

# Selecting data envelopment analysis models: A data-driven application to EU countries

**Mehdi Toloo<sup>1</sup>**

*Department of Systems Engineering, VŠB-Technical University of Ostrava, Ostrava, Czech Republic*  
*Department of Operations Management & Business Statistics, College of Economics and Political Science, Sultan Qaboos University, Muscat, Oman*  
E-mails: [mehdi.toloo@vsb.cz](mailto:mehdi.toloo@vsb.cz); URL: <http://homel.vsb.cz/~tol0013/>

**Esmail Keshavarz**

*Department of Mathematics, Sirjan Branch, Islamic Azad University, Sirjan, Iran*  
E-mail: [esmailk212@gmail.com](mailto:esmailk212@gmail.com)

**Adel Hatami-Marbini**

*Department of Management and Entrepreneurship, Leicester Castle Business School, De Montfort University, Leicester, UK, E-mail: [adel.hatamimarbini@dmu.ac.uk](mailto:adel.hatamimarbini@dmu.ac.uk)*

## Abstract

Data envelopment analysis (DEA) is a non-parametric data-driven approach for evaluating the efficiency of a set of homogeneous decision-making units (DMUs) with multiple inputs and multiple outputs. The number of performance factors (inputs and outputs) plays a crucial role when applying DEA to real-world applications. In other words, if the number of performance factors is significantly greater than the number of DMUs, it is highly possible to arrive at a large portion of efficient DMUs, which practically may become problematic due to the lack of ample discrimination among DMUs. The current research aims to develop an array of selecting DEA models to narrow down the performance factors based upon a rule of thumb. To this end, we show that the input- and output-oriented selecting DEA models may select different factors and then present the integrated models to identify a set of common factors for both orientations. In addition to efficiency evaluation at the individual level, we study structural efficiency with a single production unit at the industry level. Finally, a case study on the EU countries is presented to give insight into business innovation, social economy and growth with regard to the efficiency of the EU countries and entire EU.

**Keywords:** Data envelopment analysis; data-driven; selective factors; Variable returns-to-scale; input and output orientations; structural efficiency.

## 1 Introduction

An effective appraisal system is warranted for continuous improvement of organizations. Based upon the productive efficiency concept of Farrell (1957), Charnes, Cooper, & Rhodes (1978) initially developed a data envelopment analysis (DEA) model by the use of the linear programming (LP) approach to estimate a production frontier and find relative efficiency scores of decision-making units (DMUs) where each DMU produces multiple outputs by using multiple inputs. The current literature consists of a wide variety of DEA models to assess the efficiencies of DMUs (Liu, Lu, Lu, &

---

<sup>1</sup> Corresponding author. Postal address: Department of Systems Engineering, Faculty of Economics, VŠB-Technical University of Ostrava, Sokolská třída 33, 702 00, Ostrava 1, Ostrava, Czech Republic, Tel: +420 792 272 272.

Lin (2013 a, 2013 b) and Emrouznejad & Yang (2017)). In addition to seeking the efficiency measure for each firm (production unit), the efficiency measurement of a group of firms is of importance to policymakers in numerous situations to monitor the performance of a certain industry<sup>1</sup>. This research stream, so-called “structural efficiency”, was commenced by Farrell (1957) and Førsund & Hjalmarsson (1979) in which the technical efficiency of an industry is determined by a weighted average of the [technical] efficiencies of the individual firms. There are some studies undertaken in the related literature to study this interesting topic (e.g., see Li & Cheng (2007) and Karagiannis (2015)).

Regardless of structural efficiency and individual efficiency over the DEA assessment process, performance factors, i.e., input and output factors, are of great importance. Practically, when the number of performance factors is significantly larger than the number of DMUs, it is highly likely that a big proportion of DMUs is identified as efficient which may not be desirable from the top-level management team (Eskelinen, 2017). In this respect, Jenkins & Anderson (2003) stated "*the greater the number of input and output variables [factors] in a DEA, the higher the dimensionality of the LP solution space, and the less discerning the analysis*". To tackle this problem, combining the performance factors and omission of some inappropriate or less influential factors are two different ways that can be set out to reduce the number of performance factors. So, there is a need to opt for the most important factors, which can build up a reasonable balance between the number of performance factors and the number of DMUs.

A predefined generic rule of thumb is often used to establish the balance of the number of units and the number of inputs and outputs (see for instance Golany & Yaakov, 1989; Friedman & Sinuany-stern, 1998; Dyson et al., 2001; Cooper, Seiford, & Tone, 2007; Charles, Aparicio, & Zhu, 2019 and Khezrimotlagh, Cook, & Zhu, 2019 ).

Two evolving streams of research have been proposed in the literature to deal with the problem of factor selection in DEA. The first approach revolves around statistical tests and the second approach tries to focus on the standard DEA models. Let us review the pertinent literature of these two approaches concisely.

The first stream firstly conducted by Lewin, Morey, & Cook (1982) who were applied a correlation and regression analysis to select a set of factors that are policy-relevant and administrative efficiency stance and omit redundant factors. As a rewarding study, Simar & Wilson (2001) recommended some statistical procedures including bootstrapping and Mont-Carlo experiments to test whether performance factors are irrelevant, as well as testing whether performance factors can be aggregated. Pastor, Ruiz, & Sirvent (2002) presented an efficiency contribution measure (ECM) for examining the importance of a factor in the process of efficiency assessment. The authors highlighted that performing the test may help the assessors to decide the involvement or deletion of a given factor into or from a given DEA model. Jenkins & Anderson (2003) employed a multivariate statistical approach based on the partial correlation to reduce the number of performance factors. The authors spelled out a systematic statistical method to identify which of the original correlated variables can

---

<sup>1</sup> From the economics viewpoint, a group of DMUs, such as a bank with multiple branches, a country with different provinces and a university with several faculties, refers to as an *industry*.

be omitted with minimum loss of information, and which of them need to be retained. The results showed that the omitted variables that are highly correlated with each other, and contain little information on performance can have a significant influence on the efficiency measures. Wagner & Shimshak (2007) developed a stepwise approach to perform factor selection that involves sequentially maximizing (or minimizing) the average change in the efficiencies as variables are added to or dropped from the evaluation process. Morita & Avkiran (2009) made use of external evaluation and discriminant analysis to study a performance factor selection method. To this end, a 3-level orthogonal layout experiment is exploited to find an appropriate combination of inputs and outputs, where experiments are independent of each other. Nataraja & Johnson (2011) classified the existing factor selection methods in DEA into eight categories but because of the similarity of some methods, they only analysed four most widely-used categories; (i) efficiency contribution measure (ECM), (ii) principal component analysis (PCA), (iii) regression-based test (RBT), and (iv) bootstrapping for factor selection using Monte Carlo simulations. Their genuine research objective was to provide insights into the performance of these four methods along with developing guidelines for selecting appropriate performance factors. The overall findings showed that PCA-DEA is a robust and reliable technique with the smallest run time. Amirteimoori, Despotis, & Kordrostami (2014) that proposed an iterative DEA approach to reduce the number of performance factors periodically by aggregating each two highly correlated factors in each iteration. Toloo & Babae (2015) extended an iterative-free method to deal with variable reduction problem in DEA.

The second stream was originated by Toloo, Barat, & Masoumzadeh (2015) by modifying the standard CRS model of DEA and developing two mixed binary linear programming (MBLP) models based upon individual and aggregate efficiency scores that are able to select the performance factors along with satisfying a predefined rule of thumb. The developed models are called “selecting DEA models” since the objective of the models is to select the performance factors (Toloo et al., 2015). Toloo & Tichý (2015) extended Toloo et al. (2015)'s approach to propose the multiplier and envelopment forms of DEA models for selecting performance factors under VRS assumption, leading to an increase in discriminatory power of DEA. It was proved that the multiplier models select performance factors from an optimistic viewpoint while the envelopment models consider a pessimistic viewpoint. Toloo & Allahyar (2018) proposed epsilon-free versions of Toloo & Tichý (2015)'s selecting DEA models under general returns-to-scale (GRS). As a matter of fact, these studies formulated MBLP models based on the standard input-oriented DEA models, which selection procedure of performance factors has been implemented by utilizing a binary variable associated with each selective factor. Recently, Hančlová (2019) modified the directional distance DEA model in the presence of undesirable output in order to develop an approach for selecting factors with more than one value which are measured by various definitions or standards.

This paper first briefly reviews the standard DEA models and input-oriented selecting DEA models and then formulate a new output-oriented selecting DEA model under the VRS assumption coupled with arguing some theories and underpinnings to uncover most important properties of the proposed models. Moreover, as a significant contribution of this research, the input- and output-oriented selecting DEA models are aggregated to unify performance factor selection in a way that

gets rid of the plausible inconsistency when the different orientations cause to select different factors. Finally, we revisit the concept of structural efficiency of an industry in order to make an informed decision with regard to factor selection. In addition to the theoretical development in this research, we contribute to the practical aspect by exploring a case study in which the relative efficiencies of 28 EU countries are assessed in terms of 50 potential performance factors. The findings demonstrate the key role of types of DEA models, either envelopment or multiplier, and its orientations, either input or output respects.

The remainder of the paper is organised as follows. In Section 2, the basic input- and output-oriented DEA models are briefly reviewed. Section 3 presents the input- and output-oriented selecting DEA models for both multiplier and envelopment forms under the VRS assumption. Section 4 presents a new approach to aggregate the input- and output-oriented selecting DEA models with the aim of unifying the factor selection. In Section 5, we extend our selective factor thrust to the concept of structural efficiency. The penultimate section meticulously illustrates the proposed models and theorems by assessing the performance of the EU countries and the EU as a whole. Ultimately, some conclusions and further remarks are provided in Section 7.

## **2 DEA models in benchmarking**

DEA is an advanced benchmarking approach aiming to identify the improvement ways of business operations and organizational performance. The results derived from this approach are informative and intended to provide applicable feedback to get rid of inefficiencies sources as well as making informed decisions about effective changes across firms. Figure 1 summarises a serious of steps being performed to implement the pertinent processes. Step 1 first defines the assessment structure including the operating units (DMUs) and performance factors (inputs and outputs) subject to the generic rule of thumb and consequently collects the primary and/or secondary data for each DMU. Step 2 aims to specify the production technology and appropriate returns to scale in terms of competitive levels embedded in the markets and industries under analysis. In Step 3, it is aimed to determine whether the radial or no-radial setting is used for the purpose of efficiency and economic measurement and benchmarking. The radial setting is divided into the input- and output-oriented models where the former model is seeking to obtain technical efficiency as a proportional reduction in input usage and the latter one sought to determine technical efficiency as a proportional increase in output production. The output- and input-orientated measures are equivalent efficiency measures when CRS is regarded, but are unequal when VRS exists (Färe & Lovell, 1978). Step 4 identifies the best practices and technical inefficiency by the use of DEA models which have either multiplier or envelopment forms. Step 5 includes the results and findings interpretation that helps the decision-makers and top management team provide insight and make informed decisions about the strategic planning and resource allocation across the organisation.

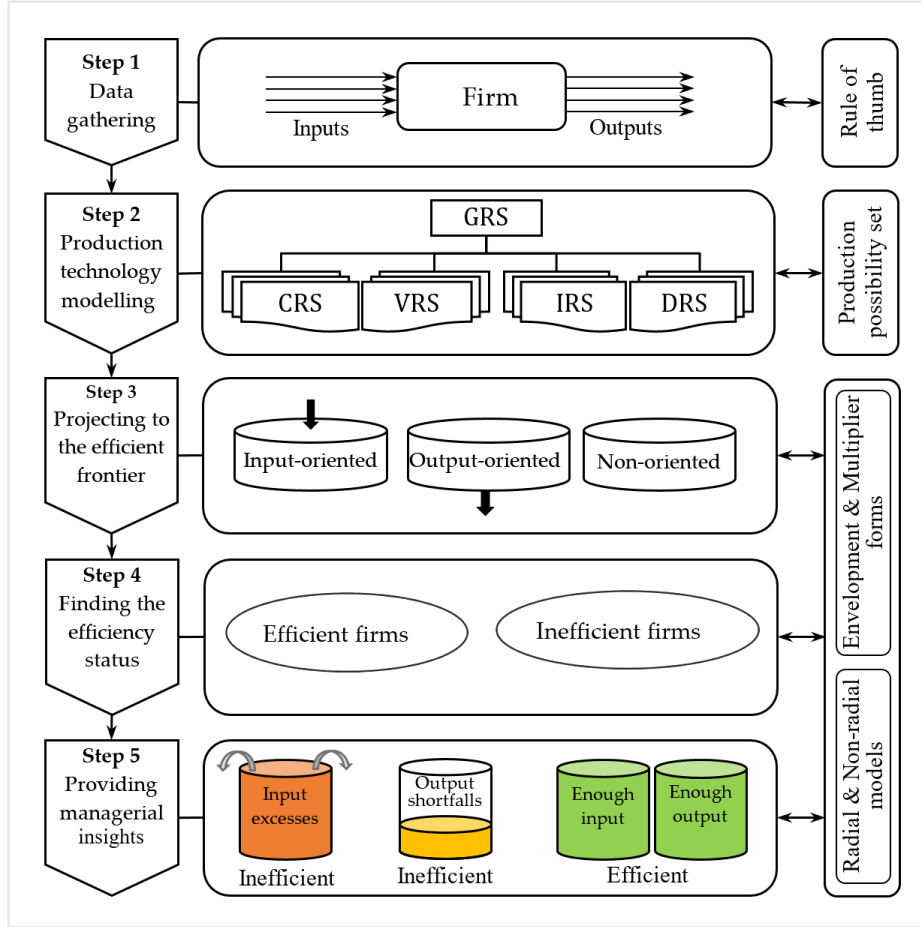


Figure 1. Summary of DEA-based procedures

Let us here review the input- and output-oriented DEA models in both multiplier and envelopment forms under different RTS assumptions. Toward this end, we assume that there are  $n$  homogenous DMUs ( $DMU_j, j = 1, \dots, n$ ). Each  $DMU_j$  consumes  $m$  different semi-positive inputs,  $\mathbf{x}_j = (x_{1j}, \dots, x_{mj})$  to produce  $s$  different semi-positive outputs,  $\mathbf{y}_j = (y_{1j}, \dots, y_{sj})$ . For the sake of brevity, the following VRS models are expressed to evaluate the relative efficiency of  $DMU_o$  (Cooper et al., 2007):

### Multiplier VRS models

#### Input-oriented

$$\begin{aligned}
 \max \varphi_o^{in} &= \sum_{r=1}^s u_r^{in} y_{ro} - w_0^{in} \\
 \text{s. t.} \\
 \sum_{i=1}^m v_i^{in} x_{io} &= 1 \\
 \sum_{r=1}^s u_r^{in} y_{rj} - w_0^{in} - \sum_{i=1}^m v_i^{in} x_{ij} &\leq 0 \quad \forall j \\
 v_i^{in} &\geq 0 \quad \forall i \\
 u_r^{in} &\geq 0 \quad \forall r \\
 w_0^{in} &\text{ free in sign}
 \end{aligned}$$

(1)

#### Output-oriented

$$\begin{aligned}
 \min \varphi_o^{out} &= \sum_{i=1}^m v_i^{out} x_{io} + w_0^{out} \\
 \text{s. t.} \\
 \sum_{r=1}^s u_r^{out} y_{ro} &= 1 \\
 \sum_{r=1}^s u_r^{out} y_{rj} - \sum_{i=1}^m v_i^{out} x_{ij} - w_0^{out} &\leq 0 \quad \forall j \\
 v_i^{out} &\geq 0 \quad \forall i \\
 u_r^{out} &\geq 0 \quad \forall r \\
 w_0^{out} &\text{ free in sign}
 \end{aligned}$$

(2)

## Envelopment VRS models

Input-oriented	Output-oriented
$\min \theta_o^{in}$	$\max \theta_o^{out}$
$\text{s. t.}$	$\text{s. t.}$
$\sum_{j=1}^n \lambda_j^{in} x_{ij} \leq \theta_o^{in} x_{io} \quad \forall i$	$\sum_{j=1}^n \lambda_j^{out} x_{ij} \leq x_{io} \quad \forall i$
$\sum_{j=1}^n \lambda_j^{in} y_{rj} \geq y_{ro} \quad \forall r$	$\sum_{j=1}^n \lambda_j^{out} y_{rj} \geq \theta_o^{out} y_{ro} \quad \forall r$
$\sum_{j=1}^n \lambda_j^{in} = 1$	$\sum_{j=1}^n \lambda_j^{out} = 1$

where  $u_r^{in}$  and  $u_r^{out}$  are the weights associated to the  $r^{th}$  output in models (1) and (2), respectively;  $v_i^{in}$  and  $v_i^{out}$  are the respective weights associated to the  $i^{th}$  input in models (1) and (2);  $w_0^{in}$  and  $w_0^{out}$  are the RTS [free] variables in models (1) and (2), respectively;  $\lambda_j^{in} \geq 0$  and  $\lambda_j^{out} \geq 0$  are intensity variables of DMU<sub>j</sub> in models (3) and (4), respectively. It is worth noting that the multiplier and envelopment forms of input (output) orientation models are mutually dual and in terms of the strong duality theorem,  $\varphi_o^{in*} = \theta_o^{in*}$  ( $\varphi_o^{out*} = \theta_o^{out*}$ ) for DMU<sub>o</sub>.

### 3 Selecting DEA models

The diagrammatical representation of the holistic decision-making framework developed in **Figure 1** shows a clear picture of applicability and importance of the performance evaluation system to achieve sustainable competitive advantages. As exemplified in **Figure 2**, DEA models play a crucial role in assessing the industry performance and making strategic decisions. In many real-world situations, the number of factors might be significantly greater than the number of operating units which is often problematic because this reduces the discrimination power of DEA models. Therefore, it has been a long-standing discussion in the related literature and reducing the number of factors and emphasising the more imperative factors are reasonable and justifiable to achieve the DEA's objectives. **In practice, the practitioner needs the number of units to be large enough in comparison with the number of performance factors so as to achieve a reasonable level of discrimination. A generic rule of thumb as  $n \geq f(m, s)$  is often used to establish the lower bound,  $f(m, s)$ , of the number of units where  $m$  and  $s$  are the number of inputs and outputs, respectively.**

The DEA literature embraces several experimental values for  $f(m, s)$  that have been suggested to the practitioners. To the best of our knowledge, Golany & Yaakov (1989) was the first study that aimed to reinforce the need of a sufficient number of inputs and outputs in DEA and  $2(m + s)$  was resultantly proposed as the lower bound of the number of units. Then, Friedman & Sinuany-stern (1998) and Dyson et al. (2001) put forth different values for  $f(m, s)$  as  $3(m + s)$  and  $2ms$ , respectively. Cooper, Seiford, & Tone (2007), with the aim of considering both aforementioned lower bounds, proposed  $f(m, s) = \max\{3(m + s), m \times s\}$ . Recently, Charles, Aparicio, & Zhu (2019) and Khezrimotlagh, Cook, & Zhu (2019) studied the discriminatory power and accuracy of DEA with respect to the number of units and performance measures.

