j=1}^{n} \lambda_j z_{tj} - d_{1,t}^{+} = \varphi_1 z_{to}, \quad t=1,...,p, \\
& \sum_{j=1}^{n} \lambda_j = 1,\ \lambda_j \geq 0, \quad j=1,...,n, \\
& s_{1,i}^{-}, d_{1,t}^{+} \geq 0, \quad i=1,...,m, t=1,...,p.
\end{aligned}
\tag{5}
$$

$$
\begin{aligned}
\text{Max} \quad & \varphi_2 + \varepsilon\left(\sum_{r=1}^{s} s_{2,r}^{+} - \varepsilon\sum_{t=1}^{p} d_{2,t}^{-}\right), \\
\text{s.t.} \quad & \sum_{j=1}^{n} \mu_j z_{tj} + d_{2,t}^{-} = z_{to}, \quad t=1,\dots,p, \\
& \sum_{j=1}^{n} \mu_j y_{rj} - s_{2,r}^{+} = \varphi_2 y_{ro}, \quad r=1,\dots,s, \\
& \sum_{j=1}^{n} \mu_j = 1,\ \mu_j \ge 0, \quad j=1,\dots,n, \\
& s_{2,r}^{+}, d_{2,t}^{-} \ge 0, \quad r=1,\dots,s, t=1,\dots,p.
\end{aligned} \tag{6}
$$

*Models (5)* and *(6)* detect the existing congestion of the first and second inputs separately, respectively. It should be noted that according to the design of the objective function in both models, three steps must be taken to solve these models. In the first step, the output-oriented radial image of the stage under evaluation is identified on the PPS. In the second step, the maximum non-radial improvement of all outputs is suggested on PPS. In the last step, the amount of congestion in each of the inputs is calculated. It should be noted that *Model (5)* corresponds to the PPS made by the first stage of DMUs but the PPS corresponding to *Model (6)* is made by the second stage of DMUs.

According to a similar idea, *Model (7)* can be proposed to calculate the congestion at the inputs of the network structure. This model similarly first specifies the output-oriented radial image of the network on the PPS. Note that here the PPS is constructed by the networks in the form of a black box, regardless of the intermediate productions. Then, the maximum non-radial improvement of all final outputs is suggested on PPS and at the end, the congestion is calculated in each of the primary inputs.

$$
\begin{aligned}
\text{Max} \quad & \varphi + \varepsilon\left(\sum_{r=1}^{s} s_{r}^{+} - \varepsilon\sum_{i=1}^{m} s_{i}^{-c}\right), \\
\text{s.t.} \quad & \sum_{j=1}^{n} \gamma_j x_{ij} + s_{i}^{-c} = x_{io}, \quad i=1,\dots,m, \\
& \sum_{j=1}^{n} \gamma_j y_{rj} - s_{r}^{+} = \varphi y_{ro}, \quad r=1,\dots,s, \\
& \sum_{j=1}^{n} \gamma_j = 1,\ \gamma_j \ge 0\ , \quad j=1,\dots,n, \\
& s_{i}^{-c}, s_{r}^{+} \ge 0, \quad i=1,\dots,m, r=1,\dots,s.
\end{aligned} \tag{7}
$$

Now according to *Models (5)*, *(6)*, and *(7)* and *Theorem 1*, the congestion in the first and second stages of $DMU_o$ along with the congestion of $DMU_o$ as a whole can be identified through *Theorems 3*, *4* and *5*, respectively:

**Theorem 3.** The first stage of $DMU_o = (X_o, Z_o, Y_o)$ exhibits congestion when at least one of the following conditions is met for the optimal solution of *Model (5)*, i.e., $(\varphi^*, \lambda^*, S^{-*}, D^{+*})$:

I. $\sum_{i=1}^{m} s_i^{-*} > 0$ and $\varphi^* > 1$.

II. $\sum_{i=1}^{m} s_i^{-*} > 0$ and $\sum_{t=1}^{p} d_t^{+*} > 0$.

Kassaei et al.| Int. J. Res. Ind. Eng. 12(1) (2023) 53-72

Proof: suppose that at least one of Cases 1 or 2 holds. Therefore, according to the constraints of *Model (5)*, the *Relation (8)* holds:

$$\sum_{j=1}^{n}\lambda_j^* x_{ij} = x_{io} - s_{1,i}^{-*} \le x_{io}, \quad i=1,\dots,m,$$
$$\sum_{j=1}^{n}\lambda_j^* z_{tj} = \varphi_1^* z_{to} + d_{1,t}^{+*} \gneqq z_{to}, \quad t=1,\dots,p, \tag{8}$$
$$\sum_{j=1}^{n}\lambda_j^* = 1, \ \lambda_j^* \ge 0, \quad j=1,\dots,n.$$

Now, according to the *Relation (8)* and the convexity principle, $(\sum_{j=1}^{n}\lambda_j^* X_j, \sum_{j=1}^{n}\lambda_j^* Z_j)$ is a member of *PPS (1)* corresponding to the first stages $(X_j, Z_j),\ j=1,\dots,n$, which can produce an output greater than $Z_o$ (at least in one component) without worsening the input $X_o$. This means that the first stage of $DMU_o$, i.e. $(X_o, Z_o)$, exhibits congestion compared to the first stages of other systems.

In this case, the amount of congestion in ith input can be achieved by the value of $s_i^{-*}$ obtained from *Model (5)*. As a managerial interpretation of *Theorem 3*, if Case 1 of this theorem is established, then decreasing the components of input X will increase all components of output Z; and, if Case 2 holds, it means that a decrease in components of input X causes an increase in at least one components of output Z without decreasing the other components of Z. Both of these mean that there is density.

According to *Definition 2,* a DMU exhibits congestion when an increase in input(s) causes a decrease in output(s). Therefore, the existence of congestion causes two main problems in DMU: one is increasing costs and the other is reducing production. Because on the one hand, congestion always reduces output and on the other hand, congestion itself is input and therefore is a cost. In this way, identifying and eliminating congestion can be useful for managing each DMU without making any changes to the production process. Hence, identifying and eliminating congestion is much more important than identifying technical inefficiencies.

