Additive decomposition in two-stage DEA: An alternative approach

Dimitrs Despotis and Gregory Koronakos and Dimitris Sotiros

University of Piraeus

13. October 2012

Additive decomposition in two-stage DEA: An alternative approach

Dimitris K. Despotis[1]
Department of Informatics, University of Piraeus,
80, Karaoli and Dimitriou, 18534, Piraeus, Greece
Tel: +30210 4142315, Fax: +30210 4142357
despotis@unipi.gr

Gregory Koronakos
Department of Informatics, University of Piraeus,
80, Karaoli and Dimitriou, 18534, Piraeus, Greece
Tel: +30210 4142378, Fax: +30210 4142357
gkoron@unipi.gr

Dimitris Sotiros
Department of Informatics, University of Piraeus,
80, Karaoli and Dimitriou, 18534, Piraeus, Greece
Tel: +30210 4142378, Fax: +30210 4142357
dsotiros@unipi.gr

Abstract

Typically, a two-stage production process assumes that the first stage transforms external inputs to a number of intermediate measures, which then are used as inputs to the second stage that produces the final outputs. The three fundamental approaches to efficiency assessment in the context of DEA (two-stage DEA) are the simple (or independent), the multiplicative and the additive. The simple approach does not assume any relationship between the two stages and estimates the overall efficiency and the individual efficiencies for the two stages independently with typical DEA models. The other two approaches assume a series relationship between the two stages and differ in the way they conceptualize the decomposition of the overall efficiency to the efficiencies of the individual stages. This paper presents an alternative approach to additive efficiency decomposition in two-stage DEA. We show that when using the intermediate measures as pivot, it is possible to aggregate the efficiency assessment models of the two individual stages in a single linear program. We test our models with data sets taken from previous studies and we compare the results with those reported in the literature.

Keywords: Data envelopment analysis (DEA), Efficiency, Decomposition, Two-stage DEA

[1] Corresponding author: Dimitris K. Despotis (despotis@unipi.gr) , University of Piraeus, Department of Informatics, 80, Karaoli and Dimitriou, 18534, Piraeus, Greece. Tel.: +302104142315, Fax: +302104142357.

Additive decomposition in two-stage DEA: An alternative approach

Dimitris K. Despotis[2]
Department of Informatics, University of Piraeus,
80, Karaoli and Dimitriou, 18534, Piraeus, Greece
Tel: +30210 4142315, Fax: +30210 4142357
despotis@unipi.gr

Gregory Koronakos
Department of Informatics, University of Piraeus,
80, Karaoli and Dimitriou, 18534, Piraeus, Greece
Tel: +30210 4142378, Fax: +30210 4142357
gkoron@unipi.gr

Dimitris Sotiros
Department of Informatics, University of Piraeus,
80, Karaoli and Dimitriou, 18534, Piraeus, Greece
Tel: +30210 4142378, Fax: +30210 4142357
dsotiros@unipi.gr

Abstract

Typically, a two-stage production process assumes that the first stage transforms external inputs to a number of intermediate measures, which then are used as inputs to the second stage that produces the final outputs. The three fundamental approaches to efficiency assessment in the context of DEA (two-stage DEA) are the simple (or independent), the multiplicative and the additive. The simple approach does not assume any relationship between the two stages and estimates the overall efficiency and the individual efficiencies for the two stages independently with typical DEA models. The other two approaches assume a series relationship between the two stages and differ in the way they conceptualize the decomposition of the overall efficiency to the efficiencies of the individual stages. This paper presents an alternative approach to additive efficiency decomposition in two-stage DEA. We show that when using the intermediate measures as pivot, it is possible to aggregate the efficiency assessment models of the two individual stages in a single linear program. We test our models with data sets taken from previous studies and we compare the results with those reported in the literature.

Keywords: Data envelopment analysis (DEA), Efficiency, Decomposition, Two-stage DEA

---

[2] Corresponding author: Dimitris K. Despotis (despotis@unipi.gr) , University of Piraeus, Department of Informatics, 80, Karaoli and Dimitriou, 18534, Piraeus, Greece. Tel.: +302104142315, Fax: +302104142357.

