Nonhomogeneous DMUs in DEA: A Directional Distance Function Approach

Abstract

Data envelopment analysis (DEA) provides performance evaluation for a set of homogeneous decision making units (DMUs) in the sense that all DMUs evaluated with the same criteria setting. In some settings, however, the assumption of having a common input and output bundle may not hold. Such can occur in universities, for example, since they may have different departments, or in hospitals where have different wards. This motivates to the issue of how to fairly evaluate efficiency when inputs and outputs configurations are different. This paper proposes a three-process methodology that aims at evaluating of a set of DMUs when the requirement of homogeneity among inputs and outputs is relaxed. In the first step, based on the duality theory a multiplier directional distance function (DDF) model is developed to determine an appropriate split of inputs and output. In step 2, the efficiency of a DMU is evaluated in terms of each scaled down inputs and outputs. Finally, the overall efficiency score of a DMU is viewed as a weighted combination of a set of product lines efficiencies. To demonstrate the validity and practicability of the proposed method, we apply it to evaluate the performance of a hypothetical data set. The results show that the methodology has the ability to discriminate performance for data with nonhomogeneous inputs and outputs.

Keywords: data envelopment analysis, non-homogeneous inputs and outputs, efficiency evaluation, combined-oriented DEA models, linear programming

1. Introduction

Data envelopment analysis (DEA), originally introduced by Charnes et al. [7], is a nonparametric linear programming (LP) methodology which measures the relative efficiencies of a set of decision making units (DMUs). Last four decades has witnessed the great theoretical developments and practical applications in DEA literature. The enthusiastic reader is referred to some useful surveys includes Cook and Seiford [8], Emrouznejad et al. [11] and Seiford [15]. In the conventional DEA models it is assumed that in a multiple-input multiple-output setting, all outputs are affected by all inputs. Moreover, these models are based on the assumption that all DMUs use the same set of inputs and produce the same set of outputs, making the set of DMUs homogeneous.

In some situations, regarding using the same technology, the assumption of homogeneity among DMUs may be violated. As an example, consider the case of a set of food manufacturing companies where certain foods do not need nutrition labeling or packaging. In such setting, if one of the inputs is labeling or packaging resources, the mentioned certain foods will not influenced by the labeling or packaging resources. Imanirad et al. [12] referred to this as partial input to output impacts and extends the conventional DEA methodology to address the problem of measuring the technical efficiency in such situations. Cook et al. [9, 10] proposed DEA-based models to demonstrate the problem of non-homogeneity of DMUs on the output side. They considered a set of steel fabrication plants for evaluating the relative efficiencies of a set of DMUs where the input set is common across all DMUs but some plants choose not to manufacture certain products. Li et al. [13] investigated the problem of lack of homogeneity on the input side and extended the earlier researches of Cook et al. [9, 10] to cover the case where different input configurations across a set of DMUs. They developed a DEA-based methodology to deal with this situation and applied it to a set of 31 provinces in China in which one of the inputs is the quantity of natural resources available to the region and not all regions have the same natural resources. Barat et al. [1] developed a three-step procedure to assess cost efficiency of nonhomogeneous DMUs with different output configuration. A network DEA methodology is proposed by Barat et al. [2] to address the problem of nonhomogeneity in settings where subunits operate in the mixed network structure. To deal with the problem of evaluating the relative efficiencies of a set of DMUs whose internal structures are nonhomogeneous Barat et al. [3] suggested a DEA methodology and applied it to a set of 40 branches of the largest private bank in a country in the Middle East.

There are other situations in which lack of homogeneity on both input and output sides prevails. As an example, comparing a set of universities where not all institutions have the same departments and hence violated the assumption of homogeneity among both inputs and outputs captures the idea. In another setting, consider a set of hospitals acting as the DMUs. Those without ICU ward cannot be directly compared to those that do have such ward. A related problem that has been widely investigated in the literature and might conceivably be used to treat this problem is the missing data problem (see e.g., Thompson et al. [16]). In the current setting, however, the issue is not that the data for some inputs and outputs is missing for some DMUs, but rather that the DMU does not have those inputs or those certain outputs are not produced. In the case of hospitals considering as DMUs, those without ophthalmology ward cannot fairly be directly compared to those that have such ward. On the other hand, in the case where a DMU for any reason cannot produce a certain products (even decides not to produce that output) or does not have a certain input, it would be leaded distorted results if artificially substituting a zero value or some average value for the missing measure.

To handle the problem of DMUs with nonhomogeneous inputs and outputs, one might potentially propose dividing the set of DMUs into multiple groups in which all of the members of a group using the same inputs and producing the same outputs, and then applying a separate DEA evaluation on each group. By using this approach, a DMU is evaluated in comparison to those DMUs whose inputs and outputs profiles are identical to its own, specifically only true peers. Cook et al. [10] claimed that at least two problems may arise with this approach. The first problem is that in some situations to reflect true peers the set of DMUs may be required to be split into multiple small subsets. This would cause difficulty to assess meaningful evaluation. The other problem is that excluding considerations of partial peers, whose inputs and outputs profiles overlap with but not identical to, those of under evaluation DMU may cause failure in identifying true best practices. This gives rise to the issue of how to include all DMUs in the comparison set to fairly compare a DMU to the others. In this paper we extend the previous researches of Cook et al. [10], and Li et al [13] to encompass the general case which is non-homogeneity on both input and output sides. Generally, this is brought about by viewing the process of inputs generating outputs and the process of outputs producing by inputs as being divided into some separate processes. We develop a DEA type methodology based on directional distance function approach to evaluate these processes.

The rest of the paper is organized as follows. Section 2 is devoted to the development of DEA-based model for dealing with the general case of non-homogeneity on both input and output sides. Section 3 applies the new methodology to a data set of 25 hypothetical DMUs. Conclusions and recommendations appear in section 4.