**Theorem 4.** The second stage of $DMU_o = (X_o, Z_o, Y_o)$ exhibits congestion when at least one of the following conditions is met for the optimal solution of *Model (6),* i.e., $(\varphi^*, \lambda^*, D^{-*}, S^{+*})$:

I. $\sum_{t=1}^{p} d_t^{-*} > 0$ and $\varphi^* > 1$.

II. $\sum_{t=1}^{p} d_t^{-*} > 0$ and $\sum_{r=1}^{s} s_r^{+*} > 0$.

Proof: Assume that at least one of Cases 1 or 2 holds. Then, according to the constraints of *Model (6),* the *Relation (9)* holds:

$$\sum_{j=1}^{n}\mu_j^* z_{tj} = z_{to} - d_{2,t}^{-*} \le z_{to}, \quad t=1,\dots,p,$$
$$\sum_{j=1}^{n}\mu_j^* y_{rj} = \varphi_2^* y_{ro} + s_{2,r}^{+*} \gneqq y_{ro}, \quad r=1,\dots,s, \tag{9}$$
$$\sum_{j=1}^{n}\mu_j^* = 1, \ \mu_j^* \ge 0, \quad j=1,\dots,n.$$

The *Relation (9)* and the convexity principle resulted in $(\sum_{j=1}^{n}\mu_j^* Z_j, \sum_{j=1}^{n}\mu_j^* Y_j)$ is a member of *PPS (1)* corresponding to the second stages $(Z_j, Y_j)$s ( $j = 1, \ldots, n$), which can produce an output greater than $Y_o$ (at least in one component) without worsening the input $Z_o$. This means that the second stage of $DMU_o$, i.e., $(Z_o, Y_o)$, exhibits congestion compared to the second stages of other systems.

In this way, the amount of congestion in pth intermediate production can be achieved by the value of $d_t^{-*}$ obtained from *Model (6)*. Similar to *Theorem 3*, if the first case of *Theorem 4* is established then decreasing the components of Z (in the role of input of Stage 2) will increase all components of output Y; and, if Case 2 holds, it means that a decrease in components of input Z causes an increase in at least one components of output Y without decreasing the other components of Z.

**Theorem 5.** $DMU_o = (X_o, Z_o, Y_o)$ exhibits congestion when at least one of the following conditions is met for the optimal solution of *Model (7)*, i.e., $(\varphi^*, \Lambda^*, \tilde{Z}^*, S^{-*}, S^{+*})$:

I. $\sum_{i=1}^{m} s_i^{-*} > 0$ and $\varphi^* > 1$.

II. $\sum_{i=1}^{m} s_i^{-*} > 0$ and $\sum_{r=1}^{s} s_r^{+*} > 0$.

Proof: suppose that Case 1 or 2 holds. Therefore, the constraints of *Model (7)* lead to the *Relation (10)*:

$$
\begin{aligned}
&\sum_{j=1}^{n} \gamma_j^* x_{ij} = x_{io} - s_i^{-c*} \leq x_{io}, \quad i = 1, \ldots, m, \\
&\sum_{j=1}^{n} \gamma_j^* y_{rj} = \varphi^* y_{ro} + s_r^{+*} \gneqq y_{ro}, \quad r = 1, \ldots, s, \\
&\sum_{j=1}^{n} \gamma_j^* = 1, \quad \gamma_j^* \geq 0, \quad j = 1, \ldots, n.
\end{aligned} \tag{10}
$$

The *Relation (10)* and the convexity principle show that $(\sum_{j=1}^{n}\gamma_j^* X_j, \sum_{j=1}^{n}\gamma_j^* Y_j)$ is a member of *PPS (1)* corresponding to the whole systems $(X_j, Y_j)$s ($j = 1, \ldots, n$) which can produce an output greater than $Y_o$ (at least in one component) without worsening the input $X_o$. This means that $DMU_o = (X_o, Y_o)$ exhibits congestion compared to the other systems.

In this case, the amount of congestion in ith input can be achieved by the value of $s_i^{-*}$ obtained from *Model (7)*. Note that the establishment of Case 1 in *Theorem 5* implies that decreasing the components of primary input X will increase all components of final output Y; and, if Case 2 holds, it means that a decrease in components of input X causes an increase in at least one components of output Y without decreasing the other components of Y. This means that the whole system exhibits congestion.

So far, there is no particular objection related to *Theorems 3* to *5* and it is possible to identify and evaluate the congestion in each of the first and second stages along with the congestion in the whole system by solving the triple *Models (5)*, *(6)* and *(7)*. But the noteworthy point is the neglect of the relationship between the congestion of the stages and the congestion of the whole system. In other words, it is possible to exhibit congestion in a stage (or even both stages) according to *Models (5)* and *(6)*, while the same unit has no congestion according to *Model (7)*. Accordingly, in the next section, *Model (7)* will be modified to identify the congestion in the whole system without facing the mentioned problem.

## 3 | Detection and Evaluation of Congestion in Series Two-Stage Network Structure

In this section, by developing *Model (4)* proposed by Cooper et al [1], the congestion in DMUs with a series two-stage network structures is investigated. In fact, since congestion always reduces the final output and elimination of congestion leads to an increase in the final output, congestion in a two-stage structure should also be defined according to the single-stage structure. Now, it should be noted that the final output ($Y$) is a function of the primary input vector ($X$) and the intermediate output vector ($Z$) in the two-stage structure. In other words, changes in the primary input as well as changes in the intermediate output can affect the final output. Therefore, an excessive increase in one of these two factors can cause congestion and therefore reduce the final output. Accordingly, the primary input vector ($X$) and the intermediate output vector ($Z$) are the factors that should be considered in defining the congestion in the two-stage structure.

**Definition 5.** A DMU with the series two-stage network structure exhibits overall congestion when at least one of the following conditions is met:

Case 1: A decrease (increase) in one/more components of the primary input $X$ leads to an increase (decrease) in one/more components of the final output $Y$; of course, without worsening (improving) other input and output components and assuming that the intermediate production $Z$ remains constant.