1. Introduction

Data envelopment analysis (DEA) is the leading technique for measuring the efficiency of decision making units (DMU) in the presence of multiple inputs and outputs. The two milestone DEA models, namely the CCR [4] and the BCC [1] models have become standards in the literature of performance measurement under the assumptions of constant and variable returns-to-scale respectively. Typically, a single stage production process is assumed, that transforms inputs to final outputs. However, there is an increasing literature body that is devoted to the efficiency assessment in multistage production processes. Castelli et al. [2] provide a comprehensive categorized overview of models and methods developed for different multi-stage production architectures. In this paper, however, we focus on the typical architecture of a two-stage production process, which assumes that the external inputs entering the first stage of the process are transformed to a number of intermediate measures that are then used as inputs to the second stage to produce the final outputs. In this model, nothing but the external inputs to the first stage enters the system and nothing but the outputs of the second stage leaves the system. Seiford and Zhu [11] studied such a production process in the banking sector by treating the two stages independently, i.e. without assuming any relationship between the two stages. Kao and Hwang [8] introduced a novel approach that takes into account a series relationship of the two stages and developed a model that estimates the overall efficiency of the production process as the product of the efficiencies of the two individual stages. Their approach is based on the reasonable assumption that the values of the intermediate measures (virtual intermediate measures) are the same, no matter if they are considered as outputs of the first stage or inputs to the second stage. This multiplicative approach to efficiency decomposition is restricted to constant returns-to-scale (CRS) situations. Chen et al. [5] introduced the additive efficiency decomposition in two-stage process under the assumption of series relationship. They derive the overall efficiency of the production process as a weighted average of the efficiencies of the individual stages. Their modeling approach facilitates the linearization of a non-linear mathematical program and is based on the assumption that the weighting of the two stages derives endogenously by the optimization process, in a manner that reflects the size of the two stages. The additive decomposition approach is extendable to variable returns-to-scale (VRS) situations. Liang et al. [10] view the efficiency assessments in two-stage process in terms of a game approach.

In this paper we present an alternative additive decomposition approach in two-stage DEA under the common assumption of the series relationship of the two stages. In such a setup, we maintain the assumption that the virtual intermediate measures are common in both stages. Selecting an output orientation for the first stage and an input orientation for the second stage, we show that it is possible to aggregate additively the efficiency measures of the two individual stages in a bi-objective linear program. Our model estimates simultaneously optimal efficiency scores for the two stages, which then are used to calculate the overall efficiency of the production process as a simple average. However, if it is to assign different importance to the two stages, a weighted average could be calculated with a priori and externally defined weights. Our model is easily extended to a VRS variant. Our experiments show that efficiency scores obtained by our approach for the individual stages are comparable to those obtained in [5].

The paper unfolds as follows. In section two we outline the two basic approaches for the two-stage DEA: The multiplicative approach [8] and the additive approach [5]. In section three we present our approach and we formulate a linear model that assesses efficiency scores for the two stages under the CRS assumption. Then we give its VRS variant. In section four we apply our models to two data sets obtained from the literature and we compare our results with those reported in [5]. In section five we discuss some further issues raised in the literature as for the deficiencies and limitations observed in two-stage DEA models. Concluding remarks are given in section six.

2. Multiplicative and additive decomposition in two-stage DEA

Consider the generic case where each DMUj, j=1,…, n transforms inputs x to final outputs y with a two-stage process as shown in Fig.1.

>> Figure 1 about here <<

Assume n units (j=1,…,n), each using m inputs x_{ij}, i=1,…,m to the first stage to produce q outputs z_{pj}, p=1,…,q from that stage. The outputs obtained from the first stage are then used as inputs to the second stage to produce s final outputs y_{rj}, r=1,…,s. Treating the two stages independently, the stage 1 and stage 2 CRS efficiency scores for the evaluated unit j_0 are obtained from the following two conventional CCR DEA models (1) and (2) respectively:

Stage 1

$$
\begin{aligned}
& E_{j_0}^1 = \max \frac{\sum_{p=1}^{q} \varphi_p z_{pj_0}}{\sum_{i=1}^{m} \eta_i x_{ij_0}} \\
& s.t. \\
& \sum_{p=1}^{q} \varphi_p z_{pj} - \sum_{i=1}^{m} \eta_i x_{ij} \leq 0, \quad j = 1,...,n \\
& \eta_i \geq 0, \varphi_p \geq 0, \quad i = 1,...,m;\ p = 1,...,q
\end{aligned}
\qquad (1)
$$

Stage 2

$$E_{j_0}^{2} = \max \frac{\sum_{r=1}^{s} \omega_r y_{rj_0}}{\sum_{p=1}^{q} \hat{\varphi}_p z_{pj_0}}$$

$$\begin{aligned} & s.t. \\ & \sum_{r=1}^{s} \omega_r y_{rj} - \sum_{p=1}^{q} \hat{\varphi}_p z_{pj} \le 0, \quad j = 1,\dots,n \\ & \hat{\varphi}_p \ge 0, \omega_r \ge 0 \quad p = 1,\dots,q;\ r = 1,\dots,s \end{aligned} \tag{2}$$

The independent overall efficiency score of unit j_0 is similarly obtained by the following CCR DEA model (3):