2. DEA Model for DMUs with Different Input and Output Configurations

2.1 Background

Consider a set J consists of n DMUs, with input levels, $x_{ij}, i=1,\ldots,m$, output levels $y_{rj}, r=1,\ldots,s$. In particular, an under evaluation unit is denoted by $o \in J$. Suppose that the constant returns to scale (CRS) technology is deemed and $\lambda_j \geq 0, j=1,\ldots,n$ are the intensity variables. The production possibility set (PPS) is then defined as:

$$
\begin{aligned}
PPS = \{(x_1,\ldots,x_m,y_1,\ldots,y_s) : & \sum_{j=1}^{n} \lambda_j x_{ij} \leq x_i, i=1,\ldots,m \\
& \sum_{j=1}^{n} \lambda_j y_{rj} \geq y_r, r=1,\ldots,s \\
& \lambda_j \geq 0, j=1,\ldots,n\}
\end{aligned}
\tag{1}
$$

By introducing a directional function $\mathbf{d}=(\mathbf{d}_x,\mathbf{d}_y)=(d_{1x},\ldots,d_{mx},d_{1y},\ldots,d_{sy})\neq 0$ Chambers et al. [4, 5] defined the directional distance function on PPS and proposed the generic directional distance model as follows:

$$
\begin{aligned}
\max \quad & \beta_o \\
\text{s.t.} \quad & \sum_{j=1}^{n} \lambda_j x_{ij} \leq x_{io} - \beta_o d_{io}, \qquad i=1,\ldots,m, \\
& \sum_{j=1}^{n} \lambda_j y_{rj} \geq y_{ro} + \beta_o d_{ro}, \qquad r=1,\ldots,s, \\
& \lambda_j \geq 0, \qquad j=1,\ldots,n, \\
& d_{ix}, d_{ry} \geq 0, \qquad i=1,\ldots,m, r=1,\ldots,s.
\end{aligned}
\tag{2}
$$

Model (2) is known as combined-oriented CCR model and $(\mathbf{d}_x, \mathbf{d}_y)$ shows the moving direction in which leads DMU_o to lie down on efficient frontier. Input-oriented and output-oriented models can easily be derived from model (2) by considering $\mathbf{d}_y = \mathbf{0}$ and $\mathbf{d}_x = \mathbf{0}$, respectively.

In the combined-oriented CCR model (2), $0 \le \beta_o < 1$, and $1-\beta_o$ is the efficiency score of DMU_o. If $0 < \beta_o < 1$, DMU_o is inefficient, moreover, $\beta_o x_{io}$ shows the amount in which DMU_o should apply to decrease input i, and $\beta_o y_{ro}$ indicates the amount of extension that has to be applied to output r, to make DMU_o efficient.

It is worth mentioning that when input and output measures are positive, the observed inputs and outputs are the usual choice for the directional vectors $(\mathbf{d}_x, \mathbf{d}_y)$ (Portela et al. [14]). The specific value of number β_o is the inefficiency value obtained as the optimal solution of the next linear problem:

$$\begin{aligned} \max \quad & \beta_o \\ \text{s.t.} \quad & \sum_{j=1}^{n} \lambda_j x_{ij} \le x_{io} - \beta_o x_{io}, \qquad i = 1, \ldots, m, \\ & \sum_{j=1}^{n} \lambda_j y_{rj} \ge y_{ro} + \beta_o y_{ro}, \qquad r = 1, \ldots, s, \\ & \lambda_j \ge 0, \qquad j = 1, \ldots, n. \end{aligned} \tag{3}$$

Model (3) is the envelopment form of the combined-oriented CCR model. Based on the duality theory, the fractional multiplier form of the combined-oriented model is as follows:

$$\begin{aligned} \min \quad & z_o = \frac{-uy_o + vx_o}{uy_o + vx_o} \\ \text{s.t.} \quad & \frac{uy_j}{vx_j} \le 1, \qquad j = 1, \ldots, n, \\ & u \ge 1\varepsilon, v \ge 1\varepsilon. \end{aligned} \tag{4}$$

where, the input and output weights are denoted by v and u, respectively. $\varepsilon > 0$, is a small non-Archimedean number used to avoid ignoring any factor in calculating efficiency (Charnes and Cooper [6]). In optimality, z_o^* gives the inefficiency score of DMU_o, and $1 - z_o^*$ is the efficiency score of DMU_o.

In models (3) and (4) it is assumed that all the DMUs are homogeneous. Now, consider the problem where not all the resources are held by all DMUs and more over not all the products are produced by all DMUs. This leads to the case where DMUs have different inputs and outputs configurations. We wish to derive the efficiency scores of DMUs in such settings; since not all inputs and outputs are common to all DMUs, using the conventional model (3) would not seem to be appropriate. In the following subsection we proposed a methodology to address this problem.

2.2 A Directional Distance Function Approach

Now, we wish to examine a general setting, as in a case of hospitals with various wards, and evaluate the efficiencies of a set of DMUs in a situation where not all inputs and outputs are common to all DMUs. Assume that DMUs with similar inputs and outputs have been fall into P mutually exclusive groups which is denoted by $N_p, p = 1, \ldots, P$. Suppose that I_{N_p} is the subset of inputs that is held by DMUs in N_p and R_{N_p} denotes the subset of outputs that is produced by all DMUs in N_p. To illustrate, consider a simple example where five DMUs have the following profiles:

	Inputs				Outputs		
DMU no.	I1	I2	I3	I4	Y1	Y2	Y3
1	*	*	-	*	*	*	-
2	-	*	*	*	*	-	*
3	*	*	-	*	*	*	
4	-	*	*	*	*	-	*
5	*	*	*	*	*	-	*

These five DMUs are organized into three subgroups N_1, N_2 and N_3, with those in N_1 consume three inputs i_1, i_2, i_4 to produce two outputs y_1, y_2, whereas those in N_2 use i_2, i_3, i_4 to produce y_1, y_3, and DMUs in N_3 have all four resources and two outputs y_1, y_2.

To handle the problem of evaluating the efficiency of a given DMU in such settings, we propose proceeding in a three-step procedure. In step 1, we determine an appropriate split of the inputs and outputs and denote the proportions by α_{ijr} and β_{rji}, respectively. In fact, for DMU_j, α_{ijr} is the appropriate proportion of input i which is consumed by output r. Similarly, β_{rji} is the appropriate proportion of output r which is produced by input i. In step 2 the efficiency of a DMU in terms of each of its scaled down input and output is evaluated. Step 3 takes a weighted average of the efficiency scores as derived in step 2, to get the overall efficiency score of the DMU. In the following, we discuss the three steps in detail.