The question here is, what are the most influential performance factors and how does one select them prior to running DEA models? In this regard, Cooper et al. (2007) recommended a before-modelling approach involving a process of selecting a small set of input and output factors at the beginning and gradually enlarging the member of the set to observe the effects of the added factors. Toloo et al. (2015) extended a by-modelling approach where a selecting DEA model chooses a subset of the inputs and outputs with the aim of meeting a predefined rule of thumb.

This section endeavours to ameliorate the benchmarking and decision-making system by opting for the more appropriate performance factors for situations where the system includes numerous inputs and outputs and cannot be used altogether in DEA models. The proposed models give the decision-makers a big advantage over the existing models by providing reliable results in the presence of more affecting factors. To achieve this, we first start with the input orientation of the selecting model and widening our scope by covering the output orientation for both the multiplier and envelopment forms. Then, we turn our discussion to the output orientation to supplement our view.

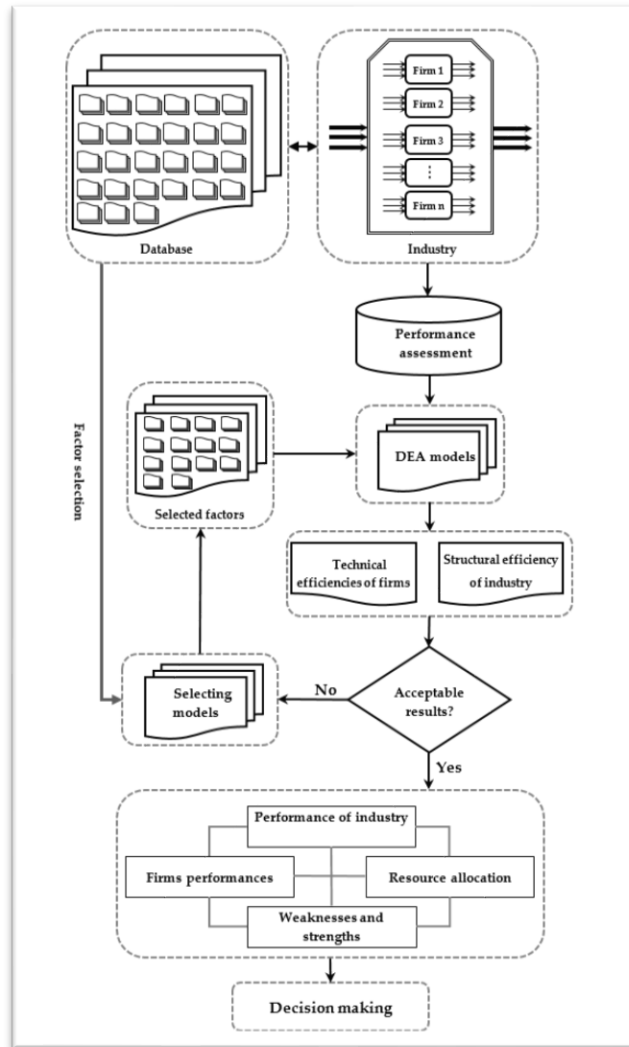


Figure 2. A decision-making framework

### 3.1 Selecting DEA models under input-oriented VRS

Let  $I^f, I^s, R^f$ , and  $R^s$  denote the index sets of the fixed-input, selective-input, fixed-output, and selective-output factors, respectively. Fixed [input/output] factors are those factors, which have been already selected by the decision maker and must be utilized in the evaluation process while selective factors provide greater leeway to either opt out of or opt into performance assessment.

Toloo & Allahyar (2018) proposed the input-oriented selecting DEA models in both multiplier and envelopment forms with the aim of satisfying a predefined rule of thumb. In their proposed selecting DEA models, which are MBLPs, all factors have been assumed to be selective factors. At present, let us focus on the modified selecting DEA models of Toloo & Allahyar (2018) by setting fixed and selective factors into the models as shown below:

### Selecting DEA models under input-oriented VRS

Multiplier form	Envelopment form
$\max \bar{\varphi}_o^{in} = \sum_{r=1}^s u_r^{in} y_{ro} - w_0^{in}$	$\min \bar{\theta}_o^{in}$
s. t.	s. t.
$\sum_{i=1}^m v_i^{in} x_{io} = 1$	$\sum_{j=1}^n \lambda_j^{in} x_{ij} \leq \bar{\theta}_o^{in} x_{io} \quad \forall i \in I^f$
$\sum_{r=1}^s u_r^{in} y_{rj} - w_0^{in} - \sum_{i=1}^m v_i^{in} x_{ij} \leq 0 \quad \forall j$	$\sum_{j=1}^n \lambda_j^{in} y_{rj} \geq y_{ro} \quad \forall r \in R^f$
$\sum_{i \in I^s} d_i^x + \sum_{r \in R^s} d_r^y \leq K - ( I^f  +  R^f )$	$\sum_{j=1}^n \lambda_j^{in} x_{ij} \leq \bar{\theta}_o^{in} x_{io} + M(1 - d_i^x) \quad \forall i \in I^s$
$\sum_{i \in I^s} d_i^x + \sum_{r \in R^s} d_r^y \geq 1$	$\sum_{j=1}^n \lambda_j^{in} y_{rj} \geq y_{ro} - M(1 - d_r^y) \quad \forall r \in R^s \quad (6)$
$v_i^{in} \leq M d_i^x \quad \forall i \in I^s$	$\sum_{i \in I^s} d_i^x + \sum_{r \in R^s} d_r^y \leq K - ( I^f  +  R^f )$
$u_r^{in} \leq M d_r^y \quad \forall r \in R^s$	$\sum_{i \in I^s} d_i^x + \sum_{r \in R^s} d_r^y \geq 1$
$d_i^x, d_r^y \in \{0,1\} \quad \forall i \in I^s, r \in R^s$	$d_i^x, d_r^y \in \{0,1\} \quad \forall i \in I^s, \forall r \in R^s$
$v_i^{in}, u_r^{in} \geq 0 \quad \forall i, r$	$\sum_{j=1}^n \lambda_j^{in} = 1$
$w_0^{in}$ free in sign	

where  $u_r^{in}$  and  $v_i^{in}$ ,  $w_0^{in}$  and  $\lambda_j^{in}$  are defined as the same as models (1) and (3), and  $\bar{\varphi}_o^{in}$  and  $\bar{\theta}_o^{in}$  are the objective functions (corresponding to radial efficiencies) of DMU<sub>o</sub> associated with the input-oriented multiplier and envelopment selecting DEA models, respectively. It must be emphasized that decision variables of selecting models (5) and (6) are different from those in models (1) and (3).  $\bar{\varphi}_o^{in}$  and  $\bar{\theta}_o^{in}$  are the objective functions (corresponding to radial efficiencies) of DMU<sub>o</sub> associated with the input-oriented multiplier and envelopment selecting DEA models, respectively.  $M$  is a sufficiently large positive number,  $|\cdot|$  indicates the cardinal of a set, and  $d_i^x$  and  $d_r^y$  are the indicator binary variables associated to the  $i^{th}$  selective input factor and the  $r^{th}$  selective output factor as defined below:

$$d_i^x = \begin{cases} 1, & i^{th} \text{ selective input is chosen} \\ 0, & \text{otherwise} \end{cases}; \quad d_r^y = \begin{cases} 1, & r^{th} \text{ selective output is chosen} \\ 0, & \text{otherwise} \end{cases}$$

In addition,  $K$  is a pre-determined value in the above models indicating the maximum number of factors that must be involved in the assessment process. The desired value of  $K$  can be computed by the generic rule of thumb as  $n \geq f(m, s)$ . In the DEA literature, the following four values for  $K$  can be regarded based on experimentally suggested lower bounds  $f(m, s)$ :

$$K = \begin{cases} \lfloor n/2 \rfloor & \text{if } f(m, s) = 2(m + s) \\ \lfloor n/3 \rfloor & \text{if } f(m, s) = 3(m + s) \\ \lfloor \sqrt{n} \rfloor & \text{if } f(m, s) = 2ms \\ \min\{\lfloor n/3 \rfloor, \lfloor 2\sqrt{n} \rfloor\} & \text{if } f(m, s) = \max\{3(m + s), ms\} \end{cases}$$

where  $\lfloor a \rfloor$  is called the *floor function* indicating the largest integer less than or equal to  $a$ .

Golany & Yaakov (1989), Friedman & Sinuany-stern (1998), Dyson et al. (2001), and Cooper et al. (2007) were respectively put forth the above lower bounds  $f(m, s)$  of the number of DMUs. Constraint  $\sum_{i \in I^s} d_i^x + \sum_{r \in R^s} d_r^y \geq 1$  enforces the models to choose at least one selective factor in the evaluation process. It is easy to verify that if  $d_i^x = 0$ , then constraint  $v_i^{in} \leq M d_i^x$  in model (5) turn



into  $v_i^{in} = 0$ , and also constraint  $\sum_{j=1}^n \lambda_j^{in} x_{ij} \leq \bar{\theta}_o^{in} x_{io} + M(1 - d_i^x)$  in model (6) will become redundant, which means that the  $i^{th}$  input has no role in performance analysis of  $DMU_o$  and it can be opted out. When  $d_i^x = 1$ , the  $i^{th}$  input plays a part due to the fact that constraint  $v_i^{in} \leq M d_i^x$  in model (5) is redundant, and constraint  $\sum_{j=1}^n \lambda_j^{in} x_{ij} \leq \bar{\theta}_o^{in} x_{io} + M(1 - d_i^x)$  in model (6) is converted to  $\sum_{j=1}^n \lambda_j^{in} x_{ij} \leq \bar{\theta}_o^{in} x_{io}$  which shows that the  $i^{th}$  input must be available throughout the evaluation process. Likewise, the value of  $d_r^y$  controls the selection status of the  $r^{th}$  output.

It should be noted that the selecting DEA models (5) or (6) are solved  $n$  times (one time for each DMU) to select a set of appropriate factors. Therefore, ultimately each performance factor has  $n$  different statuses as either *selected* or *unselected*. One way to identify the overall status of a factor (*selected* or *unselected* factor) is to consider the majority of a status that has been observed among  $n$  different statuses. After having chosen the performance factors, the conventional DEA models can be applied to measure the relative efficiencies of DMUs.

It is worth noting that the multiplier and envelopment forms of the conventional DEA models are mutually dual and the duality theorem for linear programming holds. However, the multiplier and envelopment DEA models in the presence of selective measures which are changed to the MBLP models are no longer the primal-dual pair of problems. On the other hand, similar to conventional DEA models, different orientations in each multiplier/envelopment form of selecting models may result in different selections under non-CRS conditions. Therefore, it seems worthy of further study the relation between models (5) and (6) by way of the following Theorems 1 and 2.

**Theorem 1.** The optimal objective value of the envelopment model (6) is less than or equal to the optimal value of the multiplier model (5), i.e.,  $\bar{\theta}_o^{in*} \leq \bar{\varphi}_o^{in*}$ .

**Proof.** Let  $(\bar{\varphi}_o^{in*}, \mathbf{u}^{in*}, \mathbf{v}^{in*}, \mathbf{d}^{x*}, \mathbf{d}^{y*}, w_0^{in*})$  and  $(\bar{\theta}_o^{in*}, \boldsymbol{\lambda}^{in*}, \bar{\mathbf{d}}^{x*}, \bar{\mathbf{d}}^{y*})$  be the optimal solutions to models (5) and (6), respectively. We define the index sets  $\bar{I} = \{i \in I^s : \bar{d}_i^{x*} = 1\}$ ,  $\bar{R} = \{r \in R^s : \bar{d}_r^{y*} = 1\}$  and observe the following LP model:

$$\begin{aligned}
& \min \theta \\
& \text{s. t.} \\
& \sum_{j=1}^n \lambda_j x_{ij} \leq \theta x_{io} \quad \forall i \in I^f \cup \bar{I} \\
& \sum_{j=1}^n \lambda_j y_{rj} \geq y_{ro} \quad \forall r \in R^f \cup \bar{R} \\
& \sum_{j=1}^n \lambda_j = 1
\end{aligned} \tag{7}$$

It can be easily verified that  $(\bar{\theta}_o^{in*}, \boldsymbol{\lambda}^{in*})$  is the optimal solution of model (7). Let us now write the dual problem of (7) as follows:

$$\begin{aligned}
& \max \varphi = \sum_{r \in R^f \cup \bar{R}} u_r y_{ro} - w_0 \\
& \text{s. t.} \\
& \sum_{i \in I^f \cup \bar{I}} v_i x_{io} = 1 \\
& \sum_{r \in R^f \cup \bar{R}} u_r y_{rj} - w_0 - \sum_{i \in I^f \cup \bar{I}} v_i x_{ij} \leq 0 \quad \forall j \\
& v_i, u_r \geq 0 \quad i \in I^f \cup \bar{I}, r \in R^f \cup \bar{R} \\
& w_0 \text{ free in sign}
\end{aligned} \tag{8}$$

Suppose that  $(\varphi^*, \mathbf{u}'^*, \mathbf{v}'^*, w_0'^*)$  is the optimal solution of model (8), where  $\mathbf{u}'^* = (\dots, u_r'^*, \dots)_{r \in R^f \cup \bar{R}}$  and  $\mathbf{v}'^* = (\dots, v_i'^*, \dots)_{i \in I^f \cup \bar{I}}$ . By virtue of the strong duality theorem of LP (Bazaraa, Jarvis, & Sherali,

2010), we result in  $\bar{\theta}_o^{in*} = \varphi^*$ . Let  $\hat{d}_i^x = \begin{cases} 1, & i \in \bar{I} \\ 0, & i \in I^s - \bar{I} \end{cases}$ ,  $\hat{d}_r^y = \begin{cases} 1, & r \in \bar{R} \\ 0, & r \in R^s - \bar{R} \end{cases}$ ,  $\hat{u}_r^{in} = \begin{cases} u_r^*, & r \in R^f \cup \bar{R} \\ 0, & \text{otherwise} \end{cases}$ ,  $\hat{v}_i^{in} = \begin{cases} v_i^*, & i \in I^f \cup \bar{I} \\ 0, & \text{otherwise} \end{cases}$ ,  $\hat{w}_0 = w_0^*$  and  $\hat{\varphi} = \varphi^*$ . A straightforward verification indicates that  $(\hat{\varphi}, \hat{u}^{in}, \hat{v}^{in}, \hat{d}^x, \hat{d}^y, \hat{w}_0)$  is a feasible solution of model (5) and hence  $\bar{\theta}_o^{in*} = \hat{\varphi} \leq \bar{\varphi}_o^{in*}$ , which completes the proof. ■

The selecting DEA models (5) and (6) include a binary variable associated with each selective factor. Therefore, if there are  $|I^s \cup R^s|$  selective factors, then mathematically many different subsets ( $k$ -combinations) from the selective factors, i.e.,  $\Pi = \sum_{k=1}^{|I^s \cup R^s|} \binom{|I^s \cup R^s|}{k}$ , can be drawn while we are seeking those which meet the generic rule of thumb. Note that each combination of  $|I^s \cup R^s|$  leads to a subset of selected factors which is able to provide us with a given set of efficiency scores. The following theorem proves that the subset of the selective factors is identified by solving models (5) and (6).

**Theorem 2.** Let  $e_o^{p*}, p = 1, \dots, \Pi$ , be the efficiency score of DMU<sub>o</sub> associated with the  $p^{th}$  subset (combination) of  $|I^s \cup R^s|$  selective factors, which is obtained by the conventional input-oriented models (1) or (3). We have (i)  $\bar{\varphi}_o^{in*} = \max\{e_o^{p*} : p = 1, \dots, \Pi\}$  and (ii)  $\bar{\theta}_o^{in*} = \min\{e_o^{p*} : p = 1, \dots, \Pi\}$  where  $\bar{\varphi}_o^{in*}$  and  $\bar{\theta}_o^{in*}$  are the optimal objective values of (5) and (6), respectively.

**Proof.** (i) Let  $(\bar{\varphi}_o^{in*}, \mathbf{u}^{in*}, \mathbf{v}^{in*}, \mathbf{d}^{x*}, \mathbf{d}^{y*}, w_0^{in*})$  be the optimal solution of model (5). It is clear that  $\bar{\varphi}_o^{in*} \in \{e_o^{p*} : p = 1, \dots, \Pi\}$ . In order to prove the theorem, without loss of generality, let  $e_o^{q*} = \max\{e_o^{p*} : p = 1, \dots, \Pi\}$  and on the contrary  $\bar{\varphi}_o^{in*} < e_o^{q*}$ . We assume that  $I^q (\subseteq I^s)$  and  $R^q (\subseteq R^s)$  are respectively the index sets of selected inputs and outputs, which leads to the maximum efficiency score  $e_o^{q*}$ . In what follows, we shall think of the following multiplier DEA problem:

$$\begin{aligned} \max e_o^q &= \sum_{r \in R^f \cup R^q} u_r y_{ro} - w_0 \\ \text{s. t.} & \\ \sum_{i \in I^f \cup I^q} v_i x_{io} &= 1 \\ \sum_{r \in R^f \cup R^q} u_r y_{rj} - w_0 - \sum_{i \in I^f \cup I^q} v_i x_{ij} &\leq 0 \quad \forall j \\ v_i, u_r &\geq 0 \quad i \in I^f \cup I^q, r \in R^f \cup R^q \\ w_0 &\text{free in sign} \end{aligned} \quad (9)$$

Suppose that  $(e_o^{q*}, \mathbf{u}^*, \mathbf{v}^*, w_0^*)$  is the optimal solution of model (9), where  $\mathbf{u}^* = (\dots, u_r^*, \dots)_{r \in R^f \cup R^q}$  and  $\mathbf{v}^* = (\dots, v_i^*, \dots)_{i \in I^f \cup I^q}$ . Obviously, the vector  $(\hat{\varphi}^{in}, \hat{u}^{in}, \hat{v}^{in}, \hat{d}^x, \hat{d}^y, \hat{w}_0^{in})$ , whose components are defined as follows, is a feasible solution of model (5).

$$\begin{aligned} \hat{d}_i^x &= \begin{cases} 1, & i \in I^q \\ 0, & i \in I^s - I^q \end{cases}, \hat{d}_r^y = \begin{cases} 1, & r \in I^q \\ 0, & r \in I^s - I^q \end{cases}, \hat{u}_r^{in} = \begin{cases} u_r^*, & r \in R^f \cup R^q \\ 0, & \text{otherwise} \end{cases}, \hat{v}_i^{in} = \begin{cases} v_i^*, & i \in I^f \cup I^q \\ 0, & \text{otherwise} \end{cases}, \\ \hat{w}_0^{in} &= w_0^*, \text{ and } \hat{\varphi}^{in} = e_o^{q*}. \end{aligned}$$

The objective value of this feasible solution, i.e.  $\hat{\varphi}^{in} = e_o^{q*}$ , is greater than  $\bar{\varphi}_o^{in*}$ , which occasions a contradiction with the optimality of  $(\bar{\varphi}_o^{in*}, \mathbf{u}^{in*}, \mathbf{v}^{in*}, \mathbf{d}^{x*}, \mathbf{d}^{y*}, w_0^{in*})$  for model (5).