Case 2: A decrease (increase) in one/more components of the middle input $Z$ from the second stage leads to an increase (decrease) in one/more final output components of $Y$; of course, without worsening (improving) other input and output components and assuming that the primary input $X$ does not worsen.

**Definition 6.** The first stage of a DMU with the series two-stage network structure exhibits congestion whenever decreasing (increasing) in one/more components of primary input $X$ leads to increasing (decreasing) in one/more final output components $Y$.

**Definition 7.** The second stage of a DMU with the series two-stage network structure exhibits congestion whenever a decrease (increase) in one/more components of the intermediate input $Z$ from the second stage leads to an increase (decrease) in one/more components of the final output $Y$.

Now, according to *Definitions 5*, *6*, and *7*, the single linear programming model is proposed that considers the relationship between the stages and the whole unit and provides a logical relationship between the stages' congestion and overall congestion. This model is as *Model (11)* (where $\varepsilon$ is a small positive non-Archimedean value):

$$\begin{aligned}
\text{Max} \quad & \varphi+\varepsilon\left(\sum_{r=1}^{s} s_r^{+}-\varepsilon\sum_{i=1}^{m} s_i^{-}-\varepsilon^2\sum_{t=1}^{p} d_t^{-}\right),\\
\text{s.t.} \quad & \sum_{j=1}^{n}\lambda_j x_{ij}+s_i^{-}=x_{io}, \quad i=1,\dots,m,\\
& \sum_{j=1}^{n}\lambda_j z_{tj}\ge z_{to}, \quad t=1,\dots,p,\\
& \sum_{j=1}^{n}\mu_j z_{tj}+d_t^{-}=z_{to}, \quad t=1,\dots,p,\\
& \sum_{j=1}^{n}\lambda_j=1, \sum_{j=1}^{n}\mu_j=1,
\end{aligned} \tag{11}$$

$\lambda_j \geq 0, \mu_j \geq 0, \qquad j=1,\ldots,n,$
$s_i^-, s_r^+, d_t^- \geq 0, \qquad i=1,\ldots,m, t=1,\ldots,p, r=1,\ldots,s.$

It should be noted that although *Model (11)* is presented in the form of a single linear model, it requires four steps of optimization. These steps include maximizing the variable $\varphi$ and then the variables $s_r^+ (r=1,\ldots,s)$, respectively; next, the variables $s_i^- (i=1,\ldots,m)$ and then $d_t^- (t=1,\ldots,p)$ should be minimized, respectively. Now, according to *Definition 5* to *7*, *Theorem 6* is presented to identify and evaluate the congestion in the series two-stage network structure.

**Theorem 6.** $DMU_o = (X_o, Z_o, Y_o)$ exhibits overall congestion if for the optimal solution of *Model (11)*, i.e., $(\varphi^*, \lambda^*, \mu^*, S^{-*}, D^{-*}, S^{+*})$, at least one of the following conditions holds:

I. $\sum_{i=1}^{m} s_i^{-*} + \sum_{t=1}^{p} d_t^{-*} > 0$ and $\varphi^* > 1$.
II. $\sum_{i=1}^{m} s_i^{-*} + \sum_{t=1}^{p} d_t^{-*} > 0$ and $\sum_{r=1}^{s} s_r^{+*} > 0$.

The amount of congestion in the primary inputs and the intermediate inputs of the second stage can be determined by the optimal solution $s_i^{-*} (i=1,\ldots,m)$ and $d_t^{-*} (t=1,\ldots,p)$, respectively.

The first stage of $DMU_o = (X_o, Z_o, Y_o)$ exhibits congestion if at least one of the following conditions is met for an optimal solution of *Model (11)* such as $(\bar{\varphi}^*, \bar{\lambda}^*, \bar{S}^{-*}, \bar{D}^{-*}, \bar{S}^{+*})$:

I. $\sum_{i=1}^{m} \bar{s}_i^{-*} > 0$ and $\varphi^* > 1$.
II. $\sum_{i=1}^{m} \bar{s}_i^{-*} > 0$ and $\sum_{r=1}^{s} \bar{s}_r^{+*} > 0$.

The values of $\bar{s}_i^{-*}$ s $(i=1,\ldots,m)$ indicate the amount of congestion in primary inputs.

The second stage of $DMU_o = (X_o, Z_o, Y_o)$ exhibits congestion if at least one of the following conditions is met for an optimal solution of *Model (11)* such as $(\bar{\bar{\varphi}}^*, \bar{\bar{\lambda}}^*, \bar{\bar{S}}^{-*}, \bar{\bar{D}}^{-*}, \bar{\bar{S}}^{+*})$:

I. $\sum_{t=1}^{p} \bar{\bar{d}}_t^{-*} > 0$ and $\bar{\bar{\varphi}}^* > 1$.
II. $\sum_{t=1}^{p} \bar{\bar{d}}_t^{-*} > 0$ and $\sum_{r=1}^{s} \bar{\bar{s}}_r^{+*} > 0$.

The values of $\bar{\bar{d}}_t^{-*}$ s $(t=1,\ldots,p)$ indicate the amount of congestion in intermediate inputs of the second stage.

Proof (Case (a)): Suppose that *Model (11)* achieves an optimal solution such that $\sum_{i=1}^{m} s_i^{-*} + \sum_{t=1}^{p} d_t^{-*} > 0$ and meets at least one of the conditions $\varphi^* > 1$ or $\sum_{r=1}^{s} s_r^{+*} > 0$. In this case, two situations may occur:

Kassaei et al.| Int. J. Res. Ind. Eng. 12(1) (2023) 53-72

Situation 1: $\sum_{t=1}^{p} d_t^{-*} > 0$. In this case, according to the constraints of *Model (11)*, the *Relation (12)* holds for the optimal solution $(\varphi^*, \lambda^*, \mu^*, S^{-*}, D^{-*}, S^{+*})$:

$$
\begin{aligned}
&\sum_{j=1}^{n} \lambda_j^* X_j = X_o - S^{-*}, \sum_{j=1}^{n} \lambda_j^* Z_j \geq Z_o \geq Z_o - D^{-*},\\
&\sum_{j=1}^{n} \mu_j^* Z_j = Z_o - D^{-*} \lneqq Z_o, \sum_{j=1}^{n} \mu_j^* Y_j = \varphi^* Y_o + S^{+*} \gneqq Y_o,\\
&\sum_{j=1}^{n} \lambda_j^* = 1, \sum_{j=1}^{n} \mu_j^* = 1, \lambda_j^* \geq 0, \mu_j^* \geq 0 \; j = 1,\dots,n,\\
&S^{+*}, S^{-*} \geq 0, D^{-*} \gneqq 0.
\end{aligned}
\tag{12}
$$

From the *Relation (12)* it can be seen that $(X_o - S^{-*}, Z_o - D^{-*}, \varphi^* Y_o + S^{+*})$ is a member of the *PPS (3)* that can produce an output greater than the final output $Y_o$ (at least in one input component) along with the intermediate output less than $Z_o$ (at least in one component), without worsening the primary input $X_o$. This means that $DMU_o = (X_o, Z_o, Y_o)$ exhibits congestion according to Case 2 of *Definition 5*.

Situation 2: $\sum_{i=1}^{m} s_i^{-*} > 0, \sum_{t=1}^{p} d_t^{-*} = 0$. In this case, similar to Case 1, according to the constraints of *Model (11)* in the optimal solution, *Relation (13)* is obtained:

$$
\begin{aligned}
&\sum_{j=1}^{n} \lambda_j^* X_j = X_o - S^{-*} \lneqq X_o, \sum_{j=1}^{n} \lambda_j^* Z_j \geq Z_o,\\
&\sum_{j=1}^{n} \mu_j^* Z_j = Z_o, \sum_{j=1}^{n} \mu_j^* Y_j = \varphi^* Y_o + S^{+*} \gneqq Y_o,\\
&\sum_{j=1}^{n} \lambda_j^* = 1, \sum_{j=1}^{n} \mu_j^* = 1, \lambda_j^* \geq 0, \mu_j^* \geq 0 \; j = 1,\dots,n,\\
&S^{+*} \gneqq 0, S^{-*} \geq 0.
\end{aligned}
\tag{13}
$$

and from the *Relation (13)* it is concluded that $(X_o - S^{-*}, Z_o, \varphi^* Y_o + S^{+*})$ is a member of the *PPS (3)* that using a primary input less than $X_o$ (at least in one input component) along with the intermediate output $Z_o$ leads to the production of the final output greater than $Y_o$ (at least in one input component). This means that $DMU_o = (X_o, Z_o, Y_o)$ exhibits congestion according to Case 1 of *Definition 5*.

Conversely, suppose that $DMU_o = (X_o, Z_o, Y_o)$ exhibits congestion according to *Definition 5*. So two cases can happen:

Case 1: Assume that Case 2 of *Definition 5* holds. Then a unit such as $(\bar{X}, Z_o, \bar{Y})$ is available in the *PPS (3)* such that $\bar{X} \lneqq X_o, \bar{Y} \gneqq Y_o$. In this case, according to the membership condition of $(\bar{X}, Z_o, \bar{Y})$ in the *PPS (3)*, there are $\bar{\lambda}_j, \bar{\mu}_j \; (j = 1,\dots,n)$ which satisfy the *Relations (14)*:

$$\sum_{j=1}^{n}\bar{\lambda}_j X_j = \bar{X} \lneqq X_o,$$

$$\sum_{j=1}^{n}\bar{\lambda}_j Z_j \geq Z_o,$$

$$\sum_{j=1}^{n}\bar{\mu}_j Z_j = Z_o, \tag{14}$$

$$\sum_{j=1}^{n}\bar{\mu}_j Y_j \geq \bar{Y} \gneqq Y_o,$$

$$\sum_{j=1}^{n}\bar{\lambda}_j = 1, \sum_{j=1}^{n}\bar{\mu}_j = 1, \bar{\lambda}_j \geq 0, \bar{\mu}_j \geq 0\, j = 1,\ldots,n.$$

By defining the slack variables $\bar{s}_i^-(i=1,\ldots,m)$, $\bar{D}^-$ and $\bar{s}_r^+(r=1,\ldots s)$ in the first, third, and fourth constraints of the *Relations (14)*, respectively, this relation can be rewritten as follows:

$$\sum_{j=1}^{n}\bar{\lambda}_j x_{ij} + \bar{s}_i^- = x_{io}, i = 1,\ldots,m,$$

$$\sum_{j=1}^{n}\bar{\mu}_j Z_j - \bar{D}^- = Z_o, (\bar{D}^- = 0) \tag{15}$$

$$\sum_{j=1}^{n}\bar{\mu}_j Y_j - \bar{s}_r^+ = \bar{\varphi} y_{ro}, r = 1,\ldots,s.$$

Therefore, it is concluded that $(\bar{\lambda},\bar{\mu},\bar{\varphi} \geq 1,\bar{S}^- \gneqq 0,\bar{S}^+ \geq 0,\bar{D}^- = 0)$ is a feasible solution for the *Model (11)* such that $\sum_{i=1}^{m} s_i^{-*} + \sum_{t=1}^{p} d_t^{-*} > 0$ and at least one condition $\bar{\varphi} \gneqq 1$ or $\bar{S}^+ \gneqq 0$ is established.

Case 2: Assume that Case 2 of *Definition 5* holds. In this case, a unit such as $(\bar{X},\bar{Z},\bar{Y})$ is available in the production *Possibility (3)* such that $\bar{X} \lneqq X_o, \bar{Z} \lneqq Z_o, \bar{Y} \gneqq Y_o$. In this case, the proof is similar to the previous case and we are done.