Step 1. ***Deriving the split of Inputs and Outputs***

In the first step, for any given DMU, determining proper allocations of each of its input to each of its output is of interest. It is worth mentioning that, according to the rule of product there are

$m \times s$ ways to match each input to each of the outputs. We refer to each of the way as *splitting product line*. Since DMUs are nonhomogeneous through different inputs and outputs, all the splitting product line do not belong to all the DMUs. Moreover, DMUs in the same DMU group (N_p) have the same product lines. Now, for any DMU in each of the DMU groups we have split the production function into the number of its splitting product lines. We argue herein that for under evaluation unit, DMU_o, if dividing up the inputs and outputs is done in a way that results in the best overall or aggregate efficiency score across all of its splitting product lines, it may be the best reasonable and acceptable technique to allocate the most appropriate values to alpha and beta variables. Additionally, we propose to reasonably present the overall efficiency score of DMU_o as the weighted average (convex combination) of the individual splitting product lines efficiencies (across all splitting product lines in which are related to $I_{N_{p^o}}$ and $R_{N_{p^o}}$). We should emphasize that this discussion is on the base of the assumption that a DMU is the sum of its parts and there are no economies or dis-economies of scope; in cases where such economies or dis-economies exist the idea of considering the aggregate efficiency may not be applicable. This situation has a connection to concept of non-homogeneity of outputs and non-homogeneity of inputs as discussed in Cook et al. [10] and Li et al. [13], respectively.

As is mentioned earlier, to derive the aggregate or overall efficiency of DMU_o we consider representing it as a convex combination of the splitting product line efficiencies. Since both inputs and outputs are nonhomogeneous we argue to develop the determination of the $\alpha-$split and $\beta-$split using the fractional multiplier form of the combined-oriented model and via the objective of maximizing the overall score. Additionally, the optimal objective value of model (4) gives the inefficiency value associated to a specific unit of the sample. Hence, in general terms, the following combined-oriented model for DMU_o captures the idea:

$$
\begin{aligned}
e_o = \min \quad & \sum_i \sum_r w_{iN_{p^o}r} \times \frac{-u_r \beta_{rN_{p^o}i} y_{ro} + v_i \alpha_{iN_{p^o}r} x_{io}}{u_r \beta_{rN_{p^o}i} y_{ro} + v_i \alpha_{iN_{p^o}r} x_{io}} \\
\text{s.t.} \quad & \sum_i \sum_r \frac{u_r \beta_{rN_p i} y_{rj}}{v_i \alpha_{iN_p r} x_{ij}} \le 1, \qquad j \in N_p,\, p = 1,\dots,P,\, i \in I_{N_p},\, r \in R_{N_p}, \\
& \sum_i \beta_{rN_p i} = 1, \qquad p = 1,\dots,P,\, r \in R_{N_p}, \\
& \sum_r \alpha_{iN_p r} = 1, \qquad p = 1,\dots,P,\, i \in I_{N_p}, \\
& \sum_i \sum_r w_{iN_{p^o}r} = 1, \quad i \in I_{N_{p^o}},\, r \in R_{N_{p^o}}, \\
& u_r, v_i, \alpha_{iN_p r}, \beta_{rN_p i} \ge \varepsilon, \qquad p = 1,\dots,P,\, i \in I_{N_p},\, r \in R_{N_p}
\end{aligned}
\tag{5}
$$

where, $\beta_{rN_p i}$ denotes the proportion of output r for a DMU, $j \in N_p$ which is produced by input *i*, and $\alpha_{iN_p r}$ is the proportion of input *i* which is used by output *r*.

In deriving the inefficiency score for any $\text{DMU}_j, j \in N_p$ by the proposed model (5), we suggest that the multipliers be chosen such that the weighted ratio of outputs to inputs for each of the splitting product lines to be at or under unity. By doing this, while our model captures the overall inefficiency of the DMU, it derives the inefficiency of the splitting product lines for those DMUs, simultaneously. In model (5), we impose constraints on the separated splitting product lines. At the same time we connect the product lines through the splitting variables α and β, specifically by imposing the convex constraints.

As mentioned earlier, weights $w_{iN_{p^o}r}$ should be designated in a way that reflects the relative importance or contribution of the respective production line (splitting production line that uses input i to produce output r). One reasonable choice of weights would be to choose them according to the contribution of respective production line make to overall production. As we adopt combined-oriented model herein, an appropriate and logical choice for the weights from an accounting points of view, would be the proportion of sum up the total outputs generated and the aggregate inputs consumed by the products line. Hence, we define the weights $w_{iN_{p^o}r}$ to be assigned to any products line as

$$w_{iN_{p^o}r} = \frac{u_r \beta_{rN_{p^o}i} y_{ro} + v_i \alpha_{iN_{p^o}r} x_{io}}{\sum_i \sum_r (u_r \beta_{rN_{p^o}i} y_{ro} + v_i \alpha_{iN_{p^o}r} x_{io})}, \quad i \in I_{N_{p^o}}, r \in R_{N_{p^o}}. \tag{6}$$

It is worth noting that model (5) provides different set of alpha and beta splitting variables for each DMU_j in comparison with those of the other sets.

Model (5) is nonlinear in the present form. Along the lines of the Charnes and Cooper [6], transformation model (5) is equivalently converted into a linear problem. First, we should note that by the definition of the weights $w_{iN_{p^o}r}$ as given by (6), the objective function of model (5) mathematically becomes:

$$e_o = \min \sum_i \sum_r \frac{-u_r \beta_{rNi} y_{ro} + v_i \alpha_{iNr} x_{io}}{\sum_i \sum_r (u_r \beta_{rNi} y_{ro} + v_i \alpha_{iNr} x_{io})}. \tag{7}$$

Then, make the change of variables $z_{iNr} = v_i \alpha_{iNr}$ and $\phi_{rNi} = u_r \beta_{rNi}$, and note that

$$\sum_r \alpha_{iNr} = 1 \;\Rightarrow\; v_i \sum_r \alpha_{iNr} = v_i \;\Rightarrow\; \sum_r z_{iNr} = v_i\,,$$
$$\sum_i \beta_{rNi} = 1 \;\Rightarrow\; u_r \sum_i \beta_{rNi} = u_r \;\Rightarrow\; \sum_i \beta_{rNi} = u_r\,.$$