Overall

$$E_{j_0}^{o} = \max \frac{\sum_{r=1}^{s} \omega_r y_{rj_0}}{\sum_{i=1}^{m} \eta_i x_{ij_0}}$$

$$\begin{aligned} & s.t. \\ & \sum_{r=1}^{s} \omega_r y_{rj} - \sum_{i=1}^{m} \eta_i x_{ij} \le 0, \quad j = 1,\dots,n \\ & \eta_i \ge 0, \omega_r \ge 0 \quad i = 1,\dots,m;\ r = 1,\dots,s \end{aligned} \tag{3}$$

To link the efficiency assessments of the two stages and to obtain jointly the overall efficiency score of the unit j_0, Kao and Hwang [8] assumed that the total virtual output $\sum_{p=1}^{q} \varphi_p z_{pj_0}$ of the first stage equals the total virtual input $\sum_{p=1}^{q} \hat{\varphi}_p z_{pj_0}$ that feeds the second stage (i.e. $\hat{\varphi}_p = \varphi_p, p = 1,\dots,q$). Based on this assumption, the overall efficiency score of unit j_0 is obtained by aggregating multiplicatively the efficiencies of the two stages as follows:

$$
\begin{aligned}
& e_{j_0}^{o} = \max \frac{\sum_{p=1}^{q}\varphi_p z_{pj_0}}{\sum_{i=1}^{m}\eta_i x_{ij_0}} \cdot \frac{\sum_{r=1}^{s}\omega_r y_{rj_0}}{\sum_{p=1}^{q}\varphi_p z_{pj_0}} = \frac{\sum_{r=1}^{s}\omega_r y_{rj_0}}{\sum_{i=1}^{m}\eta_i x_{ij_0}} \\
& s.t. \\
& \sum_{r=1}^{s}\omega_r y_{rj} - \sum_{p=1}^{q}\varphi_p z_{pj} \le 0, \quad j=1,\dots,n \\
& \sum_{p=1}^{q}\varphi_p z_{pj} - \sum_{i=1}^{m}\eta_i x_{ij} \le 0, \quad j=1,\dots,n \\
& \eta_i \ge 0, \varphi_p \ge 0, \omega_r \ge 0 \quad i=1,\dots,m\,;\; p=1,\dots,q\,;\; r=1,\dots,s
\end{aligned} \tag{4}
$$

Notice that the constraint $\sum_{r=1}^{s}\omega_r y_{rj} - \sum_{i=1}^{m}\eta_i x_{ij} \le 0$ included in the original model has been omitted in (4) as it is redundant. Applying the Charnes and Cooper [3] transformation (C-C transformation hereafter) to the fractional program (4), the following linear equivalent is obtained and solved for one unit at a time:

$$
\begin{aligned}
& e_{j_0}^{o} = \max \sum_{r=1}^{s} u_r y_{rj_0} \\
& s.t. \\
& \sum_{i=1}^{m} v_i x_{ij_0} = 1 \\
& \sum_{r=1}^{s} u_r y_{rj} - \sum_{p=1}^{q} w_p z_{pj} \le 0, \quad j=1,\dots,n \\
& \sum_{p=1}^{q} w_p z_{pj} - \sum_{i=1}^{m} v_i x_{ij} \le 0, \quad j=1,\dots,n \\
& v_i \ge 0, w_p \ge 0, u_r \ge 0 \quad i=1,\dots,m\,;\; p=1,\dots,q\,;\; r=1,\dots,s
\end{aligned} \tag{5}
$$

Once an optimal solution (u_r^*, v_i^*, w_p^*) of model (5) is obtained, the first stage, the second stage and the overall efficiency scores $e_{j_0}^1, e_{j_0}^2, e_{j_0}^o$ of the evaluated unit j_0 are obtained respectively by the following relations:

$$
e_{j_o}^1 = \frac{\sum_{p=1}^{q} w_p^* z_{pj_0}}{\sum_{i=1}^{m} v_i^* x_{ij_0}}, \; e_{j_o}^2 = \frac{\sum_{r=1}^{s} u_r^* y_{rj_0}}{\sum_{p=1}^{q} w_p^* z_{pj_0}}, \; e_{j_o}^o = e_{j_o}^1 \cdot e_{j_o}^2 = \frac{\sum_{r=1}^{s} u_r^* y_{rj_0}}{\sum_{i=1}^{m} v_i^* x_{ij_0}}
$$

Model (5) cannot be readily extended to treat DEA assessments under the VRS assumption. Working with BCC models of different orientations for the individual stages, Kao and Hwang [9] proposed an approach to decompose technical and scale efficiencies under the multiplicative decomposition model.