(ii) Let  $(\bar{\theta}_o^{in*}, \boldsymbol{\lambda}^{in*}, \mathbf{d}^{x*}, \mathbf{d}^{y*})$  be the optimal solution of model (6); it is obvious that  $\bar{\theta}_o^{in*} \in \{e_o^{p*} : p = 1, \dots, \Pi\}$ . If  $e_o^{q*} = \min\{e_o^{p*} : p = 1, \dots, \Pi\}$  and  $\bar{\theta}_o^{in*} > e_o^{q*}$ , then we reach a contradiction. There is no loss of generality in assuming that in the selection process which leads to the minimum

efficiency score  $e_o^{q*}$ , the index sets of the selected inputs and outputs are  $I^q$  and  $R^q$ , respectively. We now deem the following model to continue the proof:

$$\begin{aligned}
& \min \theta \\
& \text{s. t.} \\
& \sum_{j=1}^n \lambda_j x_{ij} \leq \theta x_{io} \quad \forall i \in I^f \cup I^q \\
& \sum_{j=1}^n \lambda_j y_{rj} \geq y_{ro} \quad \forall r \in R^f \cup R^q \\
& \sum_{j=1}^n \lambda_j = 1
\end{aligned} \tag{10}$$

Clearly,  $\theta^* = e_o^{q*}$  is the optimal objective value of model (10). Let  $(\theta^*, \lambda^*)$  be the optimal solution of model (10) and consider the vector  $(\hat{\theta}, \hat{\lambda}, \hat{d}^x, \hat{d}^y)$ , whose components are defined as  $\hat{d}_i^x = \begin{cases} 1, & i \in I^q \\ 0, & i \in I^s - I^q \end{cases}$ ,  $\hat{d}_r^y = \begin{cases} 1, & r \in R^q \\ 0, & r \in R^s - R^q \end{cases}$ ,  $\hat{\lambda} = \lambda^*$  and  $\hat{\theta} = e_o^{q*}$ . It is easy to observe that  $(\hat{\theta}, \hat{\lambda}, \hat{d}^x, \hat{d}^y)$  is a feasible solution of model (6) and its objective value, i.e.  $e_o^{q*}$ , is greater than  $\bar{\theta}_o^{in*}$ , which is a contradiction to the optimality of  $(\bar{\theta}_o^{in*}, \lambda^{in*}, d^{x*}, d^{y*})$ . ■

Due to the fact that the multiplier and envelopment forms of selecting models are no longer the primal-dual pair of problems, Theorem 2 shows that the developed models opt for the selective factors from two varying views that can be interesting to explore further in this study. In this respect, the multiplier model (5) seeks the factors out based upon enhancing performance (optimistic perspective) whereas the envelopment model (6) is intended to significantly enhance discrimination power of DEA (pessimistic perspective). The next section aims at extending the output-oriented version of selecting DEA models coupled with investigating some related properties.

### 3.2 Output-oriented selecting DEA model

Contrary to the input-oriented DEA models, the output-oriented DEA models make an attempt to increase outputs proportionally while keeping the existing inputs unchanged (Cooper et al., 2007). Given the notations used in Section 3, we formulate the output-oriented selecting DEA models in the multiplier and envelopment forms as follows:

$$\begin{aligned}
& \min \bar{\varphi}_o^{out} = \sum_{i=1}^m v_i^{out} x_{io} + w_0^{out} \\
& \text{s. t.} \\
& \sum_{r=1}^s u_r^{out} y_{ro} = 1 \\
& \sum_{r=1}^s u_r^{out} y_{rj} - \sum_{i=1}^m v_i^{out} x_{ij} - w_0^{out} \leq 0 \quad \forall j \\
& \sum_{i \in I^s} d_i^x + \sum_{r \in R^s} d_r^y \leq K - (|I^f| + |R^f|) \\
& \sum_{i \in I^s} d_i^x + \sum_{r \in R^s} d_r^y \geq 1 \\
& v_i^{out} \leq M d_i^x \quad \forall i \in I^s \\
& u_r^{out} \leq M d_r^y \quad \forall r \in R^s \\
& d_i^x, d_r^y \in \{0,1\} \quad \forall i \in I^s, r \in R^s \\
& v_i^{out}, u_r^{out} \geq 0 \quad \forall i, r \\
& w_0^{out} \text{ free in sign}
\end{aligned} \tag{11}$$

$$\begin{aligned}
& \max \bar{\theta}_o^{out} \\
& \text{s. t.} \\
& \sum_{j=1}^n \lambda_j^{out} x_{ij} \leq x_{io} \quad \forall i \in I^f \\
& \sum_{j=1}^n \lambda_j^{out} y_{rj} \geq \bar{\theta}_o^{out} y_{ro} \quad \forall r \in I^f \\
& \sum_{j=1}^n \lambda_j^{out} x_{ij} \leq x_{io} + M(1 - d_i^x) \quad \forall i \in I^s \\
& \sum_{j=1}^n \lambda_j^{out} y_{rj} \geq \bar{\theta}_o^{out} y_{ro} - M(1 - d_r^y) \quad \forall r \in R^s \\
& \sum_{i \in I^s} d_i^x + \sum_{r \in R^s} d_r^y \leq K - (|I^f| + |R^f|) \\
& \sum_{i \in I^s} d_i^x + \sum_{r \in R^s} d_r^y \geq 1 \\
& d_i^x, d_r^y \in \{0,1\} \quad \forall i \in I^s, r \in R^s \\
& \sum_{j=1}^n \lambda_j^{out} = 1
\end{aligned} \tag{12}$$

Let us now investigate the theorems for the above output-oriented models similar to those in Section 3.1<sup>1</sup>.

<sup>1</sup> Since the following theorems can be proved identical to Theorem 1 and Theorem 2 in the previous section, let us leave their proof to the readers.

**Theorem 3.** The optimal objective value of the multiplier model (11) is less than or equal to the optimal value of the envelopment model (12), i.e.  $\bar{\varphi}_o^{out*} \leq \bar{\theta}_o^{in*}$ .

**Theorem 4.** Let  $\xi_o^{p*}, p = 1, \dots, \Pi$ , be the optimal objective value of  $DMU_o$  calculated from the conventional output-oriented models (2) or (4) in terms of the  $p^{th}$  combination of selective factors. We have (i)  $\bar{\varphi}_o^{out*} = \min\{\xi_o^{p*} : p = 1, \dots, \Pi\}$  and (ii)  $\bar{\theta}_o^{out*} = \max\{\xi_o^{p*} : p = 1, \dots, \Pi\}$ .

It would be very interesting to study the relationship between input- and output-oriented selecting DEA models. In this regard, the following theorems prove that the multiplier and envelopment forms of the input- and output-oriented selecting DEA models are equivalent under CRS assumption.

**Theorem 5.** In the light of CRS assumption, (i) The input- and output-oriented multiplier selecting DEA models (5) and (11) are equivalent. (ii) The input- and output-oriented envelopment selecting DEA models (6) and (12) are equivalent.

**Proof.** (i) Let  $(\bar{\varphi}_o^{in*}, \mathbf{u}^{in*}, \mathbf{v}^{in*}, \mathbf{d}^{x*}, \mathbf{d}^{y*}) \in \mathbb{R}^{1+2(m+s)}$  be an optimal solution of model (5) under CRS assumption. Since  $\bar{\varphi}_o^{in*}$  is positive, it can be easily shown that  $(\frac{1}{\bar{\varphi}_o^{in*}}, \frac{\mathbf{u}^{in*}}{\bar{\varphi}_o^{in*}}, \frac{\mathbf{v}^{in*}}{\bar{\varphi}_o^{in*}}, \mathbf{d}^{x*}, \mathbf{d}^{y*})$  is a feasible solution for (11). If  $(\frac{1}{\bar{\varphi}_o^{in*}}, \frac{\mathbf{u}^{in*}}{\bar{\varphi}_o^{in*}}, \frac{\mathbf{v}^{in*}}{\bar{\varphi}_o^{in*}}, \mathbf{d}^{x*}, \mathbf{d}^{y*})$  is not optimal for model (11), then there is an alternative feasible solution  $(\hat{\varphi}_o^{out}, \hat{\mathbf{u}}^{out}, \hat{\mathbf{v}}^{out}, \hat{\mathbf{d}}^x, \hat{\mathbf{d}}^y)$  such that:

$$\sum_{i=1}^m v_i^{out} x_{io} = \sum_{i=1}^m \frac{v_i^{in*}}{\bar{\varphi}_o^{in*}} x_{io} = \frac{1}{\bar{\varphi}_o^{in*}} > \sum_{i=1}^m \hat{v}_i^{out} x_{ik} = \hat{\varphi}_o^{out} \quad (13)$$

It is plain to verify that the vector  $(\varphi_o, \mathbf{u}, \mathbf{v}, \mathbf{d}^x, \mathbf{d}^y) = (\frac{1}{\hat{\varphi}_o^{out}}, \frac{\hat{\mathbf{u}}^{out}}{\hat{\varphi}_o^{out}}, \frac{\hat{\mathbf{v}}^{out}}{\hat{\varphi}_o^{out}}, \hat{\mathbf{d}}^x, \hat{\mathbf{d}}^y)$  is a feasible solution of the input-oriented model (5) and its objective value is  $\sum_{r=1}^s \frac{\hat{u}_r^{out}}{\hat{\varphi}_o^{out}} y_{ro} = \frac{1}{\hat{\varphi}_o^{out}}$ . In the light of equation (13),  $\frac{1}{\hat{\varphi}_o^{out}} > \bar{\varphi}_o^{in*}$ , which contradicts the optimality of  $\bar{\varphi}_o^{in*}$ . In a similar manner, it can be shown that if  $(\bar{\varphi}_o^{out*}, \mathbf{u}^{out*}, \mathbf{v}^{out*}, \mathbf{d}^{x*}, \mathbf{d}^{y*})$  is an optimal solution of output-oriented model (11), then  $(\frac{1}{\bar{\varphi}_o^{out*}}, \frac{\mathbf{u}^{out*}}{\bar{\varphi}_o^{out*}}, \frac{\mathbf{v}^{out*}}{\bar{\varphi}_o^{out*}}, \mathbf{d}^{x*}, \mathbf{d}^{y*})$  is an optimal solution of input-oriented model (5).

(ii) Let  $(\bar{\theta}_o^{in*}, \boldsymbol{\lambda}^{in*}, \mathbf{d}^{x*}, \mathbf{d}^{y*})$  be an optimal solution of the input-oriented envelopment selecting DEA model (6). It is clear that  $(\frac{1}{\bar{\theta}_o^{in*}}, \frac{\boldsymbol{\lambda}^{in*}}{\bar{\theta}_o^{in*}}, \mathbf{d}^{x*}, \mathbf{d}^{y*})$  is a feasible solution of model (12). We prove that  $(\frac{1}{\bar{\theta}_o^{in*}}, \frac{\boldsymbol{\lambda}^{in*}}{\bar{\theta}_o^{in*}}, \mathbf{d}^{x*}, \mathbf{d}^{y*})$  is an optimal solution of model (12). Suppose, contrary to our claim that  $(\frac{1}{\bar{\theta}_o^{in*}}, \frac{\boldsymbol{\lambda}^{in*}}{\bar{\theta}_o^{in*}}, \mathbf{d}^{x*}, \mathbf{d}^{y*})$  is not optimal solution for model (12) and therefore there is a distinct optimal solution  $(\hat{\theta}_o^{out}, \hat{\boldsymbol{\lambda}}^{out}, \hat{\mathbf{d}}^x, \hat{\mathbf{d}}^y)$ , hence  $\hat{\theta}_o^{out} > \frac{1}{\bar{\theta}_o^{in*}}$ . Obviously,  $(\frac{1}{\hat{\theta}_o^{out}}, \frac{\hat{\boldsymbol{\lambda}}^{out}}{\hat{\theta}_o^{out}}, \hat{\mathbf{d}}^x, \hat{\mathbf{d}}^y)$  is feasible for input-oriented model (6) with an objective value which is less than the optimal objective value, i.e.,  $\bar{\theta}_o^{in*} > \frac{1}{\hat{\theta}_o^{out}}$  which is a contradiction. Likewise, we can conclude that each optimal solution of output-oriented model (12) is equivalent with an optimal solution of input-oriented model (6). ■

As mentioned earlier, the process of eventual decision-making on factor selection is intricate. Although it is supposed that problem examination from different angles simplifies the factor

selection, producing many conflicting results by means of these examinations leads to another problem. For example, regardless of multiplier and envelopment form, Theorem 5 emphasizes that both of the input- and output-oriented selecting DEA models provide identical results when the CRS condition is assumed. However, this does not necessarily observe in VRS assumptions.

Unlike conventional DEA models, the shape of PPS is not fixed and identical for both input and output orientations in our selecting models. In other words, the PPS constructed by the input-oriented selecting DEA model (either multiple or envelopment form) may differ from those created by the output-oriented ones. Therefore, the input- and output-oriented selecting DEA models might result in various sets of performance factors that can be disputable on the concept of PPS. Therefore, the decision-maker's view and characteristics embedded in the production process will be involved in the choice of orientations through the factor selection problem. In the next section, we propose a new approach to integrate input- and output-oriented selecting models for both multiplier and envelopment forms to constitute an identical PPS for both input and output orientations.

#### 4 Integrated input- and output-oriented selecting DEA models

The input- and output-oriented selecting DEA models presented in the previous sections are capable of selecting the performance factors, even though those selected factors may be different based on these models. This issue can be controversial when the performance evaluation has dissimilar sets of performance factors. In this section, we develop a pair of multiplier and envelopment selecting DEA models in which each one integrates both input- and output orientation models. The aim of the proposed models is to determine the same set of selective factors for both input- and output-oriented models. To do this, we firstly integrate both the conventional input- and output-oriented models to turn into a single model and then prove that the optimal solution of the integrated model leads to those of individual models and vice versa. Finally, we extend this idea to selecting DEA models with the aim of unifying factor selection in both orientations.

In the light of the input- and output-oriented multiplier models (1) and (2), we propose the following integrated model:

$$\begin{aligned}
& \max (\sum_{r=1}^s u_r^{in} y_{ro} - w_0^{in}) - (\sum_{i=1}^m v_i^{out} x_{io} + w_0^{out}) \\
& \text{s. t.} \\
& \sum_{i=1}^m v_i^{in} x_{io} = 1 \\
& \sum_{r=1}^s u_r^{out} y_{ro} = 1 \\
& \sum_{r=1}^s u_r^p y_{rj} - w_0^p - \sum_{i=1}^m v_i^p x_{ij} \leq 0 \quad \forall j, p = in, out \\
& v_i^p, u_r^p \geq 0 \quad \forall i, r, p = in, out \\
& w_0^p \text{ free in sign} \quad p = in, out
\end{aligned} \tag{14}$$

Let  $S_{in}$  and  $S_{out}$  be the feasible region of the input- and output-oriented models (1) and (2), respectively, i.e.,

$$\begin{aligned}
S_{in} &= \{(\mathbf{v}^{in}, \mathbf{u}^{in}, w_0^{in}) \mid \mathbf{v}^{in} \mathbf{x}_o = 1, \mathbf{u}^{in} \mathbf{Y} - w_0^{in} \mathbf{1}_n - \mathbf{v}^{in} \mathbf{X} \leq \mathbf{0}_n, \mathbf{v}^{in} \geq \mathbf{0}_m, \mathbf{u}^{in} \geq \mathbf{0}_s, w_0^{in} \text{ free in sign}\} \\
S_{out} &= \{(\mathbf{v}^{out}, \mathbf{u}^{out}, w_0^{out}) \mid \mathbf{u}^{out} \mathbf{y}_o = 1, \mathbf{u}^{out} \mathbf{Y} - \mathbf{v}^{out} \mathbf{X} - w_0^{out} \mathbf{1}_n \leq \mathbf{0}_n, \mathbf{v}^{out} \geq \mathbf{0}_m, \mathbf{u}^{out} \geq \mathbf{0}_s, w_0^{out} \text{ free in sign}\}
\end{aligned}$$

where  $\mathbf{v}^p = (v_1^p, \dots, v_m^p)$ ,  $\mathbf{u}^p = (u_1^p, \dots, u_s^p)$  for  $p = in, out$ , are the input and output weights;  $\mathbf{X} = [x_{ij}] \in \mathbb{R}^{n \times m}$ , and  $\mathbf{Y} = [y_{rj}] \in \mathbb{R}^{s \times n}$  are the matrices associated with the input and output sets,

respectively;  $\mathbf{1}_n$  is a column vector with all components equal to one, and  $\mathbf{0}_n$  is the origin in  $\mathbb{R}^n$ . It is obvious that the feasible region of model (14) is  $S_{in} \cup S_{out}$ . Moreover, the objective function of (14) is  $\varphi_o^{in} - \varphi_o^{out}$ , where  $\varphi_o^{in}$  and  $\varphi_o^{out}$  are respectively the objective function of the input- and output-oriented models (1) and (2). The following theorem shows that the integrated model (14) is equivalent with models (1) and (2).

**Theorem 6.**  $(\mathbf{v}^{in*}, \mathbf{u}^{in*}, w_0^{in*}, \mathbf{v}^{out*}, \mathbf{u}^{out*}, w_0^{out*})$  is an optimal solution of the integrated model (14) if and only if  $(\mathbf{v}^{in*}, \mathbf{u}^{in*}, w_0^{in*})$  and  $(\mathbf{v}^{out*}, \mathbf{u}^{out*}, w_0^{out*})$  are the optimal solutions of models (1) and (2), respectively.