Now, under the usual transformation, $t = 1 \big/ [\sum_i \sum_r (u_r \beta_{rNi} y_{ro} + v_i \alpha_{iNr} x_{io})]$, and defining $\nu_i = t v_i$, $\mu_r = t u_r$, $\gamma_{iNr} = t z_{iNr}$, and $\delta_{rNi} = t\phi_{rNi}$ problem (5) becomes:

$$\begin{aligned}
e_o = \min \quad & -\sum_{r \in R_{N_{p^o}}} \mu_r y_{ro} + \sum_{i \in I_{N_{p^o}}} \nu_i x_{io} \\
\text{s.t.} \quad & \sum_{r \in R_{N_p}} \mu_r y_{ro} + \sum_{i \in I_{N_p}} \nu_i x_{io} = 1, \\
& \delta_{rNi} y_{rj} - \gamma_{iNr} x_{ij} \le 0, \qquad j \in N_p, p = 1,\ldots,P, i \in I_{N_p}, r \in R_{N_p}, \\
& \sum_i \delta_{rN_p i} = \mu_r, \qquad p = 1,\ldots,P, r \in R_{N_p}, \\
& \sum_r \gamma_{iN_p r} = \nu_i, \qquad p = 1,\ldots,P, i \in I_{N_p}, \\
& \mu_r, \nu_i, \gamma_{iN_p r}, \delta_{rN_p i} \ge \varepsilon, \qquad p = 1,\ldots,P, i \in I_{N_p}, r \in R_{N_p}
\end{aligned} \tag{8}$$

It is worth noting that the solution of model (8), gives a set of optimal variables $\mu_r^*, \nu_i^*, \gamma_{iN_p r}^*, \delta_{rN_p i}^*$ which are specific to under evaluation DMU. The optimal proportion of inputs and outputs are derived from the foregoing variable transformations, which are given by, $\alpha_{iNr}^* = \gamma_{iN_p r}^* \big/ \nu_i^*$ and $\beta_{rNi}^* = \delta_{rN_p i}^* \big/ \mu_r^*$. Using these variables the scaled down inputs and outputs can then be allocated to the respective splitting product line, namely $\tilde{x}_{ijr} = \alpha_{iNr}^* x_{ij}$ and $\tilde{y}_{rji} = \beta_{rN_p i}^* y_{rj}$, $j \in N_p$.

Step 2. *Deriving the splitting product line efficiency scores*

The purpose of the first step is to determine an appropriate set of scaled down inputs and outputs of each product lines for an under evaluation DMU. In step 2, the conventional combined-oriented CCR model can be applied to each of the splitting product lines. Specifically, we form mutually exclusive splitting line subgroup $P_k, k = 1,\ldots,K$, where P_k denotes the subset of splitting product lines with the property that all of its members appear as the measures of exactly the same set of DMUs. Then, let P_k^i and P_k^r denote inputs and outputs of the splitting product line set P_k, respectively. Moreover, define L_{N_p} contains those P_k forming the full splitting product line set for any DMU in N_p, and T_{P_k} is the set of all DMU groups that have P_k as a member, specifically

$$T_{P_k} = \{N_p \text{ such that } P_k \in L_{N_p}\} \tag{9}$$

Now, for each DMU_o, and each $P_k \in L_{N_{p^o}}$, solve the following combined-oriented DEA model:

$$
\begin{aligned}
z_{P_{k^o}} = \text{Minimize} \quad & -\sum_{r \in P^r_{k^o}} u_r \tilde{y}_{roi} + \sum_{i \in P^i_{k^o}} v_i \tilde{x}_{ior} \\
\text{subject to} \quad & \sum_{r \in P^r_{k^o}} u_r \tilde{y}_{roi} + \sum_{i \in P^i_{k^o}} v_i \tilde{x}_{ior} = 1, \\
& \sum_{r \in P^r_{k^o}} u_r \tilde{y}_{rji} - \sum_{i \in P^i_{k^o}} v_i \tilde{x}_{ijr} \le 0, \qquad j \in N_p, N_p \in T_{P_{k^o}}, \\
& u_r, v_i \ge \varepsilon, \quad r \in P^r_{k^o}, i \in P^i_{k^o}.
\end{aligned} \tag{10}
$$

The subgroup efficiency score of DMU_o, is derived by $e_{P_{k^o}} = 1 - z_{P_{k^o}}$.

Step 3. ***Deriving the aggregate efficiency scores***

In the final step the overall efficiency score of DMU_o is obtained by taking a weighted average of the splitting product lines scores derived from step 2 using the weights defined in (6). It should be noted that in computing $w_{iN_{p^o}r}$, a proper set of u_r and v_i should be used. These values are calculated by solving model (8). Moreover, the total value of all resources and the total value of all products which are respectively consumed and produced by DMU_o is given by $\sum_{r \in R_{N_p}} \mu_r y_{ro} + \sum_{i \in I_{N_p}} \nu_i x_{io}$ which is scaled to unity as per the first constraint of model (8). Therefore, the reduced weights $w_{iN_{p^o}r} = \delta_{rN_{p^o}i} y_{ro} + \gamma_{iN_{p^o}r} x_{io}$ can be considered as an appropriate set of weights. Specifically, $w_{P_{k^o}} = \sum_{i \in P^i_{k^o}} \sum_{r \in P^r_{k^o}} w_{iN_{p^o}r}$.

Theorem. A DMU can be efficient if and only if all of its splitting product line sets are efficient as well.