Proof: Suppose that *Model (11)* achieves the optimal solution that satisfies $\sum_{i=1}^{m} s_i^{-*} > 0$ along with at least one of the conditions $\varphi^* > 1$ or $\sum_{r=1}^{s} s_r^{+*} > 0$. In this case, according to the constraints of *Model (11)* in the optimal solution, *Relation (16)* is obtained:

$$\sum_{j=1}^{n}\lambda_j^* X_j = X_o - S^{-*} \lneqq X_o, \sum_{j=1}^{n}\lambda_j^* Z_j \geq Z_o \geq Z_o - D^{-*},$$

$$\sum_{j=1}^{n}\mu_j^* Z_j = Z_o - D^{-*}, \sum_{j=1}^{n}\mu_j^* Y_j = \varphi^* Y_o + S^{+*} \gneqq Y_o, \tag{16}$$

$$\sum_{j=1}^{n}\lambda_j^* = 1, \sum_{j=1}^{n}\mu_j^* = 1, \lambda_j^* \geq 0, \mu_j^* \geq 0, j = 1,\ldots,n,$$

$$S^{+*} \gneqq 0, S^{-*} \geq 0, D^{-*} \geq 0.$$

Now, from *Relation (16)* it follows that $(X_o - S^{-*}, Z_o - D^{-*}, \varphi^* Y_o + S^{+*})$ is a member of the production possibilities *Set (3)* with a primary input less than $X_o$ (at least in one input component), which leads to

Kassaei et al.| Int. J. Res. Ind. Eng. 12(1) (2023) 53-72

producing a final output greater than $Y_o$ (at least in one input component). This means that the first stage exhibits congestion according to *Definition 6*.

Conversely, suppose the first stage of $DMU_o = (X_o, Z_o, Y_o)$ exhibits congestion. Then, there is a unit like $(\bar{X}, \bar{Z}, \bar{Y})$ in the *PPS (3)* such that $\bar{X} \lneqq X_o, \bar{Y} \gneqq Y_o$. In this case, according to the membership condition of the *PPS (3)*, there are $\bar{\lambda}_j, \bar{\mu}_j$ which satisfy the *Relations (17)*:

$$\begin{aligned}
&\sum_{j=1}^{n} \bar{\lambda}_j X_j = \bar{X} - S^- \lneqq X_o,\ S^- \geq 0,\\
&\sum_{j=1}^{n} \bar{\lambda}_j Z_j \geq \bar{Z},\quad \sum_{j=1}^{n} \bar{\mu}_j Z_j = \bar{Z},\\
&\sum_{j=1}^{n} \bar{\mu}_j Y_j \geq \bar{Y} \gneqq Y_o, \rightarrow (\sum_{j=1}^{n} \bar{\mu}_j Y_j + \bar{S}^+ = \bar{\varphi} Y_o),\ (\bar{\varphi} > 1 \text{ or } \bar{S}^+ \gneqq 0),\\
&\sum_{j=1}^{n} \bar{\lambda}_j = 1, \sum_{j=1}^{n} \bar{\mu}_j = 1, \bar{\lambda}_j \geq 0, \bar{\mu}_j \geq 0, j = 1, \dots, n.
\end{aligned} \tag{17}$$

From the *Relation (17)* it is concluded that a feasible solution can be found for *Model (11)* that satisfies $\sum_{i=1}^{m} \bar{s}_i^{-*} > 0$ and at least one of the conditions $\bar{\varphi}^* > 1$ or $\sum_{r=1}^{s} \bar{s}_r^{+*} > 0$.

Remark 2: For each feasible solution of *Model (11)*, according to the constraints $\sum_{j=1}^{n} \lambda_j z_{tj} \geq z_{to}$, $\sum_{j=1}^{n} \mu_j z_{tj} + d_t^- = z_{to}$ and $d_t^- \geq 0$ it turns out that $\sum_{j=1}^{n} \lambda_j z_{tj} \geq \sum_{j=1}^{n} \mu_j z_{tj}$.

Remark 3: According to *Theorem 6*, it is clear that finding the congestion in the first and second stages is not enough to find an optimal solution for *Model (11)*. Rather, to ensure that the mentioned conditions in Cases 2 and 3 of *Theorem 6* are met, it is better to investigate the multiple optimal solutions of *Model (11)*.

Remark 4: According to the sequence of optimization steps in *Model (11)*, it is clear that the detection and evaluation of input congestion in the first stage has a higher priority than the second stage. The reason can be seen in the sensitivity of the primary input and final output components. In fact, detection and evaluation of congestion are not pleasant for decision-makers without allowing improvement in the primary input and final output.

## 4 | Numerical Example and Case Study

In this section, first, with a numerical example, the distinguishing power of the proposed model is compared with conventional congestion-detecting models in the series two-stage network structure. Then, the proposed model is applied to a case study.

### 4.1 | Numerical Example

Consider 4 DMUs with the series two-stage network structure according to *Fig. 3* with a single primary input x, single intermediate output z, and single final output y.

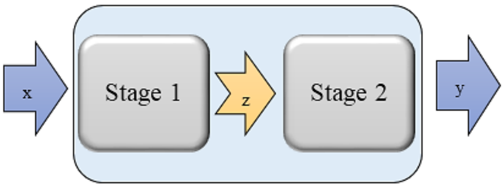


**Fig. 3. A DMU with the series two-stage structure.**

The related data for these 4 DMUs are listed in *Table 1*.

**Table 1. Data of 4 DMUs with the series two-stage network structure.**

| | $DMU_1$ | $DMU_2$ | $DMU_3$ | $DMU_4$ |
|---|---|---|---|---|
| X | 1 | 2 | 3 | 5 |
| Z | 0.5 | 2 | 2 | 1 |
| Y | 2 | 3 | 2 | 3 |

The results of *Models (5)* and *(6)* indicate the presence of congestion in the first stage of $DMU_4$, while this DMU does not show overall congestion according to *Model (7)*. However, according to *Model (11)*, the same DMU, in addition to the first stage, shows overall congestion too.

## 4.2 | Case Study

In this section, the existence of congestion in 24 Taiwanese insurance companies will be examined that are active in the non-life insurance industry [43]. As you know, the non-life insurance industry, like other service industries, expects to make a profit in exchange for providing services to its customers. But the remarkable point is that the profit of these companies is not only obtained through insurance services. Non-life insurance companies use premiums derived from systems like agencies, brokers, and lawyers as capital to support investment. With this account, the entire production process of the non-life insurance industry can be divided into two stages: 1) the premium business, and 2) the profit generation. In other words, in the first stage, each insurance company is attracted through customer insurance marketing to pay direct written premiums and premiums are also received from the other insurance companies. Then, in the second stage, the collected premiums are placed in a portfolio to make a profit.