**Proof.** Let  $(\mathbf{v}^{in*}, \mathbf{u}^{in*}, w_0^{in*}, \mathbf{v}^{out*}, \mathbf{u}^{out*}, w_0^{out*})$  be the optimal solution of model (14). Therefore,  $(\mathbf{v}^{in*}, \mathbf{u}^{in*}, w_0^{in*})$  and  $(\mathbf{v}^{out*}, \mathbf{u}^{out*}, w_0^{out*})$  are the feasible solutions of models (1) and (2), respectively. On the contrary, if  $(\mathbf{v}^{in*}, \mathbf{u}^{in*}, w_0^{in*})$  is not the optimal solution of model (1), then there is a feasible solution  $(\hat{\mathbf{v}}^{in}, \hat{\mathbf{u}}^{in}, \hat{w}_0^{in}) \neq (\mathbf{v}^{in*}, \mathbf{u}^{in*}, w_0^{in*})$  such that  $\sum_{r=1}^s u_r^{in*} y_{rk} - w_0^{in*} < \sum_{r=1}^s \hat{u}_r^{in} y_{rk} - \hat{w}_0^{in}$ . In what follows,  $(\hat{\mathbf{v}}^{in}, \hat{\mathbf{u}}^{in}, \hat{w}_0^{in}, \mathbf{v}^{out*}, \mathbf{u}^{out*}, w_0^{out*})$  is a feasible solution of model (14) and  $\sum_{r=1}^s u_r^{in*} y_{rk} - \sum_{i=1}^m v_i^{out*} x_{ik} - w_0^{in*} - w_0^{out*} < \sum_{r=1}^s \hat{u}_r^{in} y_{rk} - \sum_{i=1}^m v_i^{out*} x_{ik} - \hat{w}_0^{in} - w_0^{out*}$  which is in contradiction to the optimality of  $(\mathbf{v}^{in*}, \mathbf{u}^{in*}, w_0^{in*}, \mathbf{v}^{out*}, \mathbf{u}^{out*}, w_0^{out*})$ . Likewise, the optimality of  $(\mathbf{v}^{out*}, \mathbf{u}^{out*}, w_0^{out*})$  is concluded. Moreover, the converse is straightforward. ■

It should be noted that the first component associated with the objective function of model (14) is always equal to or smaller than one and the second one is invariably equal to or larger than one. Thereby the optimal value of the objective function of (14) will be zero if and only if the value of both components are equal to one. Subsequently, we arrive at the following corollaries as direct results of Theorem 6.

**Corollary 1.** The optimal objective value of model (14) is always non-positive.

**Corollary 2.**  $DMU_o$  is an efficient DMU if and only if the optimal objective value of model (14) is zero. In the following, we employ the above idea to integrate the envelopment form models (3) and (4).

$$\begin{aligned}
& \min \theta_o^{in} - \theta_o^{out} \\
& \text{s. t.} \\
& \sum_{j=1}^n \lambda_j^{in} x_{ij} \leq \theta_o^{in} x_{io} \quad \forall i \\
& \sum_{j=1}^n \lambda_j^{in} y_{rj} \geq y_{ro} \quad \forall r \\
& \sum_{j=1}^n \lambda_j^{out} x_{ij} \leq x_{io} \quad \forall i \\
& \sum_{j=1}^n \lambda_j^{out} y_{rj} \geq \theta_o^{out} y_{ro} \quad \forall r \\
& \sum_{j=1}^n \lambda_j^p = 1 \quad p = in, out
\end{aligned} \tag{15}$$

As can be seen, the objective function of model (15) is the difference of the objective functions of models (3) and (4), and the constraint sets of model (15) are derived from the union of the constraint sets of models (3) and (4). The following theorem shows that optimal solution of model (15) gives the optimal solution of models (3) and (4) and vice versa.

**Theorem 7.** The necessary and sufficient conditions for  $(\theta_o^{in*}, \boldsymbol{\lambda}^{in*}, \theta_o^{out*}, \boldsymbol{\lambda}^{out*})$  to be an optimal solution of the integrated model (15) are the optimality of  $(\theta_o^{in*}, \boldsymbol{\lambda}^{in*})$  and  $(\theta_o^{out*}, \boldsymbol{\lambda}^{out*})$  for models (3) and (4), respectively.

**Proof.** In order to prove the necessary condition, let the optimal solution of model (15) be  $(\theta_o^{in*}, \lambda^{in*}, \theta_o^{out*}, \lambda^{out*})$ . It is not difficult to show that  $(\theta_o^{in*}, \lambda^{in*})$  and  $(\theta_o^{out*}, \lambda^{out*})$  are the feasible solutions of models (3) and (4), respectively. To obtain a contradiction, we think of two cases (i)  $(\delta, \mu)$  is the optimal solution of model (3) such that  $\delta < \theta_o^{in*}$ , and (ii)  $(\delta, \mu)$  is the optimal solution of model (4) such that  $\delta > \theta_o^{out*}$ . Regarding case (i),  $(\delta, \mu, \theta_o^{out*}, \lambda^{out*})$  as a feasible solution can result in  $\delta - \theta_o^{out*} < \theta_o^{in*} - \theta_o^{out*}$ , which is impossible to occur due to the optimality condition for  $(\theta_o^{in*}, \lambda^{in*}, \theta_o^{out*}, \lambda^{out*})$ , and in case (ii)  $(\theta_o^{in*}, \lambda^{in*}, \delta, \mu)$  is a feasible a feasible solution of model (15) and  $\theta_o^{in*} - \delta < \theta_o^{in*} - \theta_o^{out*}$  which also cannot occur thanks to the optimality of  $(\theta_o^{in*}, \lambda^{in*}, \theta_o^{out*}, \lambda^{out*})$ . Consequently, we observe a contradiction in each case, which proves the necessary condition. Analogously, the sufficient condition can be proved. ■

Let us now come to the following corollaries as evident properties of model (15).

**Corollary 3.** The optimal objective value of model (15) is always non-positive.

**Corollary 4.** DMU<sub>o</sub> is efficient if and only if the optimal objective value of model (15) is zero.

By the use of the above idea, we propose two new multiplier- and envelopment-based selecting DEA models to identify the identical set of selective factors for both input- and output-orientations. At first, consider the following multiplier-based selecting DEA model.

$$\begin{aligned}
& \max (\sum_{r=1}^s u_r^{in} y_{ro} - w_0^{in}) - (\sum_{i=1}^m v_i^{out} x_{io} + w_0^{out}) \\
& \text{s. t.} \\
& \sum_{i=1}^m v_i^{in} x_{io} = 1 \\
& \sum_{r=1}^s u_r^{out} y_{ro} = 1 \\
& \sum_{r=1}^s u_r^p y_{rj} - w_0^p - \sum_{i=1}^m v_i^p x_{ij} \leq 0 \quad \forall j, p = in, out \\
& \sum_{i \in I^s} d_i^x + \sum_{r \in R^s} d_r^y \leq K - (|I^f| + |R^f|) \\
& \sum_{i \in I^s} d_i^x + \sum_{r \in R^s} d_r^y \geq 1 \\
& v_i^p \leq M d_i^x \quad \forall i \in I^s, p = in, out \\
& u_r^p \leq M d_r^y \quad \forall r \in R^s, p = in, out \\
& d_i^x, d_r^y \in \{0, 1\} \quad \forall i \in I^s, \forall r \in R^s \\
& v_i^p, u_r^p \geq 0 \quad \forall i, r, p = in, out \\
& w_0^p \text{ free in sign} \quad p = in, out.
\end{aligned} \tag{16}$$

The objective function of model (16) contains two distinct components where the first component  $\sum_{r=1}^s u_r^{in} y_{ro} - w_0^{in}$  and the second component  $\sum_{i=1}^m v_i^{out} x_{io} + w_0^{out}$  are the objective function of models (5) and (11), respectively. The set of variables and constraints of model (16) are the union of all variables and constraints of both models (5) and (11). The binary variables  $d_i^x$  and  $d_r^y$  associated with the selective factors in model (16) are assumed to be in common. Note that these binary variables are crucial in model (16) because they play the role of a bridge between models (5) and (11). The optimal value of the binary variables may differ when solving models (5) and (11) separately, although model (16) enables us to obtain an identical value for each binary variable, resulting in a unique selection of performance factors. The following theorem shows the relationship between the integrated selecting DEA model (16) and distinct models (5) and (11).

**Theorem 8.** Let  $(v^{in*}, u^{in*}, w_0^{in*}, v^{out*}, u^{out*}, w_0^{out*}, d^{x*}, d^{y*})$  be the optimal solutions of model (16).

Then, (i)  $\sum_{r=1}^s u_r^{in*} y_{ro} - w_0^{in*}$  is a lower bound for the optimal objective value of model (5) (ii)  $\sum_{i=1}^m v_i^{out*} x_{io} + w_0^{out*}$  is an upper bound for the optimal objective value for model (11).

**Proof.** Since  $(\mathbf{v}^{in*}, \mathbf{u}^{in*}, w_0^{in*}, \mathbf{d}^{x*}, \mathbf{d}^{y*})$  and  $(\mathbf{v}^{out*}, \mathbf{u}^{out*}, w_0^{out*}, \mathbf{d}^{x*}, \mathbf{d}^{y*})$  are feasible solutions for models (5) and (11), respectively, the proof is straightforward. ■

Similarly, the envelopment form of the selecting model can be studied, and the proposed selecting model, which integrates the input- and output-oriented selecting models (6) and (12) is expressed below:

$$\begin{aligned}
& \min \bar{\theta}_o^{in} - \bar{\theta}_o^{out} \\
& \text{s. t.} \\
& \sum_{j=1}^n \lambda_j^{in} x_{ij} \leq \bar{\theta}_o^{in} x_{io} & \forall i \in I^f \\
& \sum_{j=1}^n \lambda_j^{in} y_{rj} \geq y_{ro} & \forall r \in R^f \\
& \sum_{j=1}^n \lambda_j^{in} x_{ij} \leq \bar{\theta}_o^{in} x_{io} + M(1 - d_i^x) & \forall i \in I^s \\
& \sum_{j=1}^n \lambda_j^{in} y_{rj} \geq y_{ro} - M(1 - d_r^y) & \forall r \in R^s \\
& \sum_{j=1}^n \lambda_j^{out} x_{ij} \leq x_{io} & \forall i \in I^f \\
& \sum_{j=1}^n \lambda_j^{out} y_{rj} \geq \bar{\theta}_o^{out} y_{ro} & \forall r \in R^f \\
& \sum_{j=1}^n \lambda_j^{out} x_{ij} \leq x_{io} + M(1 - d_i^x) & \forall i \in I^s \\
& \sum_{j=1}^n \lambda_j^{out} y_{rj} \geq \bar{\theta}_o^{out} y_{ro} - M(1 - d_r^y) & \forall r \in R^s \\
& \sum_{i \in I^s} d_i^x + \sum_{r \in R^s} d_r^y \leq K - (|I^f| + |R^f|) \\
& \sum_{i \in I^s} d_i^x + \sum_{r \in R^s} d_r^y \geq 1 \\
& d_i^x, d_r^y \in \{0,1\} & \forall i \in I^s, r \in R^s \\
& \sum_{j=1}^n \lambda_j^p = 1 & p = in, out
\end{aligned} \tag{17}$$

Contrary to model (16), the input- and output-oriented models (6) and (12) are aggregated to make the above model. Over and above, the common binary variables  $d_i^x$  and  $d_r^y$  are also introduced to join the respective constraints of models (6) and (12) and these variables assist model (17) to seek the most influential factors of both input- and output-orientations. At present, it is of interest to explore the details of the relationship between the developed model (17) and models (6) and (12).

**Theorem 9.** Let  $(\bar{\theta}_o^{in*}, \boldsymbol{\lambda}^{in*}, \bar{\theta}_o^{out*}, \boldsymbol{\lambda}^{out*}, \mathbf{d}^{x*}, \mathbf{d}^{y*})$  be the optimal solution of model (17). Then, (i)  $\bar{\theta}_o^{in*}$  is an upper bound for the optimal objective value of model (6) (ii)  $\bar{\theta}_o^{out*}$  is a lower bound for the optimal objective value of model (12).

**Proof.** Similar to the proof of Theorem 8 (omitted). ■

It is worth noting that the multiplier integrated model (16) as the aggregation of the multiplier models (5) and (11) yields an optimistic results for efficiency and, during an economic boom, can be applied to decisional problems related to performance analysis, while the envelopment integrated model (17) builds on the envelopment models (6) and (12) aiming at improving the discrimination power of our assessment with the focus on a pessimistic view during an economic recession.

The proposed models (5), (6), (11), (12), (16), and (17) have been formulated at individual level analysis, that is, factor selection is made in a way that is of interest to the DMU under evaluation. However, many situations with a centralized organizational structure such as sector, region or industry levels consist of a homogenous set of firms that are managed by a top management team. In essence, the performance and efficient allocation of resources across different firms play an important role for an aggregate entity such as sector, industry or country. In the ensuing section, we propose some new selecting DEA models, which consider factor selection apropos of the industry, region or country level analysis (so-called structural efficiency hereafter).



## 5 Selective factors for structural efficiency

For policy-makers and regulators, an industry is a group of companies and organisations that are active in a particular field and their performance individually or in combination has a direct impact on economy and society. The isolated assessment of a group of firms, despite the importance and applicability, may not provide deeper insights into the performance of industry. The small- and medium-size firms are often interested in evaluating itself in comparison with their rivals at the industry level and as a consequence of such evaluation, the firms enable to find out that some firms are relatively efficient in terms of their effective resource allocation and importantly this practice helps inefficient firms to gain a terrific picture of their future potential change for enhancing the performance and competitiveness in the domestic and international market. However, the perspective and concerns of policy-makers are wholly different and they often attempt to set out more comfortable rule and regulations associated to the industry in regard to economic uncertainty, technological advances and the altering skills requirement in the workforce. By focusing on the efficiencies of both firms and industry, not only it can improve efficiency and resource allocation of respective stakeholders, but also helps the economic growth of the country. To achieve the strategic goals of an industry or organization, it needs to effectively benchmark, make improvement and make informed decisions and recommendations for the future. Farrell (1957, pp. 261–262) firstly presented the *structural efficiency* of an industry as the technical efficiency of its component organizations measured by the output-weighted average of the individual efficiencies. This thrust was extended and generalized by several researchers to make explicit the relationship between the production technology of the group and that of an individual DMU (see e.g., Førsund & Hjalmarsson (1979); Li & Cheng (2007) and Karagiannis (2015)). The aim of this section is to extend an alternative approach to choose the performance factors when one takes structural efficiency of the industry into account in a multiple inputs and multiple outputs technology.

Assume that there are  $n$  DMUs in a certain industry in which all DMUs have an identical production technology. Let  $\mathbf{X} = [x_{ij}]_{m \times n}$  and  $\mathbf{Y} = [y_{rj}]_{s \times n}$  be the input and output matrixes, respectively, and  $(\mathbf{x}_j, \mathbf{y}_j)$  represents the input-output vector associated with DMU $_j$ . The total input-output of the industry is defined as  $(\mathbf{x}, \mathbf{y}) = (\sum_{j=1}^n \mathbf{x}_j, \sum_{j=1}^n \mathbf{y}_j)$  and production technology set of the industry is the sum of the individual technology of all DMUs (for more details, see Li & Cheng, 2007). More precisely, the question is whether the industry has been able to produce the highest possible level of productions and services (overall outputs) using their available internal and external resources (overall inputs). To answer this question, the modified DEA approach can be applied from the input- or output-oriented aspects. From the input-oriented aspect, we explore whether the industry output  $\mathbf{y}$  can produce by using a smaller aggregate input bundle or not. Having focused on the Shephard–Farrell concept of radial efficiency measures, we identify to what extent proportional reduction in the input bundle of the industry would be made when keeping the aggregate output unchanged (Ray & Hu, 1997). The output-oriented aspect of the industry assessment can be analogously observed. The concept of structural efficiency of the industry that makes use of the technical efficiency of the aggregate unit  $(\mathbf{x}, \mathbf{y})$  is a proper tool in the literature. The following multiplier and envelopment forms of DEA-based models in both orientations can measure the structural efficiency of a given

industry under the VRS assumption (see Ray & Hu, 1997 and Karagiannis, 2015 for preliminary models).

### Multiplier VRS models for structural efficiency

Input-oriented	Output-oriented
$\max \phi^{in} = \mathbf{U}^{in} \mathbf{y} - nW_0^{in}$	$\min \phi^{out} = \mathbf{V}^{out} \mathbf{x} + nW_0^{out}$
s. t.	s. t.
$\mathbf{V}^{in} \mathbf{x} = 1$	$\mathbf{U}^{out} \mathbf{y} = 1$
$\mathbf{U}^{in} \mathbf{Y} - W_0^{in} \mathbf{1}_n - \mathbf{V}^{in} \mathbf{X} \leq \mathbf{0}_n$	(18) $\mathbf{U}^{out} \mathbf{Y} - W_0^{out} \mathbf{1}_n - \mathbf{V}^{out} \mathbf{X} \leq \mathbf{0}_n$
$\mathbf{V}^{in} \geq \mathbf{0}_m$	$\mathbf{V}^{out} \geq \mathbf{0}_m$
$\mathbf{U}^{in} \geq \mathbf{0}_s$	$\mathbf{U}^{out} \geq \mathbf{0}_s$
$W_0^{in}$ free in sign	$W_0^{out}$ free in sign

### Envelopment VRS models for structural efficiency

Input-oriented	Output-oriented
$\min \Theta^{in}$	$\max \Theta^{out}$
s. t.	s. t.
$\Lambda^{in} \mathbf{X} \leq \Theta^{in} \mathbf{x}$	(20) $\Lambda^{out} \mathbf{X} \leq \mathbf{x}$
$\Lambda^{in} \mathbf{Y} \geq \mathbf{y}$	$\Lambda^{out} \mathbf{Y} \geq \Theta^{out} \mathbf{y}$
$\mathbf{1}_n \Lambda^{in} = n$	$\mathbf{1}_n \Lambda^{out} = n$

At present, let us study the relationship between individual and structural efficiencies. To this end, we need to present the following two theorems, which focus on the optimal solution of the input-oriented multiplier models (1) and (18) and the output-oriented multiplier models (2) and (19), respectively.