Proof. On the contrary, assume that a DMU is efficient and at least one of its splitting product line is not efficient (specifically, $e_{P_{k^o}} < 1$). According to the proposed methodology, $e_o = \sum_{k=1}^{K} w_{P_{k^o}} e_{P_{k^o}}$, $\sum_{k=1}^{K} w_{P_{k^o}} = 1$, and $e^o_{P_{k^i_o}} < 1$. Since $e_{P_{k^o}} < 1$, then $e_o < 1$, which violates the assumption of being efficient of DMU_o, hence all of the splitting product lines of under-evaluating DMU are efficient. On the other hand, suppose that all of the splitting product lines of DMU_o are efficient. The efficiency of DMU_o is the weighted average of these efficient product lines, hence the overall efficiency score of DMU is one, and the DMU is efficient. This completes the proof. □

3. Numerical Example

This section includes a numerical illustration of the use of the methodology proposed in the foregoing section. We apply the proposed methodology to a set of hypothetical data set involving 25 DMUs with three inputs and three outputs but with different inputs and outputs configurations; both on the input side and output side the commonality of inputs and outputs among DMUs is missing. In other words, some DMUs choose not to manufacture certain products, and input configuration existing in DMUs might be different from the configuration in other DMUs. The DMUs fall into 3 groups as shown in Table 1. Seven of the DMUs have inputs I1, I2, and outputs O1 and O2, consisting group N2. DMU group N3, contains inputs I1, I3, and outputs O1 and O3. The ten member of group N1 have all three resources and all three products. There are nine splitting product lines involved in generating outputs for each DMU in N1 group, namely product line 1 which uses input I1 and produce output O1, product line 2 which uses input I2 to produce output O1 and so on. DMUs in groups N2 and N3, involves four product lines. Table 2 displays data on 25 hypothetical DMUs.

Now, we use the methodology proposed in the foregoing section and calculate the efficiency scores of DMUs. Recall that the purpose of step 1 (model (8)) is to determine the $\alpha -$split and $\beta -$split for each splitting product lines of a DMU such that results in the best overall efficiency score (minimum overall inefficiency score) across all of its product lines. Applying model (8), the alpha and beta variables for each DMU_o in $N_p, p = 1,2,3$ have been derived. The results are displayed in Tables 3 and 4. The alpha and beta variables obtained from (8) describe the portions of each input and each output in a DMU group that are paired up with each other. Tables 5, 6 contain the scaled data for the three DMU groups as described in step 1.

It can be shown that for the DMU profiles in Table 1 the splitting product line sets and the respective inputs and output sets P_k^i and P_k^r are

$$
\begin{aligned}
&P_1 = \{(I_1, O_1)\},\ P_2 = \{(I_1, O_2), (I_2, O_1), (I_2, O_2)\},\\
&P_3 = \{(I_1, O_3), (I_3, O_1), (I_3, O_3)\},\ P_4 = \{(I_2, O_3), (I_3, O_2)\},\\
&P_1^i = \{1\},\ P_2^i = \{1,2\},\ P_3^i = \{1,3\},\ P_4^i = \{2,3\},\\
&P_1^r = \{1\},\ P_2^r = \{1,2\},\ P_3^r = \{1,3\},\ P_4^r = \{2,3\}.
\end{aligned}
$$

Note that in the hypothetical example described above,

$$
\begin{aligned}
&L_{N_1} = \{P_1, P_2, P_3, P_4\},\ L_{N_2} = \{P_1, P_2\},\ L_{N_3} = \{P_1, P_3\},\\
&T_{P_1} = \{N_1, N_2, N_3\},\ T_{P_2} = \{N_1, N_2\},\ T_{P_3} = \{N_1, N_3\},\ T_{P_4} = \{N_1\}.
\end{aligned}
$$

Then, using the appropriately adjusted data, model (10) is applied to each splitting product line subgroups related to under evaluation DMU. The resulting aggregate efficiency scores are demonstrated along with their relevant subgroup scores in Table 8. Data which are displayed in Table 7 are the product line weights arising from the solution of (8) and are used to derive the aggregate efficiency scores. Table 8 demonstrates that DMU 14 has the lowest efficiency score, 0.44, and DMU 19 with score 1 has the highest efficiency among all DMUs. It is worth mentioning that a DMU is efficient if and only if all of its subgroups are efficient as well. Moreover, within each subgroup T_{P_k} at least one of the DMUs is efficient. We note that 9 out of 25 DMUs show a mix of efficient and inefficient subgroup efficiencies.

To complete the analysis of this section, we compare the efficiency results obtained by the proposed methodology with what the conventional DEA analysis had been rendered by simply inserting zero data for any missing inputs and outputs. The results are displayed in the last column of Table 8. It is worth noting that having replaced all blank spaces with zeros, a major number of DMUs are yielded technically efficient. 10 out of 25 DMUs are reported as efficient in evaluating by the conventional combined-oriented DEA model (4) and DMU 15 has the lowest efficiency score, 0.70. The results from the conventional DEA analysis and the proposed method are not the same to each other. However, there still exist some consistencies, for example, if the results from our method identify one DMU is inefficient then the other method provide similar conclusions on that DMU. Another interesting phenomenon is that, except for the three DMUs, all the efficiency scores by the proposed method are less than the efficiency scores by the conventional DEA model. For these three DMUs the efficiency scores of both methods are approximately the same.

4. Conclusions and Further Directions

In this paper the usual assumption of examining the efficiency of a set of DMUs, requirement of homogeneity among DMUs, is relaxed. This motivates us to measure the relative efficiency in the presence of different inputs and different outputs configurations across a set of DMUs. This environment is related to the problem of missing data which has been extensively addressed in the literature, however in the context that the missing value exists but is not available to the DMU or the DMU intended to produce it but for a reason none was actually created. Herein, we argue that the input/output bundle can differ from one DMU to another and the assumption of homogeneity is violated. To address this non-conventional situation in DEA literature we develop a DEA-based methodology which considers a DMU as a set of splitting production lines. The overall efficiency of a DMU is derived by proceeding in three steps which allows the overall efficiency to be viewed as the weighted average of the efficiency scores for the subgroups that make up the DMU. To show the practical aspect of the proposed methodology we applied our proposed model to a hypothetical data set. The results obtained from the proposed approach shown that a DMU will be evaluated efficient if (and only if) it is efficient in all of its splitting product line sets.

The methodology developed in this paper is based on the assumption that no economies or diseconomies of scope exists. More explicitly, in this paper it is assumed that subgroups efficiencies can be aggregated via a weighted average to provide the overall efficiency score of a DMU. Further research is needed to cover the cases where scope consideration is necessary.

Another suggestion for the extension of research is to accommodate special variables such as dual-role factors, undesirable outputs, ordinal data, and bounded data into the models.

Availability of data The proposed methodology are applied to a set of hypothetical data that is inserted in the manuscript.