The structure of each insurance company is shown in *Fig. 4*. The inputs of each company, which are the inputs of the first section, are the operating and insurance expenses. The operating expenses include the salaries of employees and various types of expenses incurred in daily work. On the other hand, the insurance expenses consist of expenses paid to agencies, brokers, lawyers, and other expenses associated with marketing insurance services.

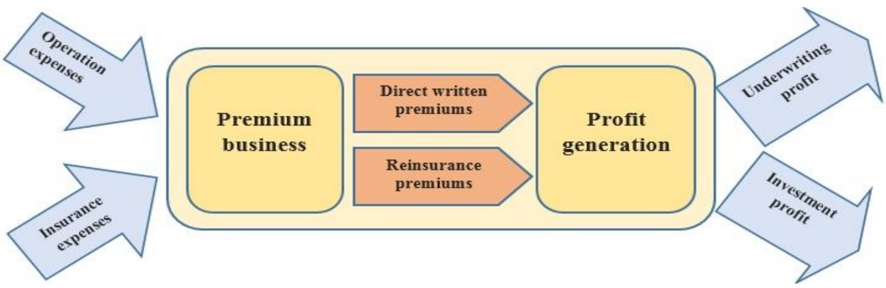


**Fig. 4. Structure of the non-life insurance companies.**

The outputs of each company, which is the output of the second stage, are the underwriting profit from the insurance trade and the investment profit from the investment portfolio. Also, two intermediate products correspond to each company as the output of the marketing process as well as the input of the investment process. They are the direct written premium received from the insured customers and the reinsurance premium, which is the premium received from the divested companies. *Table 2* shows the

Kassaei et al.| Int. J. Res. Ind. Eng. 12(1) (2023) 53-72

data related to each company. The results of the congestion investigation using the proposed *Model (11)* along with the results obtained from the one-model method of Cooper et al. [1] are listed in *Table 3*. The second and third columns show the results of overall congestion according to the method proposed by Cooper et al. [1], and columns 4 to 7 show the results of overall congestion according to the proposed *Model (11)*.

**Table 2. Data of 24 non-life insurance companies in Taiwan.**

| | Operation Expenses | Insurance Expenses | Direct Written Premiums | Reinsurance Premiums | Underwriting Profit | Investment Profit |
|---|---|---|---|---|---|---|
| 1 | 1178744 | 673512 | 7451757 | 856735 | 984143 | 681687 |
| 2 | 1381822 | 1352755 | 10020274 | 1812894 | 1228502 | 834754 |
| 3 | 1177494 | 592790 | 4776548 | 560244 | 293613 | 658428 |
| 4 | 601320 | 594259 | 3174851 | 371863 | 248709 | 177331 |
| 5 | 6699063 | 3531614 | 37392862 | 1753794 | 7851229 | 3925272 |
| 6 | 2627707 | 668363 | 9747908 | 952326 | 1713598 | 415058 |
| 7 | 1942833 | 1443100 | 10685457 | 643412 | 2239593 | 439039 |
| 8 | 3789001 | 1873530 | 17267266 | 1134600 | 3899530 | 622868 |
| 9 | 1567746 | 950432 | 11473162 | 546337 | 1043778 | 264098 |
| 10 | 1303249 | 1298470 | 8210389 | 504528 | 1697941 | 554806 |
| 11 | 1962448 | 672414 | 7222378 | 643178 | 1486014 | 18259 |
| 12 | 2592790 | 650952 | 9434406 | 1118489 | 1574191 | 909295 |
| 13 | 2609941 | 1368802 | 13921464 | 811343 | 3609236 | 223047 |
| 14 | 1396002 | 988888 | 7396396 | 465509 | 1401200 | 332283 |
| 15 | 2184944 | 651063 | 10422297 | 749893 | 3355197 | 555482 |
| 16 | 1211716 | 415071 | 5606013 | 402881 | 854054 | 197947 |
| 17 | 1453797 | 1085019 | 7695461 | 342489 | 3144484 | 371984 |
| 18 | 757515 | 547997 | 3631484 | 995620 | 692731 | 163927 |
| 19 | 159422 | 182338 | 1141950 | 483291 | 519121 | 46857 |
| 20 | 145442 | 53518 | 316829 | 131920 | 355624 | 26537 |
| 21 | 84171 | 26224 | 225888 | 40542 | 51950 | 6491 |
| 22 | 15993 | 10502 | 52063 | 14574 | 82141 | 4181 |
| 23 | 54693 | 28408 | 245910 | 49864 | 0.1 | 18980 |
| 24 | 163297 | 235094 | 476419 | 644816 | 142370 | 16976 |

Note that, as shown in *Table 3*, most of the companies which exhibit congestion (without considering intermediate outputs) according to the one-model method of Cooper et al. [1], exhibit congestion concerning the proposed *Model (11)*, too. Nonetheless, Companies 12 and 17 which exhibit no congestion according to Cooper et al. [1] method, exhibit congestion concerning the proposed *Model (11)*. In other words, according to the proposed model, the congestion value of company 12 in the first primary input is equal to 408221.85, and the congestion value of company 17 in the second primary input is equal to 196734.31. It should also be noted that according to the results of *Table 3*, none of the insurance companies in the second stage doesn't exhibit any congestion.

## 5 | Conclusion

Many real-world problems can be modeled based on the series two-stage network structure. On the other hand, congestion is one of the basic concepts in data envelopment analysis which can play an important role in reducing costs and increasing output. In this paper, it is shown that the existing classical models (especially, the one-model method of Cooper et al. [1]) are only able to detect the congestion in each stage or the whole unit independently. While ignoring the relationship between stages and the whole unit can interfere with the relationship between the congestion of stages and the congestion of the whole unit. Therefore, the definition of congestion is developed for DMUs with the series two-stage network structure. According to this definition, congestion is evaluated by providing a logical relationship between the congestion of stages and whole the system.