**Theorem 10.** Let  $(\mathbf{u}^{in*}, \mathbf{v}^{in*}, w_0^{in*}, \phi_0^{in*})$  and  $(\mathbf{U}^{in*}, \mathbf{V}^{in*}, W_0^{in*}, \phi^{in*})$  be the optimal solutions of models (1) and (18), respectively. If  $\mathbf{x}_o$  is a strictly positive vector, then  $(\mathbf{U}^{in*} \mathbf{y}_j - W_0^{in*}) \leq \phi_0^{in*}$ .

**Proof.** Given the normalization constraint of model (18), we have  $\mathbf{V}^{in*} \mathbf{x} = \mathbf{V}^{in*} \sum_{j=1}^n \mathbf{x}_j = \sum_{j=1}^n \mathbf{V}^{in*} \mathbf{x}_j = \mathbf{V}^{in*} \mathbf{x}_o + \sum_{j=1, j \neq o}^n \mathbf{V}^{in*} \mathbf{x}_j = 1$ , and consequently  $\mathbf{V}^{in*} \mathbf{x}_o \leq 1$ . On the other hand,  $\mathbf{V}^{in*}$  is a semi-positive and  $\mathbf{x}_o$  is a strictly positive vector and then  $0 < \mathbf{V}^{in*} \mathbf{x}_o \leq 1$ . Evidently,  $\frac{1}{\mathbf{V}^{in*} \mathbf{x}_o} (\mathbf{U}^{in*}, \mathbf{V}^{in*}, W_0^{in*})$  is a feasible solution for model (1) because it satisfies the normalization constraint  $\sum_{i=1}^m \frac{v_i^{in*}}{\mathbf{V}^{in*} \mathbf{x}_o} x_{io} = 1$  and the common constraints  $\sum_{r=1}^s \frac{u_r^{in*}}{\mathbf{V}^{in*} \mathbf{x}_o} y_{rj} - \frac{W_0^{in*}}{\mathbf{V}^{in*} \mathbf{x}_o} - \sum_{i=1}^m \frac{v_i^{in*}}{\mathbf{V}^{in*} \mathbf{x}_o} x_{ij} \leq 0$  for  $j = 1, \dots, n$  in model (1). Given the feasibility of  $\frac{1}{\mathbf{V}^{in*} \mathbf{x}_o} (\mathbf{U}^{in*}, \mathbf{V}^{in*}, W_0^{in*})$  and  $\mathbf{V}^{in*} \mathbf{x}_o \leq 1$ , we can conclude that  $\mathbf{U}^{in*} \mathbf{y}_o - W_0^{in*} \leq \frac{\mathbf{U}^{in*} \mathbf{y}_o - W_0^{in*}}{\mathbf{V}^{in*} \mathbf{x}_o} \leq \mathbf{u}^{in*} \mathbf{y}_o - w_0^{in*} = \phi_0^{in*}$ , which completes the proof. ■

**Theorem 11.** Let  $(\mathbf{U}^{out*}, \mathbf{V}^{out*}, W_0^{out*}, \phi^{out*})$  and  $(\mathbf{u}^{out*}, \mathbf{v}^{out*}, w_0^{out*}, \phi_0^{out*})$ , be the optimal solutions of models (19) and (2), respectively. If  $\mathbf{y}_o$  is a strictly positive vector, then  $\frac{\mathbf{V}^{out*} \mathbf{x}_o + W_0^{out*}}{\mathbf{U}^{out*} \mathbf{y}_o} \geq \phi_0^{out*}$ .

**Proof.** This is alike to the proof of Theorem 10. If  $(\mathbf{U}^{out*}, \mathbf{V}^{out*}, W_0^{out*})$  is the optimal solution of the output-oriented model (19), then  $\frac{1}{\mathbf{U}^{out*} \mathbf{y}_o} (\mathbf{U}^{out*}, \mathbf{V}^{out*}, W_0^{out*})$  is a feasible solution for model (2) and we have  $\frac{\mathbf{V}^{out*} \mathbf{x}_o + W_0^{out*}}{\mathbf{U}^{out*} \mathbf{y}_o} \geq \mathbf{v}^{out*} \mathbf{x}_o + w_0^{out*} = \phi_0^{out*}$ , which completes the proof. ■

Similarly, there exist an identical relationships and arguments between the envelopment models (1), (4), (20) and (21).

The production technology associated with each industry and its constituent firms hinges on the performance factors. In this regard, the exclusion (inclusion) of the factors from (in) performance analysis affects the structural and individual efficiencies. Accordingly, we can extend the selecting DEA models (16) and (17) in a way that aggregates the input- and output-oriented models of structural efficiency. To do so, we suggest the following MBLPs models associated with the multiplier and envelopment forms:

<b>Integrated selecting DEA model for structural efficiency (Multiplier form)</b>	<b>Integrated selecting DEA model for structural efficiency (Envelopment form)</b>
$\max (U^{in}y - nW_0^{in}) - (V^{out}x + nW_0^{out})$ <p>s. t.</p> $V^{in}x = 1$ $V^{out}y = 1$ $U^{in}Y - W_0^{in}1_n - V^{in}X \leq 0_n$ $U^{out}Y - W_0^{out}1_n - V^{out}X \leq 0_n$ $1_{ I^f }d^x + 1_{ R^s }d^y \leq K - ( I^f  +  R^f )$ $1_{ I^s }d^x + 1_{ R^s }d^y \geq 1$ $V^{in}, V^{out} \leq Md^x$ $U^{in}, U^{out} \leq Md^y$ $V^{in}, V^{out} \geq 0_m$ $U^{in}, U^{out} \geq 0_s$ $d^x \text{ and } d^y \text{ are binary vectors}$ $W_0^{in}, W_0^{out} \text{ are free in sign}$	$\min \bar{\Theta}^{in} - \bar{\Theta}^{out}$ <p>s. t.</p> $\Lambda^{in}X^f \leq \bar{\Theta}^{in}x^f$ $\Lambda^{in}Y^f \geq y^f$ $\Lambda^{in}X^s \leq \bar{\Theta}^{in}x^s + M(1_{ I^s } - d^x)$ $\Lambda^{in}Y^s \geq y^s - M(1_{ R^s } - d^y)$ $\Lambda^{out}X^f \leq x^f$ $\Lambda^{out}Y^f \geq \bar{\Theta}^{out}y^f$ $\Lambda^{out}X^s \leq x^s + M(1_{ I^s } - d^x)$ $\Lambda^{out}Y^s \geq \bar{\Theta}^{out}y^s - M(1_{ R^s } - d^y)$ $1_{ I^s }d^x + 1_{ R^s }d^y \leq K - ( I^f  +  R^f )$ $1_{ I^s }d^x + 1_{ R^s }d^y \geq 1$ $d^x \text{ and } d^y \text{ are binary vectors}$ $1_n\Lambda^{in} = n, 1_n\Lambda^{out} = n$
(22)	(23)

where the input matrix  $X$  is portioned into fixed input matrix ( $X^f$ ) and selective input matrix ( $X^s$ ), i.e.  $X = [X^f, X^s]$ , and  $Y = [Y^f, Y^s]$  indicates the output matrix where  $Y^f$  and  $Y^s$  denote the fixed output and selective output matrixes, respectively.

We find models (22) and (23) helpful when the selection of factors must be made to render the structural efficiency of the industry. Moreover, such models can be applied when ties are encountered at the individual selecting DEA models.

## 6 A case study

The Europe 2020 strategy lays emphasis on sustainable and all-encompassing economy growth and development with the aim of tackling the structural barriers, and improving the competitiveness and productivity across Europe<sup>1</sup>. Europe has five interconnected objectives to reach its strategic vision. The first objective is to create “a highly competitive social market economy aiming at full employment and social progress...<sup>2</sup>”. The second key and long-standing objective of the EU is to boost spending on research and development (R&D) to 3% of the EU's gross domestic product (GDP) in order to help countries to improve productivity and living standards. The third integral objective is set to get around barriers to climate change and energy both in Europe and globally. In this regard, the Europe 2020 targets are to reduce greenhouse gas emissions by 20%, increase energy efficiency by 20%, and

<sup>1</sup> <https://ec.europa.eu>

<sup>2</sup> Article 3 of the Treaty on European Union (TEU).

generate 20% energy from renewable resources by 2020. The fourth objective is concerned with the EU education and training that should be managed, executed and monitored by each member state, and the EU as a supporting role funds different educational programs such as Erasmus to encourage EU countries to develop quality education. Currently, the EU has more than 120 million people at stake of poverty or social exclusion and the objective of EU leaders is to raise public consciousness and lift at least 20 million people out of poverty.

In complex EU environment, it is extremely difficult to check the overall strategy implemented and its coherence. Therefore, the development of performance indicators helps the EU examine whether activities has contributed to the achievement of specific objectives. Despite a myriad of performance indicators, the consideration of each indicator independently to track progress of the strategic performance of the EU members has not only been often contentious debate but also seems like a herculean task. In addition, there is also a tendency to focus heavily on some recognised indicators of performance such as GDP per capita rather than less familiar indicators. To deal with the existing partial analysis of the EU countries, we take 50 performance indicators into account to provide a better picture of each country's performance. This study looks into the relative performance of 28 EU countries in the year of 2013 using a DEA methodology. In an effort to estimate a best practice frontier, DEA envelops observed data as tightly as possible to measure the relative efficiency of each country to the estimated best practice. All data used in the analysis are taken from the World Bank database<sup>1</sup>.

The specification of the DEA model is required to define the inputs and outputs in which inputs have to be minimized and outputs have to be maximized. One of the key assumptions in conventional DEA models is that input and output observations are non-negative data but some data of the performance indicators are characterized by negative values. To combat the negativity, we adopt Scheel (2001)'s method in which the inverse of absolute values of negative outputs are deemed as inputs and the inverse of the absolute values of negative inputs are deemed as outputs. Finally, the data set involves fourteen inputs and thirty-six outputs. The descriptive statistics and definitions of input and output variables are given in Table 1 and Table 2, respectively.

The 28 EU countries are evaluated using conventional DEA models in both input and output orientations under CRS and VRS assumptions and the results show that all countries are efficient. This bespeaks the lowest level of discrimination power which is, of course, unable to provide insights into the strategic alignment of the EU. Technically, this problem may arise due to violating the generic rule of thumb  $n > f(m, n)$ . If Cooper et al., (2007)'s suggestion is utilized i.e.,  $n < \max \{3(m + s), m \times s\}$ , we have  $28 < 504$ . To combat this problem, it is inevitable to reduce the number of performance factors the extent to which the desirable and formative results are observed. To this end, we apply the proposed models to identify the upmost factors in our analysis. We assume that "CO2 emissions",  $x_2$ , and "GDP",  $y_9$ , are the fixed input and output, respectively, and the remaining 48 factors are regarded as selective factors (13 inputs and 35 outputs) because it seems to be of importance in relation to the EU strategy. Therefore, we need to make  $K - (|I^f| + |R^f|) = \min\{\lfloor n/3 \rfloor, \lfloor 2\sqrt{n} \rfloor\} - (|I^f| + |R^f|) = 7$  factors selection among 48 selective factors where  $n =$

---

<sup>1</sup> <http://data.worldbank.org/>. The raw data used in this paper is available upon request.

28,  $|I^f| = 1$  and  $|R^f| = 1$ . In the ensuing discussion, we seek the performance factors from individual and industrial viewpoints and in turn measure the relative individual and structural efficiencies.

**Table 1.** Descriptive statistics and definitions of input factors

Inputs	Definition	Min	Mean	Median	Sd.	Max
$x_1$	Adolescent fertility rate (births per 1,000 women ages 15-19)	4.357	12.701	10.780	8.845	42.673
$x_2$	CO2 emissions (metric tons per capita)	3.518	7.346	6.225	3.703	18.498
$x_3$	External debt stocks, total (DOD, 1.E+05US\$)	5.15E+10	8.77E+10	8.77E+10	9.69E+09	1.24E+11
$x_4$	Fertility rate, total (births per woman)	1.210	1.546	1.500	0.216	1.990
$x_5$	Foreign direct investment, net inflows (BoP, current US\$)	3.04E-12	8.12E-10	2.31E-10	1.78E-09	9.62E-09
$x_6$	GDP growth (annual %)	0.169	1.097	0.641	1.232	5.254
$x_7$	Imports of goods and services (% of GDP)	21.286	61.119	52.523	32.982	158.613
$x_8$	Military expenditure (% of GDP)	0.419	1.342	1.247	0.509	2.358
$x_9$	Mortality rate, under-5 (per 1,000 live births)	2.600	4.568	4.100	1.784	10.000
$x_{10}$	Population density (people per sq. km of land area)	3.013	172.040	110.926	247.214	1331.147
$x_{11}$	Population growth (annual %)	0.433	3.933	1.993	5.994	30.148
$x_{12}$	Poverty headcount ratio at \$1.90 a day (2011 PPP) (% of population)	1.00E-11	4.41E-01	2.00E-01	4.97E-01	1.70E+00
$x_{13}$	Poverty headcount ratio at national poverty lines (% of population)	9.700	17.927	17.927	2.758	25.100
$x_{14}$	Prevalence of HIV, total (% of population ages 15-49)	0.100	0.200	0.200	0.110	0.600

**Table 2.** Descriptive statistics and description of output factors

Outputs	Definition	Min	Mean	Median	Sd.	Max
$y_1$	Agriculture, forestry, and fishing, value added (% of GDP)	0.277	2.334	2.189	1.239	5.397
$y_2$	Births attended by skilled health staff (% of total)	98.300	99.322	99.322	0.484	100.000
$y_3$	Contraceptive prevalence, any methods (% of women ages 15-49)	65.100	68.300	68.300	1.111	73.000
$y_4$	Domestic credit provided by financial sector (% of GDP)	41.658	133.049	139.555	64.436	306.105
$y_5$	Electric power consumption (MWh per capita)	2.49E+03	6.43E+03	5.41E+03	3.23E+03	1.55E+04
$y_6$	Energy use (100kg of oil equivalent per capita)	1.60E+03	3.34E+03	2.90E+03	1.44E+03	7.31E+03
$y_7$	Exports of goods and services (% of GDP)	19.988	65.071	56.739	38.115	190.629
$y_8$	Forest area (sq. km)	3.50E+00	1.00E+05	2.89E+04	2.31E+05	1.24E+06
$y_9$	GDP (current million US\$)	1.01E+10	6.85E+11	2.33E+11	9.88E+11	3.75E+12
$y_{10}$	GNI per capita, Atlas method (current US\$)	7.61E+03	3.27E+04	2.51E+04	1.90E+04	7.55E+04
$y_{11}$	GNI per capita, PPP (current international \$)	1.63E+04	3.40E+04	2.96E+04	1.10E+04	6.38E+04
$y_{12}$	GNI, Atlas method (current million US\$)	9.65E+09	6.85E+11	2.13E+11	9.98E+11	3.82E+12
$y_{13}$	GNI, PPP (current international million \$)	1.27E+10	6.82E+11	2.78E+11	9.41E+11	3.73E+12
$y_{14}$	Gross capital formation (% of GDP)	11.601	20.140	19.499	3.685	27.810
$y_{15}$	High-technology exports (% of manufactured exports)	4.309	12.796	10.439	7.349	38.553
$y_{16}$	Immunization, measles (% of children ages 12-23 months)	86.000	94.929	95.500	3.432	99.000
$y_{17}$	Income share held by lowest 20%	5.700	7.689	7.694	1.124	9.500
$y_{18}$	Industry (including construction), value added (% of GDP)	10.006	22.378	22.763	5.637	32.850
$y_{19}$	Inflation, GDP deflator (annual %)	0.281	1.144	0.754	1.229	6.333
$y_{20}$	Life expectancy at birth, total (years)	73.915	79.388	80.539	2.795	83.078

$y_{21}$	Merchandise trade (% of GDP)	31.462	89.029	72.917	44.625	176.693
$y_{22}$	Mobile cellular subscriptions (per 100 people)	99.068	124.258	122.402	15.975	162.337
$y_{23}$	Net barter terms of trade index (2000 = 100)	42.660	96.005	93.538	20.955	176.992
$y_{24}$	Personal remittances, received (current million US\$)	2.46E+08	4.01E+09	1.97E+09	5.37E+09	2.44E+10
$y_{25}$	Population, total (million)	4.26E+05	1.86E+07	9.75E+06	2.28E+07	8.06E+07
$y_{26}$	Primary completion rate, total (% of relevant age group)	81.984	97.906	97.906	4.291	108.675
$y_{27}$	Revenue, excluding grants (% of GDP)	2.162	33.726	35.996	10.125	47.524
$y_{28}$	School enrolment, primary (% gross)	95.315	102.166	100.679	4.723	120.429
$y_{29}$	School enrolment, primary and secondary (gross), gender parity index (GPI)	0.966	1.006	1.005	0.031	1.098
$y_{30}$	School enrolment, secondary (% gross)	90.806	112.053	107.469	15.893	160.925
$y_{31}$	Statistical Capacity score (Overall average)	84.444	86.667	86.667	0.664	88.889
$y_{32}$	Surface area (sq. km)	3.20E+02	4.30E+05	8.55E+04	1.42E+06	7.74E+06
$y_{33}$	Tax revenue (% of GDP)	1.287	20.123	22.018	6.620	33.820
$y_{34}$	Time required to start a business (days)	2.500	13.018	10.750	8.905	38.500
$y_{35}$	Total debt service (% of exports of goods, services and primary income)	16.704	49.483	49.483	120.086	95.459
$y_{36}$	Urban population growth (annual %)	0.374	4.633	1.776	45.903	24.762

### 6.1 Performance factor selection with an individual view

We first make an effort to select the performance factors by those proposed selecting DEA models that focus the degree to which a unit can preserve its efficiency score in favorable respect.

For the sake of brevity, let us consider Austria ( $DMU_1$ ). Here, we employ the input- and output-oriented selecting DEA models, i.e., models (5), (6), (11), and (12), in relation to the multiplier and envelopment forms under assumptions of CRS and VRS. The results are shown in Table 3 in this respect.