References

1. Barat, M., Tohidi, G.H., Sanei, M.: Cost Efficiency Measures in Data Envelopment Analysis with Nonhomogeneous DMUs. Int. J. Ind. Math. 10(1), 75-85 (2018)
2. Barat, M., Tohidi, G.H., Sanei, M.: DEA for Nonhomogeneous Mixed Networks. Asia Pac. Manag. Rev. 24(2), 161-166 (2019)
3. Barat, M., Tohidi, G.H., Sanei, M., Razaviyan, S.H.: Data Envelopment Analysis for Decision Making Units with Nonhomogeneous Internal Structures: An Application to the Banking Industry. J Oper Res Soc. 70(5), 760-769 (2019)
4. Chambers, R.G., Chung, Y., Fare, R.: Benefit and Distance Functions. J. Econ. Theory. 70, 407-419 (1996).
5. Chambers, R.G., Chung, Y., Fare, R.: Profit, directional distance functions, and nerlovian efficiency. J Optimiz Theory App. 98, 351-364 (1998)
6. Charnes, A., Cooper, W.W.: Programming with linear fractional functional. Nav. Res. Logist. Q. 15, 333-334 (1962)
7. Charnes, A., Cooper, W.W., Rhodes, E.: Measuring the efficiency of decision making units. Eur. J. Oper. Res. 2(6), 428-444 (1978)
8. Cook, W.D., Seiford, L.: Data envelopment analysis (DEA)-Thirty years on. Eur. J. Oper. Res. 192(1), 1-17 (2009)
9. Cook, W.D., Harrison, J., Rouse, P., Zhu, J.: Relative efficiency measurement: The problem of missing output in a subset of decision making units. Eur. J. Oper. Res. 220(1), 79-84 (2012)
10. Cook, W.D., Harrison, J., Imanirad, R., Rouse, P., Zhu, J.: Data envelopment analysis with non-homogeneous DMUs. Oper. Res. 61, 666-676 (2013)
11. Emrouznejad, A., Parker, B.R., Tavares, G.: Evaluation of research in efficiency and productivity: A survey and analysis of the first 30 years of scholarly literature in DEA. Socio-Econ Plan Sci. 42, 151-157 (2008)
12. Imanirad, R., Cook, W.D., Zhu, J.: Partial input to output impacts in DEA: production considerations and resource sharing among business subunits. Nav Res Log. 60, 190-207 (2013)
13. Li, W.H., Liang, L., Cook, W.D., Zhu, J.: DEA models for non-homogeneous DMUs with different input configurations. Eur. J. Oper. Res. 254(3), 946-956 (2016)
14. Portela, M.C.A.S., Thanassoulis, E., Simpson, G.: A directional distance approach to deal with negative data in DEA: An application to bank branches. J Oper Res Soc, 55(10), 1111-1121 (2004)
15. Seiford, L.: Data envelopment analysis: The evaluation of the state of art (1978-1995). J Prod Anal. 7, 99-138 (1996)
16. Thompson, R.G., Dharmapala, P.S., Thrall, R.M.: Importance for DEA of zeros in the data, multipliers, and solutions. J Prod Anal. 4, 379-390 (1993)

Table 1. DMU profiles						
	Inputs			Outputs		
DMU Group	I1	I2	I3	O1	O2	O3
N1	*	*	*	*	*	*
N2	*	*	-	*	*	-
N3	*	-	*	*	-	*

Table 2. Data

DMUs	I1	I2	I3	O1	O2	O3
DMU1	70	123	44	10	3	20
DMU2	89	162	20	14	9	20
DMU3	62	137	39	20	5	20
DMU4	69	136	40	15	4	24
DMU5	51	147	22	18	8	20
DMU6	55	132	24	18	5	20
DMU7	43	140	50	16	7	20
DMU8	39	170	30	11	8	24
DMU9	61	145	33	18	6	22
DMU10	62	147	50	18	4	25
DMU11	81	146	-	7	9	-
DMU12	84	117	-	5	8	-
DMU13	44	161	-	11	7	-
DMU14	78	135	-	6	3	-
DMU15	91	167	-	5	5	-
DMU16	52	162	-	9	7	-
DMU17	87	149	-	10	7	-
DMU18	42	-	45	9	-	22
DMU19	42	-	28	14	-	21
DMU20	88	-	36	10	-	23
DMU21	82	-	48	13	-	22
DMU22	75	-	30	14	-	22
DMU23	58	-	33	9	-	24
DMU24	92	-	45	16	-	22
DMU25	37	-	45	11	-	25

Table 3. Alpha values

DMUs	α_{1N1}	α_{1N2}	α_{1N3}	α_{2N1}	α_{2N2}	α_{2N3}	α_{3N1}	α_{3N2}	α_{3N3}
DMU1	0.03	0.37	0.60	0.82	0.03	0.15	0.94	0.03	0.03
DMU2	0.04	0.37	0.59	0.81	0.04	0.15	0.93	0.04	0.04
DMU3	0.03	0.37	0.60	0.81	0.03	0.15	0.94	0.03	0.03
DMU4	0.03	0.37	0.60	0.81	0.03	0.15	0.93	0.03	0.03
DMU5	0.03	0.37	0.60	0.81	0.03	0.15	0.94	0.03	0.03
DMU6	0.03	0.37	0.60	0.82	0.03	0.15	0.94	0.03	0.03
DMU7	0.03	0.37	0.60	0.81	0.03	0.15	0.94	0.03	0.03
DMU8	0.03	0.37	0.60	0.81	0.03	0.15	0.94	0.03	0.03
DMU9	0.03	0.37	0.60	0.81	0.03	0.15	0.93	0.03	0.03
DMU10	0.03	0.37	0.59	0.81	0.03	0.15	0.93	0.03	0.03
DMU11	0.97	0.03	-	0.03	0.97	-	-	-	-
DMU12	0.98	0.02	-	0.02	0.98	-	-	-	-