It should be noted that the proposed method in this study, like the model proposed by Cooper et al. [1], has a relatively long computational process. Because in 3 steps and by solving the linear programming problem 3 times, it identifies the congestion of the inputs. Anyway, since no modeling has been done to identify the congestion in the two-stage network system, this issue can be considered the innovation of the current study. On the other hand, the simplicity of interpreting the congestion calculation process can be considered another advantage of the proposed method.



Finally, the following topics can be suggested for future research:

I. Provide a method to identify the congestion in the two-stage network structure by solving a maximum of two problems (to reduce the computation complexity).
II. Development of the proposed method for more complex network structures like multi-stage structures or cases that also have initial input for the second stage.
III. Congestion detection in network structures that have an undesirable final output.

**Table 3. Data of 24 non-life insurance companies in Taiwan.**

| | The Results of Cooper et al. Model [1] | | The Results of the Proposed Model [2] | | | |
|---|---|---|---|---|---|---|
| | Operation Expenses | Insurance Expenses | Operation Expenses | Insurance Expenses | Direct written Premiums | Reinsurance Premiums |
| 1 | - | - | - | - | - | - |
| 2 | - | - | 0 | 435381.42 | 0 | 0 |
| 3 | - | - | - | - | - | |
| 4 | 0 | 8533.51 | 0 | 170911.69 | 0 | 0 |
| 5 | - | - | - | - | - | - |
| 6 | 289135.71 | 0 | 415652.12 | 0 | 0 | 0 |
| 7 | 0 | 129975.69 | 0 | 303989.77 | 0 | 0 |
| 8 | - | - | - | - | - | - |
| 9 | - | - | - | - | - | - |
| 10 | 0 | 490690.75 | 0 | 492293.55 | 0 | 0 |
| 11 | - | - | - | - | - | - |
| 12 | - | - | 408221.85 | 0 | 0 | 0 |
| 13 | - | - | - | - | - | - |
| 14 | 0 | 64039.83 | 0 | 132124.46 | 0 | 0 |
| 15 | - | - | - | - | - | - |
| 16 | - | - | - | - | - | - |
| 17 | - | - | 0 | 196734.31 | 0 | 0 |
| 18 | 0 | 19392.62 | 0 | 39461.59 | 0 | 0 |
| 19 | - | - | - | - | - | - |
| 20 | - | - | - | - | - | - |
| 21 | 13109.6 | 0 | 4921.85 | 0 | 0 | 0 |
| 22 | - | - | - | - | - | - |
| 23 | - | - | - | - | - | - |
| 24 | 0 | 68094.56 | 0 | 50642.59 | 0 | 0 |

# References


[1] Cooper, W. W., Deng, H., Huang, Z. M., & Li, S. X. (2002). A one-model approach to congestion in data envelopment analysis. *Socio-economic planning sciences*, *36*(4), 231-238.

[2] Färe, R., & Svensson, L. (1980). Congestion of factors of production. *Econometrica*, *48*(7), 1743-1753.

[3] Färe, R., & Grosskopf, S. (1983). Measuring congestion in production. *Zeitschrift für nationalökonomie/journal of economics*, *43*(3), 257-271.

[4] Färe, R., Grosskopf, S., & Lovell, C. K. (1985). *The measurement of efficiency of production*. Springer Science & Business Media. New York. https://link.springer.com/book/10.1007/978-94-015-7721-2

[5] Cooper, W. W., Thompson, R. G., & Thrall, R. M. (1996). Introduction: extensions and new developments in DEA. *Annals of operations research*, *66*(1), 1-45. https://doi.org/10.1007/BF02125451

[6] 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), 1-25.

[7] Brockett, P. L., Cooper, W. W., Wang, Y., & Shin, H. C. (1998). Inefficiency and congestion in Chinese production before and after the 1978 economic reforms. *Socio-economic planning sciences*, *32*(1), 1-20.

[8] Jahanshahloo, G. R., & Khodabakhshi, M. (2004). Suitable combination of inputs for improving outputs in DEA with determining input congestion: considering textile industry of China. *Applied mathematics and computation*, *151*(1), 263-273.

[9] Wei, Q., & Yan, H. (2009). Weak congestion in output additive data envelopment analysis. *Socio-economic planning sciences*, *43*(1), 40-54.

[10] Sueyoshi, T., & Sekitani, K. (2007). The measurement of returns to scale under a simultaneous occurrence of multiple solutions in a reference set and a supporting hyperplane. *European journal of operational research*, *181*(2), 549-570.

[11] Cooper, W. W., Gu, B., & Li, S. (2001). Comparisons and evaluations of alternative approaches to the treatment of congestion in DEA. *European journal of operational research*, *132*(1), 62-74.

[12] Wei, Q., & Yan, H. (2004). Congestion and returns to scale in data envelopment analysis. *European journal of operational research*, *153*(3), 641-660.

[13] Tone, K., & Sahoo, B. K. (2004). Degree of scale economies and congestion: a unified DEA approach. *European journal of operational research*, *158*(3), 755-772.

[14] Noura, A. A., Lotfi, F. H., Jahanshahloo, G. R., Rashidi, S. F., & Parker, B. R. (2010). A new method for measuring congestion in data envelopment analysis. *Socio-economic planning sciences*, *44*(4), 240-246.

[15] Sueyoshi, T., & Sekitani, K. (2009). DEA congestion and returns to scale under an occurrence of multiple optimal projections. *European journal of operational research*, *194*(2), 592-607.

[16] Mehdiloozad, M., Zhu, J., & Sahoo, B. K. (2018). Identification of congestion in data envelopment analysis under the occurrence of multiple projections: a reliable method capable of dealing with negative data. *European journal of operational research*, *265*(2), 644-654.

[17] Karimi, B., Khorram, E., & Moeini, M. (2016). Identification of congestion by means of integer-valued data envelopment analysis. *Computers & industrial engineering*, *98*, 513-521.

[18] 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.