**Table 3.** The results of factor selection by different disintegrated selecting DEA models for Austria

Binary Variable	Input-oriented				Output-oriented				Binary Variable	Input-oriented				Output-oriented			
	Model (5)		Model (6)		Model (11)		Model (12)			Model (5)		Model (6)		Model (11)		Model (12)	
	CRS	VRS	CRS	VRS	CRS	VRS	CRS	VRS		CRS	VRS	CRS	VRS	CRS	VRS	CRS	VRS
$d_1^x$	0	0	0	0	0	1	0	1	$d_{12}^y$	0	0	1	0	0	0	1	0
$d_2^x$	Fixed input				Fixed input				$d_{13}^y$	0	0	1	0	0	0	1	0
$d_3^x$	0	0	0	0	0	0	0	1	$d_{14}^y$	0	0	0	0	0	0	0	0
$d_4^x$	0	0	0	0	0	0	0	1	$d_{15}^y$	0	0	1	0	0	0	1	0
$d_5^x$	0	0	0	0	0	0	0	1	$d_{16}^y$	0	0	0	0	0	0	0	0
$d_6^x$	0	0	0	0	0	0	0	0	$d_{17}^y$	0	1	0	0	0	0	0	0
$d_7^x$	0	0	0	0	0	0	0	0	$d_{18}^y$	0	0	0	0	0	0	0	0
$d_8^x$	0	0	0	0	0	0	0	1	$d_{19}^y$	1	0	0	0	1	1	0	0
$d_9^x$	0	0	0	0	0	0	0	1	$d_{20}^y$	1	0	0	0	1	1	0	0
$d_{10}^x$	0	1	0	0	0	0	0	0	$d_{21}^y$	1	0	0	1	1	1	0	0
$d_{11}^x$	0	0	0	0	0	0	0	0	$d_{22}^y$	0	0	0	0	0	0	0	0
$d_{12}^x$	0	0	1	1	0	0	1	0	$d_{23}^y$	1	1	0	0	1	1	0	0
$d_{13}^x$	0	0	0	0	0	0	0	0	$d_{24}^y$	1	0	1	0	1	1	1	0
$d_{14}^x$	1	0	0	0	1	0	0	0	$d_{25}^y$	0	1	1	0	0	0	1	0
$d_1^y$	0	0	0	1	0	0	0	0	$d_{26}^y$	0	0	0	0	0	0	0	0
$d_2^y$	0	0	0	0	0	0	0	0	$d_{27}^y$	0	0	1	0	0	0	1	0
$d_3^y$	0	0	0	1	0	0	0	0	$d_{28}^y$	0	1	0	0	0	0	0	0

$d_4^y$	0	0	0	0	0	0	0	0	$d_{29}^y$	0	0	0	0	0	0	0	0
$d_5^y$	0	0	0	0	0	0	0	0	$d_{30}^y$	0	1	0	0	0	0	0	0
$d_6^y$	0	1	0	0	0	0	0	0	$d_{31}^y$	0	0	0	0	0	1	0	0
$d_7^y$	0	0	0	1	0	0	0	0	$d_{32}^y$	0	0	0	0	0	0	0	0
$d_8^y$	1	0	0	0	1	0	0	0	$d_{33}^y$	0	0	0	0	0	0	0	0
$d_9^y$	Fixed output				Fixed output				$d_{34}^y$	0	0	0	1	0	0	0	0
$d_{10}^y$	0	0	0	0	0	0	0	0	$d_{35}^y$	0	0	0	0	0	0	0	0
$d_{11}^y$	0	0	0	0	0	0	0	0	$d_{36}^y$	0	0	0	1	0	0	0	1
<i>Optimal objective function value</i>										1	1	0.17	0.27	1	1	5.68	2.38

As can be spotted from Table 3,  $x_{14}$ ,  $y_8, y_{19}, y_{20}, y_{21}, y_{23}$  and  $y_{24}$  are selected identically by both the input- and output-oriented multiplier models under CRS assumption, and  $x_{12}, y_{12}, y_{13}, y_{15}, y_{24}, y_{25}$  and  $y_{27}$  are the identified performance factors as the result of solving both the input- and output-oriented envelopment models under CRS assumption. These results subject to the CRS assumption are in line with Theorem 5.

To validate Theorem 1 and Theorem 3, let us look into the optimal values of the objective functions of models (5), (6), (11) and (12) apropos as appeared in the last row of Table 3. In the view of Theorem 1, the optimal objective function value calculated from the envelopment model (6), irrespective of RTS assumptions, is always less than or equal to the efficiencies obtained from the multiplier model (5). In this regard, the optimal objective function value for Austria as result of solving model (6) for CRS and VRS is **0.176 and 0.270** while it is equal to **1** for model (5). Contrary to the input-oriented case in Theorem 1, Theorem 3 stiffly focuses on the output-orientation models, in which the optimal objective function value of the multiplier model (11) is less than or equal to the optimal objective function value of the envelopment model (12).

Whilst the total combination for identifying the seven factors for each unit is  $\Pi = \sum_{k=1}^7 \binom{|I^S \cup R^S|}{k} = \binom{48}{1} + \binom{48}{2} + \dots + \binom{48}{7} = 87,825,940$  that is obviously very time consuming, Theorem 2 shows that the optimal objective values of models (5) and (6) are  $\bar{\varphi}_j^{in*} = \max\{e_j^{p*} : p = 1, \dots, \Pi\}$  and  $\bar{\theta}_j^{in*} = \min\{e_j^{p*} : p = 1, \dots, \Pi\}$ , respectively. The similar argument can be observed for models (2) or (4) according to Theorem 4. Table 4 exhibits the total frequencies of each factor as well as its rank in parentheses after solving the multiplier and envelopment [selecting] models for both input- and output orientations. Consequently, we pick out the seven factors ( $K - (|I^f| + |R^f|) = 7$ ) with the highest frequencies in Table 4 and the resultant selection of four disintegrated models are minutely summarised in Table 5.

**Table 4.** Frequencies of a performance factor and its rank derived from disintegrated selecting DEA models

Binary Variable	Input-oriented				Output-oriented				Binary Variable	Input-oriented				Output-oriented			
	Model (5)		Model (6)		Model (11)		Model (12)			Model (5)		Model (6)		Model (11)		Model (12)	
	CRS	VRS	CRS	VRS	CRS	VRS	CRS	VRS		CRS	VRS	CRS	VRS	CRS	VRS	CRS	VRS
$x_1$	8 (7 <sup>th</sup> )	2 (31 <sup>st</sup> )	14 (5 <sup>th</sup> )	7 (10 <sup>th</sup> )	8 (7 <sup>th</sup> )	7 (7 <sup>th</sup> )	14 (5 <sup>th</sup> )	16 (3 <sup>rd</sup> )	$y_{12}$	12 (6 <sup>th</sup> )	1 (39 <sup>th</sup> )	21 (1 <sup>st</sup> )	3 (15 <sup>th</sup> )	12 (6 <sup>th</sup> )	1 (23 <sup>rd</sup> )	21 (1 <sup>st</sup> )	16 (3 <sup>rd</sup> )
$x_2$	Fixed input				Fixed input				$y_{13}$	1 (29 <sup>th</sup> )	2 (31 <sup>st</sup> )	4 (15 <sup>th</sup> )	2 (18 <sup>th</sup> )	1 (29 <sup>th</sup> )	1 (23 <sup>rd</sup> )	4 (15 <sup>th</sup> )	3 (18 <sup>th</sup> )
$x_3$	3 (13 <sup>th</sup> )	1 (39 <sup>th</sup> )	13 (6 <sup>th</sup> )	2 (18 <sup>th</sup> )	3 (13 <sup>th</sup> )	7 (7 <sup>th</sup> )	13 (6 <sup>th</sup> )	16 (3 <sup>rd</sup> )	$y_{14}$	0 (39 <sup>th</sup> )	1 (39 <sup>th</sup> )	0 (29 <sup>th</sup> )	2 (18 <sup>th</sup> )	0 (39 <sup>th</sup> )	0 (36 <sup>th</sup> )	0 (29 <sup>th</sup> )	0 (26 <sup>th</sup> )
$x_4$	1 (29 <sup>th</sup> )	1 (39 <sup>th</sup> )	9 (8 <sup>th</sup> )	3 (15 <sup>th</sup> )	1 (29 <sup>th</sup> )	11 (6 <sup>th</sup> )	9 (8 <sup>th</sup> )	13 (7 <sup>th</sup> )	$y_{15}$	1 (29 <sup>th</sup> )	2 (31 <sup>st</sup> )	2 (19 <sup>th</sup> )	1 (31 <sup>st</sup> )	1 (29 <sup>th</sup> )	0 (36 <sup>th</sup> )	2 (19 <sup>th</sup> )	1 (20 <sup>th</sup> )
$x_5$	2 (16 <sup>th</sup> )	5 (15 <sup>th</sup> )	21 (1 <sup>st</sup> )	10 (7 <sup>th</sup> )	2 (16 <sup>th</sup> )	7 (7 <sup>th</sup> )	21 (1 <sup>st</sup> )	23 (1 <sup>st</sup> )	$y_{16}$	0 (39 <sup>th</sup> )	9 (4 <sup>th</sup> )	0 (29 <sup>th</sup> )	2 (18 <sup>th</sup> )	0 (39 <sup>th</sup> )	1 (23 <sup>rd</sup> )	0 (29 <sup>th</sup> )	0 (26 <sup>th</sup> )

$x_6$	4 (10 <sup>th</sup> )	3 (27 <sup>th</sup> )	6 (11 <sup>th</sup> )	15 (3 <sup>rd</sup> )	4 (10 <sup>th</sup> )	3 (11 <sup>th</sup> )	6 (11 <sup>th</sup> )	7 (11 <sup>th</sup> )	$y_{17}$	0 (39 <sup>th</sup> )	4 (19 <sup>th</sup> )	0 (29 <sup>th</sup> )	2 (18 <sup>th</sup> )	0 (39 <sup>th</sup> )	1 (23 <sup>rd</sup> )	0 (29 <sup>th</sup> )	0 (26 <sup>th</sup> )
$x_7$	2 (16 <sup>th</sup> )	2 (31 <sup>st</sup> )	20 (3 <sup>rd</sup> )	8 (8 <sup>th</sup> )	2 (16 <sup>th</sup> )	3 (11 <sup>th</sup> )	20 (3 <sup>rd</sup> )	21 (2 <sup>nd</sup> )	$y_{18}$	0 (39 <sup>th</sup> )	6 (10 <sup>th</sup> )	0 (29 <sup>th</sup> )	2 (18 <sup>th</sup> )	0 (39 <sup>th</sup> )	0 (36 <sup>th</sup> )	0 (29 <sup>th</sup> )	0 (26 <sup>th</sup> )
$x_8$	1 (29 <sup>th</sup> )	2 (31 <sup>st</sup> )	3 (17 <sup>th</sup> )	4 (12 <sup>th</sup> )	1 (29)	1 (23 <sup>rd</sup> )	3 (17 <sup>th</sup> )	8 (9 <sup>th</sup> )	$y_{19}$	20 (3 <sup>rd</sup> )	5 (15 <sup>th</sup> )	0 (29 <sup>th</sup> )	1 (31 <sup>st</sup> )	20 (3 <sup>rd</sup> )	21 (4 <sup>th</sup> )	0 (29 <sup>th</sup> )	0 (26 <sup>th</sup> )
$x_9$	4 (10 <sup>th</sup> )	0 (44 <sup>th</sup> )	5 (13 <sup>th</sup> )	0 (44 <sup>th</sup> )	4 (10 <sup>th</sup> )	2 (17 <sup>th</sup> )	5 (13 <sup>th</sup> )	8 (9 <sup>th</sup> )	$y_{20}$	19 (4 <sup>th</sup> )	10 (2 <sup>nd</sup> )	0 (29 <sup>th</sup> )	1 (31 <sup>st</sup> )	19 (4 <sup>th</sup> )	22 (3 <sup>rd</sup> )	0 (29 <sup>th</sup> )	0 (26 <sup>th</sup> )
$x_{10}$	3 (13 <sup>th</sup> )	4 (19 <sup>th</sup> )	9 (8 <sup>th</sup> )	12 (6 <sup>th</sup> )	3 (13 <sup>th</sup> )	1 (23 <sup>rd</sup> )	9 (8 <sup>th</sup> )	7 (11 <sup>th</sup> )	$y_{21}$	18 (5 <sup>th</sup> )	4 (19 <sup>th</sup> )	0 (29 <sup>th</sup> )	1 (31 <sup>st</sup> )	18 (5 <sup>th</sup> )	21 (4 <sup>th</sup> )	0 (29 <sup>th</sup> )	0 (26 <sup>th</sup> )
$x_{11}$	2 (16 <sup>th</sup> )	6 (10 <sup>th</sup> )	6 (11 <sup>th</sup> )	8 (8 <sup>th</sup> )	2 (16 <sup>th</sup> )	2 (17 <sup>th</sup> )	6 (11 <sup>th</sup> )	9 (8 <sup>th</sup> )	$y_{22}$	4 (10 <sup>th</sup> )	8 (5 <sup>th</sup> )	0 (29 <sup>th</sup> )	0 (44 <sup>th</sup> )	4 (10 <sup>th</sup> )	3 (11 <sup>th</sup> )	0 (29 <sup>th</sup> )	0 (26 <sup>th</sup> )
$x_{12}$	2 (16 <sup>th</sup> )	7 (9 <sup>th</sup> )	19 (4 <sup>th</sup> )	15 (3 <sup>rd</sup> )	2 (16 <sup>th</sup> )	1 (23 <sup>rd</sup> )	19 (4 <sup>th</sup> )	16 (3 <sup>rd</sup> )	$y_{23}$	25 (2 <sup>nd</sup> )	13 (1 <sup>st</sup> )	0 (29 <sup>th</sup> )	1 (31 <sup>st</sup> )	25 (2 <sup>nd</sup> )	25 (2 <sup>nd</sup> )	0 (29 <sup>th</sup> )	0 (26 <sup>th</sup> )
$x_{13}$	1 (29 <sup>th</sup> )	0 (44 <sup>th</sup> )	12 (7 <sup>th</sup> )	0 (44 <sup>th</sup> )	1 (29 <sup>th</sup> )	1 (23 <sup>rd</sup> )	12 (7 <sup>th</sup> )	6 (13 <sup>th</sup> )	$y_{24}$	27 (1 <sup>st</sup> )	5 (15 <sup>th</sup> )	4 (15 <sup>th</sup> )	1 (31 <sup>st</sup> )	27 (1 <sup>st</sup> )	26 (1 <sup>st</sup> )	4 (15 <sup>th</sup> )	4 (15 <sup>th</sup> )
$x_{14}$	5 (8 <sup>th</sup> )	2 (31 <sup>st</sup> )	1 (23 <sup>rd</sup> )	4 (12 <sup>th</sup> )	5 (8 <sup>th</sup> )	3 (11 <sup>th</sup> )	1 (23 <sup>rd</sup> )	1 (20 <sup>th</sup> )	$y_{25}$	2 (16 <sup>th</sup> )	4 (19 <sup>th</sup> )	5 (13 <sup>th</sup> )	1 (31 <sup>st</sup> )	2 (16 <sup>th</sup> )	0 (36 <sup>th</sup> )	5 (13 <sup>th</sup> )	4 (15 <sup>th</sup> )
$y_1$	5 (8 <sup>th</sup> )	8 (5 <sup>th</sup> )	3 (17 <sup>th</sup> )	23 (1 <sup>st</sup> )	5 (8 <sup>th</sup> )	2 (17 <sup>th</sup> )	3 (17 <sup>th</sup> )	3 (18 <sup>th</sup> )	$y_{26}$	2 (16 <sup>th</sup> )	8 (5 <sup>th</sup> )	0 (29 <sup>th</sup> )	1 (31 <sup>st</sup> )	2 (16 <sup>th</sup> )	2 (17 <sup>th</sup> )	0 (29 <sup>th</sup> )	0 (26 <sup>th</sup> )
$y_2$	0 (39 <sup>th</sup> )	4 (19 <sup>th</sup> )	1 (23 <sup>rd</sup> )	1 (31 <sup>st</sup> )	0 (39 <sup>th</sup> )	1 (23 <sup>rd</sup> )	1 (23 <sup>rd</sup> )	1 (20 <sup>th</sup> )	$y_{27}$	2 (16 <sup>th</sup> )	10 (2 <sup>nd</sup> )	1 (23 <sup>rd</sup> )	1 (31 <sup>st</sup> )	2 (16 <sup>th</sup> )	0 (36 <sup>th</sup> )	1 (23 <sup>rd</sup> )	1 (20 <sup>th</sup> )
$y_3$	0 (39 <sup>th</sup> )	6 (10 <sup>th</sup> )	1 (23 <sup>rd</sup> )	14 (5 <sup>th</sup> )	0 (39 <sup>th</sup> )	0 (36 <sup>th</sup> )	1 (23 <sup>rd</sup> )	1 (20 <sup>th</sup> )	$y_{28}$	0 (39 <sup>th</sup> )	8 (5 <sup>th</sup> )	0 (29 <sup>th</sup> )	2 (18 <sup>th</sup> )	0 (39 <sup>th</sup> )	3 (11 <sup>th</sup> )	0 (29 <sup>th</sup> )	0 (26 <sup>th</sup> )
$y_4$	1 (29 <sup>th</sup> )	2 (31 <sup>st</sup> )	1 (23 <sup>rd</sup> )	0 (44 <sup>th</sup> )	1 (29 <sup>th</sup> )	1 (23 <sup>rd</sup> )	1 (23 <sup>rd</sup> )	0 (26 <sup>th</sup> )	$y_{29}$	1 (29 <sup>th</sup> )	6 (10 <sup>th</sup> )	0 (29 <sup>th</sup> )	0 (44 <sup>th</sup> )	1 (29 <sup>th</sup> )	2 (17 <sup>th</sup> )	0 (29 <sup>th</sup> )	0 (26 <sup>th</sup> )
$y_5$	2 (16 <sup>th</sup> )	0 (44 <sup>th</sup> )	2 (19 <sup>th</sup> )	3 (15 <sup>th</sup> )	2 (16 <sup>th</sup> )	0 (36 <sup>th</sup> )	2 (19 <sup>th</sup> )	0 (26 <sup>th</sup> )	$y_{30}$	0 (39 <sup>th</sup> )	4 (19 <sup>th</sup> )	0 (29 <sup>th</sup> )	1 (31 <sup>st</sup> )	0 (39 <sup>th</sup> )	1 (23 <sup>rd</sup> )	0 (29 <sup>th</sup> )	0 (26 <sup>th</sup> )
$y_6$	0 (39 <sup>th</sup> )	3 (27 <sup>th</sup> )	2 (19 <sup>th</sup> )	2 (18 <sup>th</sup> )	0 (39 <sup>th</sup> )	0 (36 <sup>th</sup> )	2 (19 <sup>th</sup> )	0 (26 <sup>th</sup> )	$y_{31}$	1 (29 <sup>th</sup> )	6 (10 <sup>th</sup> )	0 (29 <sup>th</sup> )	1 (31 <sup>st</sup> )	1 (29 <sup>th</sup> )	6 (10 <sup>th</sup> )	0 (29 <sup>th</sup> )	0 (26 <sup>th</sup> )
$y_7$	1 (29 <sup>th</sup> )	3 (27 <sup>th</sup> )	0 (29 <sup>th</sup> )	4 (12 <sup>th</sup> )	1 (29 <sup>th</sup> )	0 (36 <sup>th</sup> )	0 (29 <sup>th</sup> )	0 (26 <sup>th</sup> )	$y_{32}$	2 (16 <sup>th</sup> )	0 (44 <sup>th</sup> )	2 (19 <sup>th</sup> )	2 (18 <sup>th</sup> )	2 (16 <sup>th</sup> )	0 (36 <sup>th</sup> )	2 (19 <sup>th</sup> )	1 (20 <sup>th</sup> )
$y_8$	2 (16 <sup>th</sup> )	0 (44 <sup>th</sup> )	8 (10 <sup>th</sup> )	7 (10 <sup>th</sup> )	2 (16 <sup>th</sup> )	0 (36 <sup>th</sup> )	8 (10 <sup>th</sup> )	6 (13 <sup>th</sup> )	$y_{33}$	2 (16 <sup>th</sup> )	4 (19 <sup>th</sup> )	0 (29 <sup>th</sup> )	2 (18 <sup>th</sup> )	2 (16 <sup>th</sup> )	1 (23 <sup>rd</sup> )	0 (29 <sup>th</sup> )	0 (26 <sup>th</sup> )
$y_9$	Fixed output				Fixed output				$y_{34}$	3 (13 <sup>th</sup> )	4 (19 <sup>th</sup> )	0 (29 <sup>th</sup> )	2 (18 <sup>th</sup> )	3 (13 <sup>th</sup> )	0 (36 <sup>th</sup> )	0 (29 <sup>th</sup> )	0 (26 <sup>th</sup> )
$y_{10}$	2 (16 <sup>th</sup> )	1 (39 <sup>th</sup> )	0 (29 <sup>th</sup> )	2 (18 <sup>th</sup> )	2 (16 <sup>th</sup> )	1 (23 <sup>rd</sup> )	0 (29 <sup>th</sup> )	0 (26 <sup>th</sup> )	$y_{35}$	1 (29 <sup>th</sup> )	2 (31 <sup>st</sup> )	0 (29 <sup>th</sup> )	2 (18 <sup>th</sup> )	1 (29 <sup>th</sup> )	3 (11 <sup>th</sup> )	0 (29 <sup>th</sup> )	0 (26 <sup>th</sup> )
$y_{11}$	0 (39 <sup>th</sup> )	3 (27 <sup>th</sup> )	0 (29 <sup>th</sup> )	1 (31 <sup>st</sup> )	0 (39 <sup>th</sup> )	0 (36 <sup>th</sup> )	0 (29 <sup>th</sup> )	0 (26 <sup>th</sup> )	$y_{36}$	2 (16 <sup>th</sup> )	5 (15 <sup>th</sup> )	1 (23 <sup>rd</sup> )	17 (2 <sup>nd</sup> )	2 (16 <sup>th</sup> )	2 (17 <sup>th</sup> )	1 (23 <sup>rd</sup> )	4 (15 <sup>th</sup> )