DMU13	0.98	0.02	-	0.02	0.98	-	-	-	-
DMU14	0.98	0.02	-	0.02	0.98	-	-	-	-
DMU15	0.97	0.03	-	0.03	0.97	-	-	-	-
DMU16	0.98	0.02	-	0.02	0.98	-	-	-	-
DMU17	0.97	0.03	-	0.03	0.97	-	-	-	-
DMU18	0.62	-	0.38	-	-	-	0.01	-	0.99
DMU19	0.62	-	0.38	-	-	-	0.01	-	0.99
DMU20	0.61	-	0.39	-	-	-	0.02	-	0.98
DMU21	0.61	-	0.39	-	-	-	0.02	-	0.98
DMU22	0.62	-	0.38	-	-	-	0.02	-	0.98
DMU23	0.62	-	0.38	-	-	-	0.02	-	0.98
DMU24	0.61	-	0.39	-	-	-	0.02	-	0.98
DMU25	0.62	-	0.38	-	-	-	0.01	-	0.99

Table 4. Beta values

DMUs	β_{1N1}	β_{1N2}	β_{1N3}	β_{2N1}	β_{2N2}	β_{2N3}	β_{3N1}	β_{3N2}	β_{3N3}
DMU1	0.02	0.97	0.02	0.97	0.02	0.02	0.52	0.47	0.02
DMU2	0.02	0.96	0.02	0.96	0.02	0.02	0.51	0.47	0.02
DMU3	0.02	0.97	0.02	0.97	0.02	0.02	0.52	0.47	0.02
DMU4	0.02	0.97	0.02	0.97	0.02	0.02	0.52	0.47	0.02
DMU5	0.02	0.97	0.02	0.97	0.02	0.02	0.52	0.47	0.02
DMU6	0.02	0.97	0.02	0.97	0.02	0.02	0.52	0.47	0.02
DMU7	0.02	0.97	0.02	0.97	0.02	0.02	0.52	0.47	0.02
DMU8	0.02	0.97	0.02	0.97	0.02	0.02	0.52	0.47	0.02
DMU9	0.02	0.97	0.02	0.97	0.02	0.02	0.52	0.47	0.02
DMU10	0.02	0.96	0.02	0.96	0.02	0.02	0.52	0.47	0.02
DMU11	0.99	0.01	-	0.01	0.99	-	-	-	-
DMU12	0.99	0.01	-	0.01	0.99	-	-	-	-
DMU13	0.99	0.01	-	0.01	0.99	-	-	-	-
DMU14	0.99	0.01	-	0.01	0.99	-	-	-	-
DMU15	0.99	0.01	-	0.01	0.99	-	-	-	-
DMU16	0.99	0.01	-	0.01	0.99	-	-	-	-
DMU17	0.99	0.01	-	0.01	0.99	-	-	-	-
DMU18	0.98	-	0.02	-	-	-	0.30	-	0.70
DMU19	0.99	-	0.01	-	-	-	0.30	-	0.70
DMU20	0.98	-	0.02	-	-	-	0.30	-	0.70
DMU21	0.98	-	0.02	-	-	-	0.30	-	0.70
DMU22	0.98	-	0.02	-	-	-	0.30	-	0.70
DMU23	0.98	-	0.02	-	-	-	0.30	-	0.70
DMU24	0.98	-	0.02	-	-	-	0.31	-	0.69
DMU25	0.98	-	0.02	-	-	-	0.30	-	0.70

Table 5. Scaled inputs

DMUs	$\tilde{x}_{1j1}$	$\tilde{x}_{1j2}$	$\tilde{x}_{1j3}$	$\tilde{x}_{2j1}$	$\tilde{x}_{2j2}$	$\tilde{x}_{2j3}$	$\tilde{x}_{3j1}$	$\tilde{x}_{3j2}$	$\tilde{x}_{3j3}$
DMU1	2.09	26.09	41.81	100.31	3.68	19.01	41.37	1.32	1.32
DMU2	3.13	32.96	52.92	131.31	5.70	25.01	18.60	0.70	0.70
DMU3	2.00	23.05	36.96	111.41	4.42	21.16	36.48	1.26	1.26
DMU4	2.25	25.64	41.12	110.56	4.43	20.99	37.39	1.30	1.30
DMU5	1.56	19.00	30.45	119.78	4.50	22.74	20.65	0.67	0.67
DMU6	1.60	20.52	32.88	107.74	3.85	20.41	22.60	0.70	0.70
DMU7	1.35	16.00	25.66	113.97	4.39	21.65	46.87	1.57	1.57
DMU8	1.25	14.50	23.25	138.31	5.43	26.25	28.08	0.96	0.96
DMU9	1.98	22.66	36.35	117.88	4.72	22.41	30.85	1.07	1.07

DMU10	2.15	22.97	36.87	119.21	5.10	22.71	46.52	1.74	1.74
DMU11	78.91	2.08	-	3.75	142.23	-	-	-	-
DMU12	82.10	1.89	-	2.64	114.36	-	-	-	-
DMU13	42.94	1.05	-	3.84	157.14	-	-	-	-
DMU14	76.21	1.79	-	3.10	131.91	-	-	-	-
DMU15	88.48	2.52	-	4.62	162.38	-	-	-	-
DMU16	50.72	1.27	-	3.95	158.03	-	-	-	-
DMU17	84.66	2.33	-	3.99	144.99	-	-	-	-
DMU18	25.92	-	16.08	-	-	-	0.65	-	44.34
DMU19	25.95	-	16.05	-	-	-	0.38	-	27.62
DMU20	54.04	-	33.95	-	-	-	0.67	-	35.33
DMU21	50.30	-	31.71	-	-	-	0.94	-	47.06
DMU22	46.13	-	28.87	-	-	-	0.52	-	29.48
DMU23	35.77	-	22.24	-	-	-	0.51	-	32.50
DMU24	56.35	-	35.66	-	-	-	0.94	-	44.06
DMU25	22.83	-	14.18	-	-	-	0.67	-	44.33