[19] Shabanpour, H., Fathi, A., Yousefi, S., & Saen, R. F. (2019). Ranking sustainable suppliers using congestion approach of data envelopment analysis. *Journal of cleaner production*, *240*, 118190. https://doi.org/10.1016/j.jclepro.2019.118190

[20] Adimi, M. E., Rostamy-Malkhalifeh, M., Hosseinzadeh Lotfi, F., & Mehrjoo, R. (2019). A new linear method to find the congestion hyperplane in DEA. *Mathematical sciences*, *13*(1), 43-52.

[21] Salehi, A., Hosseinzadehlotfi, F., & Rostamy-Malkhalifeh, M. (2020). A Method for identifying congestion using non-radial models in data envelopment analysis. *International journal of nonlinear analysis and applications*, *11*(1), 225-235.

[22] Ren, X. T., Fukuyama, H., & Yang, G. L. (2022). Eliminating congestion by increasing inputs in R&D activities of Chinese universities. *Omega*, *110*, 102618. https://doi.org/10.1016/j.omega.2022.102618

[23] Navidi, S., Rostamy-Malkhalifeh, M., & Hosseinzadeh Lotfi, F. (2021). Measuring congestion in data envelopment analysis without solving any models. *Scientia Iranica*, *28*(5), 2926-2932.

[24] Yang, G., Ren, X., Khoveyni, M., & Eslami, R. (2020). Directional congestion in the framework of data envelopment analysis. *Journal of management science and engineering*, *5*(1), 57-75.

[25] Mendoza-Velázquez, A., & Benita, F. (2019). Efficiency, productivity, and congestion performance: analysis of the automotive cluster in Mexico. *Journal of industry, competition and trade*, *19*(4), 661-678.

[26] Cho, C. C., & Yang, F. C. (2019). Incorporating the non-separable characteristic of undesirable outputs into congestion analysis: a case of regional industries in China. *Journal of industrial and production engineering*, *36*(4), 248-257.

[27] Shadab, M., Saati, S., Farzipoor Saen, R., & Mostafaee, A. (2020). Measuring congestion by anchor points in DEA. *Sādhanā*, *45*, 1-10.

[28] Khoshroo, A., Izadikhah, M., & Emrouznejad, A. (2021). Energy efficiency and congestion considering data envelopment analysis and bounded adjusted measure: a case of tomato production. *Journal of cleaner production*, *328*, 129639. https://doi.org/10.1016/j.jclepro.2021.129639

[29] Fallahnejad, R., Asadi Rahmati, S., & Moradipour, K. (In Press). Decomposition of cost efficiency, given the set of price and cost production possibilities in data envelopment analysis. *Journal of applied research on industrial engineering*. https://doi.org/10.22105/jarie.2022.310345.1393

[30] Zuo, Z., Guo, H., Li, Y., & Cheng, J. (2022). A two-stage DEA evaluation of Chinese mining industry technological innovation efficiency and eco-efficiency. *Environmental impact assessment review, 94*, 106762. https://doi.org/10.1016/j.eiar.2022.106762

[31] Silva, P. M., Moutinho, V. F., & Moreira, A. C. (2022). Do social and economic factors affect the technical efficiency in entrepreneurship activities? Evidence from European countries using a two-stage DEA model. *Socio-economic planning sciences, 82*, 101314. https://doi.org/10.1016/j.seps.2022.101314

[32] Jingxin, G., Jinbo, S., & Lufang, W. (2022). A new methodology to measure the urban construction land-use efficiency based on the two-stage DEA model. *Land use policy, 112*, 105799. https://doi.org/10.1016/j.landusepol.2021.105799

[33] Henriques, I. C., Sobreiro, V. A., Kimura, H., & Mariano, E. B. (2020). Two-stage DEA in banks: terminological controversies and future directions. *Expert systems with applications, 161*, 113632. https://doi.org/10.1016/j.eswa.2020.113632

[34] Chen, Y., Ma, X., Yan, P., & Wang, M. (2021). Operating efficiency in Chinese universities: an extended two-stage network DEA approach. *Journal of management science and engineering, 6*(4), 482-498.

[35] Izadikhah, M., Azadi, M., Toloo, M., & Hussain, F. K. (2021). Sustainably resilient supply chains evaluation in public transport: a fuzzy chance-constrained two-stage DEA approach. *Applied soft computing, 113*, 107879. https://doi.org/10.1016/j.asoc.2021.107879

[36] Mozaffari, M. R., & Ostovan, S. (2021). Finding projection in the two-stage supply chain in DEA-R with random data using (CRA) model. *Big data and computing visions, 1*(3), 146-155.

[37] Marzband, A. (2020). Precise services and supply chain prioritization in manufacturing companies using cost analysis provided in a fuzzy environment. *Journal of fuzzy extension and applications, 1*(1), 41-56. http://www.journal-fea.com/article_114139_b12f0df6b47d3095c5fd458ee6dd96fe.pdf

[38] Nematizadeh, M., & Nematizadeh, M. (2021). Profitability assessment of two-stage feedback processes with undesirable factors. *Journal of applied research on industrial engineering, 8*(Spec. Issue), 1-16.

[39] GholamAzad, M., & Pourmahmoud, J. (2021). Evaluating the congestion effect in the two-stage structures by NDEA models. *Journal of modern operations research, 1*(2). http://www.jmor.ir/article_240075.html

[40] Pritchard, R. D. (1995). *Productivity measurement and improvement: organizational case studies*. Greenwood Publishing Group.

[41] Alirezaee, M. R., & MirHassani, S. A. (2004). From the data envelopment analysis technique to the development of a productivity system. In *Data envelopment analysis and performance management* (pp 185-192). Birmingham: Aston Business School, Aston University.

[42] Bogetoft, P., & Otto, L. (2010). *Benchmarking with dea, sfa, and r* (Vol. 157). Springer Science & Business media.

[43] Kao, C., & Hwang, S. N. (2008). Efficiency decomposition in two-stage data envelopment analysis: an application to non-life insurance companies in Taiwan. *European journal of operational research, 185*(1), 418-429.



Kassaei et al.| Int. J. Res. Ind. Eng. 12(1) (2023) 53-72