**Table 5.** Seven factors selected by disintegrated models

Models		Selected factors and their frequencies						
Model (5)	CRS	$y_{24}$ (27)	$y_{23}$ (25)	$y_{19}$ (20)	$y_{20}$ (19)	$y_{21}$ (18)	$y_{12}$ (12)	$x_1$ (8)
	VRS	$y_{23}$ (13)	$y_{20}$ (10)	$y_{27}$ (10)	$y_{16}$ (9)	$y_1$ (8)	$y_{22}$ (8)	$y_{26}$ (8)
Model (6)	CRS	$x_5$ (21)	$y_{12}$ (21)	$x_7$ (20)	$x_{12}$ (19)	$x_1$ (14)	$x_3$ (13)	$x_{13}$ (12)
	VRS	$y_1$ (23)	$y_{36}$ (17)	$x_6$ (15)	$x_{12}$ (15)	$y_3$ (14)	$x_{10}$ (12)	$x_5$ (10)
Model (11)	CRS	$y_{24}$ (27)	$y_{23}$ (25)	$y_{19}$ (20)	$y_{20}$ (19)	$y_{21}$ (18)	$y_{12}$ (12)	$x_1$ (8)
	VRS	$y_{24}$ (26)	$y_{23}$ (25)	$y_{20}$ (22)	$y_{19}$ (21)	$y_{21}$ (21)	$x_4$ (11)	$x_1$ (7)
Model (12)	CRS	$x_5$ (21)	$y_{12}$ (21)	$x_7$ (20)	$x_{12}$ (19)	$x_1$ (14)	$x_3$ (13)	$x_{13}$ (12)
	VRS	$x_5$ (23)	$x_7$ (21)	$x_1$ (16)	$x_3$ (16)	$x_{12}$ (16)	$y_{12}$ (16)	$x_4$ (13)

For instance, under CRS assumption the input- and output-oriented multiplier forms are composed of six selective outputs  $y_{12}, y_{19}, y_{20}, y_{21}, y_{23}, y_{24}$  and one selective input  $x_1$ . However, under VRS assumption the input- and output-oriented multiplier forms consist of the dissimilar factors in which the former case includes seven selective outputs  $y_1, y_{16}, y_{20}, y_{22}, y_{23}, y_{26}, y_{27}$  and latter one embraces five selective outputs  $y_{19}, y_{20}, y_{21}, y_{23}, y_{24}$  and two selective inputs  $x_1, x_4$ . Notice that through selecting inputs and outputs if ties occurs, the choice may be made by either a decision-maker or applying the selecting approach developed in Subsection 6.2.

As an alternative matter, it is envisaged that the identical seven factors are identified for the input- and output-oriented selecting DEA models under a certain RTS. However, Table 5 bespeaks that the results culminate in different selected performance factors. To deal with the inconsistency and complexity, the integrated selecting DEA models (16) and (17) can be appropriately considered in a way that combine input and output orientation models. Table 6 reports the corresponding results of multiplier and envelopment forms in relation to DMU<sub>1</sub> (Austria) instigating unified performance factors in both orientations, and these factors are completely different with what we identify from



the disintegrated selecting DEA models. Table 7 releases the frequencies of selected factors after solving integrated selecting DEA models for all countries, and Table 8 consequently sums up the top seven frequently selected performance factors. For example, regardless of orientations, the multiplier integrated model embraces six selective outputs  $y_3, y_{22}, y_{26}, y_{28}, y_{31}, y_{35}$  and one selective input  $x_5$  under VRS assumption whereas that envelopment integrated model consists of two selective outputs  $y_{12}, y_{24}$  and five selective inputs  $x_3, x_5, x_7, x_{11}, x_{12}$ . It is obvious that varying outputs and inputs are resulted from the multiplier and envelopment integrated models. This is because the multiplier and envelopment integrated selecting models are no longer the primal-dual pair of problems. We can hence gain insights by returning to Theorem 2 showing that the multiplier integrated model chooses those performance factors with the object of enhancing performance (optimistic perspective) and the envelopment integrated selecting model picks out those performance factors for strengthening the discrimination power of our assessment (pessimistic perspective). Intuitively, a significant decline in economic activities in the EU such as an increase in unemployment and bankruptcy rates and vulnerable social conditions inspires us to apply the envelopment integrated selecting models because it can take a pessimistic view and avoid extreme dissatisfaction by overestimating efficiencies of counties whereas the flourishing society and prosperous economy lead us to have an optimistic view and use the multiplier integrated selecting model to analyze the countries' efficiencies.

**Table 6.** Factors selected from the integrated selecting DEA models for Austria

Binary variables	Multiplier model (16)		Envelopment model (17)		Binary variables	Multiplier model (16)		Envelopment model (17)	
	CRS	VRS	CRS	VRS		CRS	VRS	CRS	VRS
$d_1^x$	0	0	0	0	$d_{12}^y$	0	0	1	1
$d_2^x$	Fixed input		Fixed input		$d_{13}^y$	0	0	1	1
$d_3^x$	1	0	0	0	$d_{14}^y$	0	1	0	0
$d_4^x$	1	0	0	0	$d_{15}^y$	0	0	1	0
$d_5^x$	0	1	0	0	$d_{16}^y$	0	0	0	0
$d_6^x$	0	0	0	0	$d_{17}^y$	0	1	0	0
$d_7^x$	0	0	0	0	$d_{18}^y$	0	1	0	0
$d_8^x$	0	0	0	0	$d_{19}^y$	0	1	0	0
$d_9^x$	0	0	0	0	$d_{20}^y$	0	0	0	0
$d_{10}^x$	1	0	0	0	$d_{21}^y$	0	0	0	0
$d_{11}^x$	0	0	0	0	$d_{22}^y$	0	0	0	0
$d_{12}^x$	0	0	1	1	$d_{23}^y$	0	0	0	0
$d_{13}^x$	0	0	0	0	$d_{24}^y$	0	0	1	1
$d_{14}^x$	0	0	0	0	$d_{25}^y$	0	0	1	1
$d_1^y$	0	0	0	0	$d_{26}^y$	0	0	0	0
$d_2^y$	0	0	0	0	$d_{27}^y$	0	0	0	0
$d_3^y$	0	0	0	0	$d_{28}^y$	0	0	0	0
$d_4^y$	0	0	0	0	$d_{29}^y$	0	0	0	0
$d_5^y$	0	0	0	0	$d_{30}^y$	0	0	0	0
$d_6^y$	0	0	0	0	$d_{31}^y$	0	0	0	0
$d_7^y$	0	0	0	0	$d_{32}^y$	1	0	0	0
$d_8^y$	0	1	0	0	$d_{33}^y$	0	0	0	0
$d_9^y$	Fixed output		Fixed output		$d_{34}^y$	1	0	1	1
$d_{10}^y$	1	1	0	0	$d_{35}^y$	0	0	0	0
$d_{11}^y$	0	0	0	0	$d_{36}^y$	1	0	0	1
Optimal objective function value						0	0	-5.499	-2.114

**Table 7.** Frequencies of a performance factor and its rank derived from the integrated selecting DEA models

Input	Multiplier model (16)		Envelopment model (17)		Output	Multiplier model (16)		Envelopment model (17)	
	CRS	VRS	CRS	VRS		CRS	VRS	CRS	VRS
$x_1$	9 (5 <sup>th</sup> )	4 (18 <sup>th</sup> )	13(5 <sup>th</sup> )	9 (8 <sup>th</sup> )	$y_{12}$	0 (39 <sup>th</sup> )	0 (43 <sup>rd</sup> )	21(1 <sup>st</sup> )	20(1 <sup>st</sup> )
$x_2$	Fixed input		Fixed input		$y_{13}$	2 (22 <sup>nd</sup> )	0 (43 <sup>rd</sup> )	6 (13 <sup>th</sup> )	8 (10 <sup>th</sup> )
$x_3$	18(3 <sup>rd</sup> )	3 (25 <sup>th</sup> )	8 (11 <sup>th</sup> )	11(6 <sup>th</sup> )	$y_{14}$	0 (39 <sup>th</sup> )	5 (14 <sup>th</sup> )	0 (31 <sup>st</sup> )	1 (26 <sup>th</sup> )
$x_4$	26(2 <sup>nd</sup> )	5 (14 <sup>th</sup> )	4 (17 <sup>th</sup> )	3 (17 <sup>th</sup> )	$y_{15}$	3 (18 <sup>th</sup> )	2 (31 <sup>st</sup> )	2 (21 <sup>st</sup> )	2 (20 <sup>th</sup> )
$x_5$	5 (11 <sup>th</sup> )	7 (5 <sup>th</sup> )	15(3 <sup>rd</sup> )	16(3 <sup>rd</sup> )	$y_{16}$	1 (30 <sup>th</sup> )	4 (18 <sup>th</sup> )	0 (31 <sup>st</sup> )	2 (20 <sup>th</sup> )
$x_6$	3 (18 <sup>th</sup> )	4 (18 <sup>th</sup> )	6 (13 <sup>th</sup> )	6 (13 <sup>th</sup> )	$y_{17}$	0 (39 <sup>th</sup> )	2 (31 <sup>st</sup> )	0 (31 <sup>st</sup> )	1 (26 <sup>th</sup> )
$x_7$	0 (39 <sup>th</sup> )	2 (31 <sup>st</sup> )	18(2 <sup>nd</sup> )	14(4 <sup>th</sup> )	$y_{18}$	2 (22 <sup>nd</sup> )	6 (10 <sup>th</sup> )	0 (31 <sup>st</sup> )	0 (36 <sup>th</sup> )
$x_8$	1 (30 <sup>th</sup> )	4 (18 <sup>th</sup> )	3 (18 <sup>th</sup> )	5 (14 <sup>th</sup> )	$y_{19}$	2 (22 <sup>nd</sup> )	2 (31 <sup>st</sup> )	3 (18 <sup>th</sup> )	1 (26 <sup>th</sup> )
$x_9$	1 (30 <sup>th</sup> )	4 (18 <sup>th</sup> )	8 (11 <sup>th</sup> )	3 (17 <sup>th</sup> )	$y_{20}$	0 (39 <sup>th</sup> )	4 (18 <sup>th</sup> )	3 (18 <sup>th</sup> )	2 (20 <sup>th</sup> )
$x_{10}$	5 (11 <sup>th</sup> )	6 (10 <sup>th</sup> )	9 (8 <sup>th</sup> )	9 (8 <sup>th</sup> )	$y_{21}$	5 (11 <sup>th</sup> )	3 (25 <sup>th</sup> )	1 (25 <sup>th</sup> )	0 (36 <sup>th</sup> )
$x_{11}$	6 (8 <sup>th</sup> )	6 (10 <sup>th</sup> )	9 (8 <sup>th</sup> )	11(6 <sup>th</sup> )	$y_{22}$	2 (22 <sup>nd</sup> )	7 (5 <sup>th</sup> )	1 (25 <sup>th</sup> )	1 (26 <sup>th</sup> )
$x_{12}$	2 (22 <sup>nd</sup> )	5 (14 <sup>th</sup> )	15(3 <sup>rd</sup> )	20(1 <sup>st</sup> )	$y_{23}$	3 (18 <sup>th</sup> )	3 (25 <sup>th</sup> )	1 (25 <sup>th</sup> )	1 (26 <sup>th</sup> )
$x_{13}$	1 (30 <sup>th</sup> )	1 (40 <sup>th</sup> )	10(7 <sup>th</sup> )	3 (17 <sup>th</sup> )	$y_{24}$	1 (30 <sup>th</sup> )	2 (31 <sup>st</sup> )	11(6 <sup>th</sup> )	12(5 <sup>th</sup> )
$x_{14}$	6 (8 <sup>th</sup> )	0 (43 <sup>rd</sup> )	2 (21 <sup>st</sup> )	5 (14 <sup>th</sup> )	$y_{25}$	1 (30 <sup>th</sup> )	3 (25 <sup>th</sup> )	6 (13 <sup>th</sup> )	7 (11 <sup>th</sup> )
$y_1$	9 (5 <sup>th</sup> )	2 (31 <sup>st</sup> )	0 (31 <sup>st</sup> )	1 (26 <sup>th</sup> )	$y_{26}$	0 (39 <sup>th</sup> )	10(4 <sup>th</sup> )	0 (31 <sup>st</sup> )	0 (36 <sup>th</sup> )
$y_2$	0 (39 <sup>th</sup> )	3 (25 <sup>th</sup> )	0 (31 <sup>st</sup> )	1 (26 <sup>th</sup> )	$y_{27}$	5 (11 <sup>th</sup> )	6 (10 <sup>th</sup> )	0 (31 <sup>st</sup> )	1 (26 <sup>th</sup> )
$y_3$	1 (30 <sup>th</sup> )	7 (5 <sup>th</sup> )	0 (31 <sup>st</sup> )	0 (36 <sup>th</sup> )	$y_{28}$	0 (39 <sup>th</sup> )	13(2 <sup>nd</sup> )	0 (31 <sup>st</sup> )	0 (36 <sup>th</sup> )
$y_4$	3 (18 <sup>th</sup> )	2 (31 <sup>st</sup> )	0 (31 <sup>st</sup> )	0 (36 <sup>th</sup> )	$y_{29}$	0 (39 <sup>th</sup> )	7 (5 <sup>th</sup> )	0 (31 <sup>st</sup> )	0 (36 <sup>th</sup> )
$y_5$	1 (30 <sup>th</sup> )	1 (40 <sup>th</sup> )	1 (25 <sup>th</sup> )	2 (20 <sup>th</sup> )	$y_{30}$	1 (30 <sup>th</sup> )	5 (14 <sup>th</sup> )	0 (31 <sup>st</sup> )	0 (36 <sup>th</sup> )
$y_6$	0 (39 <sup>th</sup> )	0 (43 <sup>rd</sup> )	1 (25 <sup>th</sup> )	2 (20 <sup>th</sup> )	$y_{31}$	2 (22 <sup>nd</sup> )	16(1 <sup>st</sup> )	0 (31 <sup>st</sup> )	0 (36 <sup>th</sup> )
$y_7$	4 (15 <sup>th</sup> )	0 (43 <sup>rd</sup> )	0 (31 <sup>st</sup> )	0 (36 <sup>th</sup> )	$y_{32}$	2 (22 <sup>nd</sup> )	0 (43 <sup>rd</sup> )	5 (16 <sup>th</sup> )	5 (14 <sup>th</sup> )
$y_8$	2 (22 <sup>nd</sup> )	2 (31 <sup>st</sup> )	9 (8 <sup>th</sup> )	7 (11 <sup>th</sup> )	$y_{33}$	4 (15 <sup>th</sup> )	7 (5 <sup>th</sup> )	0 (31 <sup>st</sup> )	0 (36 <sup>th</sup> )
$y_9$	Fixed output		Fixed output		$y_{34}$	6 (8 <sup>th</sup> )	3 (25 <sup>th</sup> )	1 (25 <sup>th</sup> )	1 (26 <sup>th</sup> )
$y_{10}$	8 (7 <sup>th</sup> )	2 (31 <sup>rd</sup> )	0 (31 <sup>st</sup> )	0 (36 <sup>th</sup> )	$y_{35}$	4 (15 <sup>th</sup> )	11(3 <sup>rd</sup> )	0 (31 <sup>st</sup> )	0 (36 <sup>th</sup> )
$y_{11}$	11(4 <sup>th</sup> )	1 (40 <sup>th</sup> )	2 (21 <sup>st</sup> )	1 (26 <sup>th</sup> )	$y_{36}$	28(1 <sup>st</sup> )	4 (18 <sup>th</sup> )	2 (21 <sup>st</sup> )	2 (20 <sup>th</sup> )

**Table 8.** Seven selected factors for integrated models

Models		Selected factors and their frequencies						
Model (16)	CRS	$y_{36}$ (28)	$x_4$ (26)	$x_3$ (18)	$y_{11}$ (11)	$x_1$ (9)	$y_1$ (9)	$y_{10}$ (8)
	VRS	$y_{31}$ (16)	$y_{28}$ (13)	$y_{35}$ (11)	$y_{26}$ (10)	$x_5$ (7)	$y_3$ (7)	$y_{22}$ (7)
Model (17)	CRS	$y_{12}$ (21)	$x_7$ (18)	$x_5$ (15)	$x_{12}$ (15)	$x_1$ (13)	$y_{24}$ (11)	$x_{13}$ (10)
	VRS	$x_{12}$ (20)	$y_{12}$ (20)	$x_5$ (16)	$x_7$ (14)	$y_{24}$ (12)	$x_3$ (11)	$x_{11}$ (11)

Figures in parentheses in Table 4 and Table 7 represent the overall rankings of the factors for each selecting DEA model as per their frequencies. It is interesting to show that the results of six different models (groups) under each RTS are dependent. To do so, we apply the Kruskal–Wallis H test for CRS and VRS to examine whether there are statistically significant indifferences between the results of six models. As seen in Table 9, the Kruskal–Wallis H test shows that there is no statistically significant difference between the different six models under CRS and VRS technologies with  $\chi^2(2) = 0.749$ ,  $\chi^2(2) = 2.713$ , respectively.