Table 6. Scaled outputs

DMUs	$\tilde{y}_{1j1}$	$\tilde{y}_{1j2}$	$\tilde{y}_{1j3}$	$\tilde{y}_{2j1}$	$\tilde{y}_{2j2}$	$\tilde{y}_{2j3}$	$\tilde{y}_{3j1}$	$\tilde{y}_{3j2}$	$\tilde{y}_{3j3}$
DMU1	0.16	9.68	0.16	2.90	0.05	0.05	10.34	9.34	0.32
DMU2	0.26	13.47	0.26	8.66	0.17	0.17	10.30	9.32	0.37
DMU3	0.34	19.31	0.34	4.83	0.09	0.09	10.32	9.33	0.34
DMU4	0.26	14.48	0.26	3.86	0.07	0.07	12.38	11.20	0.42
DMU5	0.29	17.41	0.29	7.74	0.13	0.13	10.34	9.34	0.33
DMU6	0.28	17.44	0.28	4.84	0.08	0.08	10.35	9.34	0.31
DMU7	0.27	15.46	0.27	6.77	0.12	0.12	10.33	9.33	0.33
DMU8	0.19	10.63	0.19	7.73	0.14	0.14	12.39	11.20	0.41
DMU9	0.31	17.38	0.31	5.79	0.10	0.10	11.35	10.26	0.38
DMU10	0.33	17.33	0.33	3.85	0.07	0.07	12.88	11.66	0.46
DMU11	6.90	0.10	-	0.12	8.88	-	-	-	-
DMU12	4.94	0.06	-	0.10	7.90	-	-	-	-
DMU13	10.86	0.14	-	0.09	6.91	-	-	-	-
DMU14	5.93	0.07	-	0.04	2.96	-	-	-	-
DMU15	4.93	0.07	-	0.07	4.93	-	-	-	-
DMU16	8.88	0.12	-	0.09	6.91	-	-	-	-
DMU17	9.86	0.14	-	0.10	6.90	-	-	-	-
DMU18	8.86	-	0.14	-	-	-	6.63	-	15.37
DMU19	13.80	-	0.20	-	-	-	6.32	-	14.69
DMU20	9.80	-	0.20	-	-	-	6.99	-	16.01
DMU21	12.73	-	0.27	-	-	-	6.70	-	15.30
DMU22	13.74	-	0.26	-	-	-	6.67	-	15.33
DMU23	8.85	-	0.15	-	-	-	7.24	-	16.76
DMU24	15.65	-	0.35	-	-	-	6.71	-	15.28
DMU25	10.83	-	0.18	-	-	-	7.54	-	17.46

Table 7. Weights

DMUs	w_{1N1}	w_{1N2}	w_{1N3}	w_{2N1}	w_{2N2}	w_{2N3}	w_{3N1}	w_{3N2}	w_{3N3}
DMU1	0.008	0.107	0.193	0.407	0.014	0.105	0.154	0.005	0.006
DMU2	0.011	0.133	0.201	0.461	0.019	0.109	0.060	0.003	0.003
DMU3	0.008	0.099	0.167	0.462	0.016	0.108	0.130	0.005	0.006
DMU4	0.009	0.102	0.186	0.434	0.016	0.112	0.131	0.005	0.006
DMU5	0.007	0.101	0.153	0.516	0.017	0.121	0.079	0.003	0.004
DMU6	0.007	0.100	0.170	0.493	0.015	0.117	0.090	0.003	0.004

DMU7	0.006	0.082	0.130	0.469	0.016	0.112	0.171	0.006	0.007
DMU8	0.005	0.079	0.126	0.528	0.020	0.133	0.100	0.004	0.005
DMU9	0.008	0.100	0.167	0.475	0.017	0.115	0.109	0.004	0.005
DMU10	0.008	0.088	0.163	0.446	0.017	0.112	0.153	0.006	0.007
DMU11	0.353	0.009	-	0.016	0.622	-	-	-	-
DMU12	0.407	0.009	-	0.013	0.571	-	-	-	-
DMU13	0.241	0.005	-	0.018	0.736	-	-	-	-
DMU14	0.370	0.008	-	0.014	0.607	-	-	-	-
DMU15	0.349	0.010	-	0.018	0.624	-	-	-	-
DMU16	0.259	0.006	-	0.018	0.717	-	-	-	-
DMU17	0.374	0.010	-	0.016	0.600	-	-	-	-
DMU18	0.295	-	0.192	-	-	-	0.007	-	0.506
DMU19	0.379	-	0.213	-	-	-	0.006	-	0.403
DMU20	0.407	-	0.261	-	-	-	0.006	-	0.327
DMU21	0.382	-	0.233	-	-	-	0.007	-	0.378
DMU22	0.425	-	0.252	-	-	-	0.005	-	0.318
DMU23	0.360	-	0.238	-	-	-	0.005	-	0.397
DMU24	0.411	-	0.242	-	-	-	0.007	-	0.339
DMU25	0.285	-	0.184	-	-	-	0.007	-	0.524

Table 8. Efficiency results

DMUs	$e^{o}_{p_1}$	$e^{o}_{p_2}$	$e^{o}_{p_3}$	$e^{o}_{p_4}$	Overall Efficiency	Conventional Combined-oriented Efficiency
DMU1	0.25	0.74	0.55	0.96	0.69	0.73
DMU2	0.27	1	1	1	0.99	1
DMU3	0.49	1	0.94	0.96	0.97	0.96
DMU4	0.36	0.87	0.76	1	0.84	0.85
DMU5	0.52	1	1	1	0.99	1
DMU6	0.49	0.99	0.94	1	0.97	0.98
DMU7	0.54	1	0.9	1	0.96	1
DMU8	0.44	0.92	0.84	0.99	0.9	1
DMU9	0.46	0.96	0.92	0.99	0.94	0.92
DMU10	0.45	0.94	0.88	0.98	0.92	0.91
DMU11	0.28	0.99	-	-	0.73	1
DMU12	0.2	1	-	-	0.67	1
DMU13	0.64	1	-	-	0.91	1
DMU14	0.26	0.55	-	-	0.44	0.78
DMU15	0.19	0.63	-	-	0.47	0.7
DMU16	0.5	0.96	-	-	0.84	0.96
DMU17	0.36	0.84	-	-	0.66	1
DMU18	0.78	-	0.91	-	0.87	0.91
DMU19	1	-	1	-	1	1
DMU20	0.51	-	0.92	-	0.75	0.92
DMU21	0.64	-	0.83	-	0.75	0.76
DMU22	0.72	-	1	-	0.88	0.99
DMU23	0.64	-	0.98	-	0.85	0.98
DMU24	0.69	-	0.94	-	0.83	0.83
DMU25	0.94	-	1	-	0.98	1