**Table 9.** Comparison of the models apropos of the frequencies of selected factors

Test	CRS	VRS
Kruskal-Wallis H	0.749	2.713
df	5	5
Asymp. Sig.	0.980	0.744

**Table 10.** Validation of Theorem 9 for Austria

Technology	$\bar{\theta}_o^{in*}$	$\bar{\theta}_o^{out*}$	Model (6)	Model (12)
CRS	0.176	5.675	0.176	5.681
VRS	0.270	2.385	0.270	2.386

Let us at present emphasise on  $DMU_1 = \text{Austria}$  for CRS and VRS technologies to oversee Theorem 8 and Theorem 9. Regarding Theorem 8,  $\sum_{r=1}^{36} u_r^{in*} y_{r1} - w_0^{in*}$  and  $\sum_{i=1}^{14} v_i^{out*} x_{i1} + w_0^{out*}$  and the optimal objective values of models of (5) and (11) are equal to 1. Furthermore, in compliance with Theorem 9, Table 10 exhibits that the optimal solutions, i.e.,  $\bar{\theta}_o^{in*}$  ( $\bar{\theta}_o^{out*}$ ), calculated from model (17) is greater (less) than or equal to the optimal objective value of model (6) (model (12)).

Having narrowed down the numerous potential factors and selected the most influential ones with regard to model and situation characteristics, we, in turn, measure the relative efficiency of 28 EU countries in the year 2013. To delve into the all proposed selecting DEA models, either disintegrated or integrated models and their impacts on the efficiency, their descriptive statistics are given in Table 11 and Table 12. Let us start with Table 11 that is the result of efficiencies where the disintegrated models are employed to opt for the performance factors. According to the mean efficiencies, apart from the NDRS and VRS cases, the efficiencies in the input-oriented multiplier models are greater than those calculated by the envelopment models, while the efficiencies of the output-oriented multiplier models are invariably greater than those specified by the envelopment models. Turning to a comparison of variances, findings in Table 11 appealingly shows the higher variances in the envelopment models contrary to multiplier models. From consideration of orientations, the mean efficiencies of input-oriented multiplier (envelopments) models in VRS technology are smaller (greater) than those obtained from the output-oriented models, while the efficiencies of both input- and output-oriented multiplier (envelopment) models in CRS technology are identical (see Theorem 5). We also draw attention to the number of efficient countries in different models showing that discretion in choosing the types of models and orientation imperatively matters. In brief, the discrimination power of the envelopment models in both orientations is more than the multiplier models.

**Table 11.** Descriptive statistics of efficiencies of the EU countries using conventional DEA models

Title	Input orientation				Output orientation			
	Multiplier		Envelopment		Multiplier		Envelopment	
	CRS	VRS	CRS	VRS	CRS	VRS	CRS	VRS
Min	0.412	0.464	0.003	0.557	0.412	0.962	0.003	0.007
Mean	0.902	0.909	0.273	0.923	0.902	0.997	0.273	0.736
1 <sup>st</sup> Quartile	0.827	0.913	0.024	0.88	0.827	1	0.024	0.235
Median	1	1	0.088	1	1	1	0.088	1
3 <sup>rd</sup> Quartile	1	1	0.348	1	1	1	0.348	1
Sd.	0.154	0.162	0.359	0.13	0.154	0.008	0.359	0.42
Max	1	1	1	1	1	1	1	1
No. eff.	16	20	3	19	16	23	3	20

**Table 12.** Descriptive statistics of efficiencies of the EU countries using the integrated selecting models

Title	Input orientation				Output orientation			
	Model (16)		Model (17)		Model (16)		Model (17)	
	CRS	VRS	CRS	VRS	CRS	VRS	CRS	VRS
Min	0.778	0.514	0.010	0.804	0.778	0.982	0.010	0.015
Mean	0.965	0.864	0.333	0.988	0.965	0.997	0.333	0.691
1 <sup>st</sup> Quartile	0.953	0.693	0.081	1.000	0.953	0.997	0.081	0.109
Median	1.000	1.000	0.128	1.000	1.000	1.000	0.128	1.000
3 <sup>rd</sup> Quartile	1.000	1.000	0.609	1.000	1.000	1.000	0.609	1.000
Sd.	0.062	0.174	0.349	0.038	0.062	0.005	0.349	0.422
Max	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
No. eff.	20	16	3	22	20	17	3	18

Table 12 represents the efficiency measures of input- and output-oriented models apropos of the selecting DEA models (16) and (17) . Let us compare the results in terms of orientations and selecting DEA models. Based on Theorem 1 for factor selection, the mean efficiencies in the input-oriented multiplier models are greater than or equal to those calculated by the envelopment models, while regarding model (16) the mean efficiencies of the input-oriented models under VRS assumption are less than those from the output-oriented models. In addition, the different orientations do not affect the discrimination power between but the factors selected by model (17) clearly cause an increase in discriminatory power in DEA with greater variances. According to our findings, we would express that the greater variances in efficiency measure lead to higher discriminatory power and vice versa.

**Table 13.** The results of factor selection, obtained by proposed models (22) and (23)

	Multiplier model (22)		Envelopment model (23)			Multiplier model (22)		Envelopment model (23)	
	CRS	VRS	CRS	VRS		CRS	VRS	CRS	VRS
$d_1^x$	0	0	0	1	$d_{12}^y$	0	0	1	1
$d_2^x$	Fixed input		Fixed input		$d_{13}^y$	0	0	0	0
$d_3^x$	1	1	0	0	$d_{14}^y$	0	0	0	0
$d_4^x$	1	1	0	0	$d_{15}^y$	0	0	0	0
$d_5^x$	0	0	1	1	$d_{16}^y$	0	0	0	0
$d_6^x$	0	0	0	0	$d_{17}^y$	1	1	0	0
$d_7^x$	1	1	1	0	$d_{18}^y$	0	0	0	0
$d_8^x$	0	0	0	0	$d_{19}^y$	0	0	0	0
$d_9^x$	0	0	0	1	$d_{20}^y$	0	0	0	0
$d_{10}^x$	0	0	1	0	$d_{21}^y$	0	0	0	0
$d_{11}^x$	0	0	1	1	$d_{22}^y$	0	0	0	0
$d_{12}^x$	0	0	1	1	$d_{23}^y$	0	0	0	0
$d_{13}^x$	0	0	0	0	$d_{24}^y$	0	0	1	1
$d_{14}^x$	1	0	0	0	$d_{25}^y$	0	0	0	0
$d_1^y$	0	0	0	0	$d_{26}^y$	0	0	0	0
$d_2^y$	0	1	0	0	$d_{27}^y$	0	0	0	0
$d_3^y$	0	0	0	0	$d_{28}^y$	0	0	0	0
$d_4^y$	0	0	0	0	$d_{29}^y$	1	0	0	0
$d_5^y$	0	0	0	0	$d_{30}^y$	0	0	0	0
$d_6^y$	0	0	0	0	$d_{31}^y$	0	0	0	0
$d_7^y$	1	1	0	0	$d_{32}^y$	0	1	0	0
$d_8^y$	0	0	0	0	$d_{33}^y$	0	0	0	0
$d_9^y$	Fixed output		Fixed output		$d_{33}^y$	0	0	0	0
$d_{10}^y$	0	0	0	0	$d_{35}^y$	0	0	0	0
$d_{11}^y$	0	0	0	0	$d_{36}^y$	0	0	0	0

## 6.2 Performance factor selection for structural efficiency

Rather than evaluate efficiency at the country level, let us now focus on [structural] efficiency at the EU level to form a single production unit that allows for resource allocation across constituent countries. The structural efficiency measure identifies how well the EU performs in 2013. To this end, the most influential factors can be selected by solving the multiplier model (22) or the envelopment model (23) for CRS and VRS technologies (see Table 13). Table 14 reports the selected seven factors. Having crunched the results, inputs  $x_3, x_4, x_7$  and outputs  $y_7, y_{17}$  for model (22) and inputs  $x_5, x_{11}, x_{12}$  and outputs  $y_{12}, y_{24}$  for model (23) are always selected despite the RTS characterization. From Table 14, we do not find any common performance factors selected by either model (22) or model (23). Consequently, the structural efficiencies are reported in

Table 15. In the light of models (22) and (23), the structural efficiency values from the both orientations is at least 0.96 and 0.6, respectively, meaning that all inputs should be reduced by about 4% and 40% for the input-oriented assessment and all outputs should be augmented by about 4% and 40% for the output-oriented assessment to omit all types of inefficiency. Our findings show that if the analysis is based upon the factors identified by the multiplier selecting DEA model (22), the EU had a flourishing year (optimistic viewpoint), while the factors recognised by the envelopment selecting DEA model (23) deteriorates the EU efficiency significantly (pessimistic viewpoint).

**Table 14.** Seven selected factors for integrated models

Models		Selected factors						
Model (22)	CRS	$x_3$	$x_4$	$x_7$	$x_{14}$	$y_7$	$y_{17}$	$y_{29}$
	VRS	$x_3$	$x_4$	$x_7$	$y_2$	$y_7$	$y_{17}$	$y_{32}$
Model (23)	CRS	$x_5$	$x_7$	$x_{10}$	$x_{11}$	$x_{12}$	$y_{12}$	$y_{24}$
	VRS	$x_1$	$x_5$	$x_9$	$x_{11}$	$x_{12}$	$y_{12}$	$y_{24}$

**Table 15.** Input- and output-oriented efficiency of the EU under different technologies

Based upon factors selected from	Input-oriented		Output-oriented	
	CRS	VRS	CRS	VRS
Model (22)	0.960	0.961	0.960	0.995
Model (23)	0.609	0.676	0.609	0.847

## 7 Summary and conclusion

One of the initial steps in employing DEA is to identify the most influential performance factors (inputs and outputs), however, it is generally lost in numerous existing studies due to the fact that the literature lacks a mechanism for such important element before undertaking performance evaluation. As the number of factors increases, the problem tends to add more complexity to factor selection in DEA models. In this paper, we thoroughly look into the factor selection problem in DEA models to fill the gap in the literature. The contribution of this research is fivefold. First, the envelopment and multiplier forms of selecting DEA models under VRS assumption are developed for input and output orientations. Second, contrasting properties and relations between the developed

models are identified and in turn discussed. Our key findings reveal that the set of factors selected by the selecting DEA models often vary considerably from input- to output-oriented assessment for VRS assumption, other than CRS assumption. Third, a new mathematics programming model in relation to the DEA principles is developed to unify the results of selecting DEA models for both orientations. Fourth, two distinct levels of assessments, including individual (firm) and structural (industry) levels, are studied in this paper. Fifth, a large number of growth and development factors is utilized to present a case study on assessment of EU countries.

For future study, the developed framework can be extended to other existing DEA models such as non-radial models to identify the performance factors. In this regard, there are many challenges involved in the proposed research that open abundant scope for future research. For example, it would be also interesting to consider data irregularities and structural complexities in the factor selection problem. In addition, the literature entails a plethora of selecting techniques from varying perspectives. The comparison of existent selecting DEA approaches will be carried out in a future paper.

### Acknowledgements

This study was supported by the Czech Science Foundation (GAČR 19-13946S).

### References

- Amirteimoori, A., Despotis, D. K., & Kordrostami, S. (2014). Variables reduction in data envelopment analysis. *Optimization*, 63(5), 735–745.
- Bazaraa, M. S., Jarvis, J. J., & Sherali, H. D. (2010). *Linear programming and network flows* (4th ed.). Hoboken, New Jersey: John Wiley & Sons.
- Charles, V., Aparicio, J., & Zhu, J. (2019). The curse of dimensionality of decision-making units: A simple approach to increase the discriminatory power of data envelopment analysis. *European Journal of Operational Research*, 279(3), 929–940.
- Charnes, A., Cooper, W. W., & Rhodes, E. (1978). Measuring the efficiency of decision making units. *European Journal of Operational Research*, 2, 429–444.
- Cooper, W. W., Seiford, L. M., & Tone, K. (2007). *Data envelopment analysis: A comprehensive text with models, applications, references and DEA-solver software* (2nd ed.). Springer US.
- Dyson, R. G., Allen, R., Camanho, A. S., Podinovski, V. V., Sarrico, C. S., & Shale, E. A. (2001). Pitfalls and protocols in DEA. *European Journal of Operational Research*, 132(2), 245–259.
- Emrouznejad, A., & Yang, G. (2017). A survey and analysis of the first 40 years of scholarly literature in DEA: 1978–2016. *Socio-Economic Planning Sciences*, 61, 4–8.
- Eskelinen, J. (2017). Comparison of variable selection techniques for data envelopment analysis in a retail bank. *European Journal of Operational Research*, 259(2), 778–788.
- Färe, R., & Lovell, C. A. K. (1978). Measuring the technical efficiency of production. *Journal of Economic Theory*, 19(1), 150–162.
- Farrell, M. J. (1957). The Measurement of Productive Efficiency. *Journal of the Royal Statistical Society. Series A (General)*, 120(3), 253–290.
- Førsund, F. R., & Hjalmarsson, L. (1979). Generalised Farrell measures of efficiency: an application to milk processing in Swedish dairy plants. *The Economic Journal*, 89(354), 294–315.
- Friedman, L., & Sinuany-stern, Z. (1998). Combining ranking scales and selecting variables in the DEA context: The case of industrial branches. *Computers & Operations Research*, 25(9), 781–791.
- Golany, B., & Yaakov, R. (1989). An application procedure of DEA. *Omega*, 17(3), 237–250.
- Jenkins, L., & Anderson, M. (2003). A multivariate statistical approach to reducing the number of variables in data envelopment analysis. *European Journal of Operational Research*, 147(1), 51–

61.

- Karagiannis, G. (2015). On structural and average technical efficiency. *Journal of Productivity Analysis*, 43(3), 259–267.
- Khezrimotlagh, D., Cook, W. D., & Zhu, J. (2019). Number of performance measures versus number of decision making units in DEA. *Annals of Operations Research*.
- Lewin, A. Y., Morey, R. C., & Cook, T. J. (1982). Evaluating the administrative efficiency of courts. *Omega*, 10(4), 401–411.
- Li, S. ko, & Cheng, Y. shing. (2007). Solving the puzzles of structural efficiency. *European Journal of Operational Research*, 180(2), 713–722.
- Liu, J. S., Lu, L. Y. Y., Lu, W.-M., & Lin, B. J. Y. (2013a). A survey of DEA applications. *Omega*, 41(5), 893–902.
- Liu, J. S., Lu, L. Y. Y., Lu, W.-M., & Lin, B. J. Y. (2013b). Data envelopment analysis 1978–2010: A citation-based literature survey. *Omega*, 41(1), 3–15.
- Morita, H., & Avkiran, N. K. (2009). Selecting inputs and outputs in data envelopment analysis by designing statistical experiments. *Journal of the Operations Research Society of Japan*, 52(2), 163–173.
- Nataraja, N. R., & Johnson, A. L. (2011). Guidelines for using variable selection techniques in data envelopment analysis. *European Journal of Operational Research*, 215(3), 662–669.
- Pastor, J. T., Ruiz, J. L., & Sirvent, I. (2002). A Statistical Test for Nested Radial Dea Models. *Operations Research Publication*, 50(4), 728–735.
- Ray, S. C., & Hu, X. (1997). On the Technically Efficient Organization of an Industry: A Study of U.S. Airlines. *Journal of Productivity Analysis*, 8(1), 5–18.
- Scheel, H. (2001). Undesirable outputs in efficiency valuations. *European Journal of Operational Research*, 132(2), 400–410.
- Simar, L., & Wilson, P. W. (2001). Testing restrictions in nonparametric efficiency models. *Communications in Statistics Part B: Simulation and Computation*, 30(1), 159–184.
- Toloo, M., & Allahyar, M. (2018). A simplification generalized returns to scale approach for selecting performance measures in data envelopment analysis. *Measurement: Journal of the International Measurement Confederation*, 121, 327–334.
- Toloo, M., & Babae, S. (2015). On variable reductions in data envelopment analysis with an illustrative application to a gas company. *Applied Mathematics and Computation*, 270, 527–533.
- Toloo, M., Barat, M., & Masoumzadeh, A. (2015). Selective measures in data envelopment analysis. *Annals of Operations Research*, 226(1), 623–642.
- Toloo, M., & Hančlová, J. (2019). Multi-valued measures in DEA in the presence of undesirable outputs. *Omega*, (in press).
- Toloo, M., & Tichý, T. (2015). Two alternative approaches for selecting performance measures in data envelopment analysis. *Measurement*, 65, 29–40.
- Wagner, J. M., & Shimshak, D. G. (2007). Stepwise selection of variables in data envelopment analysis: Procedures and managerial perspectives. *European Journal of Operational Research*, 180(1), 57–67.