Chen et al. [5] developed an alternative two-stage DEA model by assuming a weighted average of the efficiencies of the two stages as follows:

$$
\begin{aligned}
&\max\ t_1 \frac{\sum_{p=1}^{q} \varphi_p z_{pj_0}}{\sum_{i=1}^{m} \eta_i x_{ij_0}} + t_2 \frac{\sum_{r=1}^{s} \omega_r y_{rj_0}}{\sum_{p=1}^{q} \varphi_p z_{pj_0}} \\
&s.t. \\
&\sum_{r=1}^{s} \omega_r y_{rj} - \sum_{p=1}^{q} \varphi_p z_{pj} \le 0, \quad j = 1,...,n \\
&\sum_{p=1}^{q} \varphi_p z_{pj} - \sum_{i=1}^{m} \eta_i x_{ij} \le 0, \quad j = 1,...,n \\
&\eta_i \ge 0, \varphi_p \ge 0, \omega_r \ge 0 \quad i = 1,...,m\,;\ p = 1,...,q\,;\ r = 1,...,s
\end{aligned}
\tag{6}
$$

To enable the transformation of (6) to a linear equivalent, they assumed further that the weights t_1 and t_2 are endogenously defined as functions of the variables, as:

$$
t_1 = \frac{\sum_{i=1}^{m} \eta_i x_{ij_0}}{\sum_{i=1}^{m} \eta_i x_{ij_0} + \sum_{p=1}^{q} \varphi_p z_{pj_0}}, \quad t_2 = \frac{\sum_{p=1}^{q} \varphi_p z_{pj_0}}{\sum_{i=1}^{m} \eta_i x_{ij_0} + \sum_{p=1}^{q} \varphi_p z_{pj_0}}
$$

Substituting t_1 and t_2 in model (6) they derive the following model under the CRS assumption:

$$
\begin{aligned}
&\max \frac{\sum_{p=1}^{q} \varphi_p z_{pj_0} + \sum_{r=1}^{s} \omega_r y_{rj_0}}{\sum_{i=1}^{m} \eta_i x_{ij_0} + \sum_{p=1}^{q} \varphi_p z_{pj_0}} \\
&s.t. \\
&\sum_{r=1}^{s} \omega_r y_{rj} - \sum_{p=1}^{q} \varphi_p z_{pj} \le 0, \quad j = 1,...,n \\
&\sum_{p=1}^{q} \varphi_p z_{pj} - \sum_{i=1}^{m} \eta_i x_{ij} \le 0, \quad j = 1,...,n \\
&\eta_i \ge 0, \varphi_p \ge 0, \omega_r \ge 0 \quad i = 1,...,m\,;\ p = 1,...,q\,;\ r = 1,...,s
\end{aligned}
\tag{7}
$$

Applying the C-C transformation, the linear equivalent of (7) is as follows:

$$
\begin{aligned}
&\max \sum_{r=1}^{s} u_r y_{rj_0} + \sum_{p=1}^{q} w_p z_{pj_0} \\
&s.t. \\
&\sum_{i=1}^{m} v_i x_{ij_0} + \sum_{p=1}^{q} w_p z_{pj_0} = 1 \\
&\sum_{r=1}^{s} u_r y_{rj} - \sum_{p=1}^{q} w_p z_{pj} \leq 0, \quad j = 1,\ldots,n \\
&\sum_{p=1}^{q} w_p z_{pj} - \sum_{i=1}^{m} v_i x_{ij} \leq 0, \quad j = 1,\ldots,n \\
&v_i \geq 0, w_p \geq 0, u_r \geq 0 \quad i = 1,\ldots,m \,;\, p = 1,\ldots,q \,;\, r = 1,\ldots,s
\end{aligned} \tag{8}
$$

The optimal solution of (8) can be used to calculate the efficiency scores $\theta^1_{j_0}, \theta^2_{j_0}$ of unit j_0 for the two individual stages and then the overall efficiency $\theta^0_{j_0} = t^*_1\theta^1_{j_0} + t^*_2\theta^2_{j_0}$, where t^*_1, t^*_2 are the weights calculated a posteriori on the basis of the optimal solution of (8). Notice, however, that the overall efficiency of j_0 derives also as the optimal value of the objective function in (8). In case of multiple optimal solutions in (8), two extra linear programs are solved to calculate $\theta^1_{j_0}, \theta^2_{j_0}$ [5]. The above additive decomposition approach enables the extension of model (8) to a variant that can be used under the VRS assumption [5].

3. An alternative additive model for two-stage DEA

Consider the linear equivalent of the output oriented variant of the first-stage model (1):

Stage 1: output oriented

$$
\begin{aligned}
&\frac{1}{E^1_{j_0}} = \min \sum_{i=1}^{m} v_i x_{ij_0} \\
&s.t. \\
&\sum_{p=1}^{q} w_p z_{pj_0} = 1 \\
&\sum_{p=1}^{q} w_p z_{pj} - \sum_{i=1}^{m} v_i x_{ij} \leq 0, \quad j = 1,\ldots,n \\
&v_i \geq 0, w_p \geq 0, \quad i = 1,\ldots,m;\ p = 1,\ldots,q
\end{aligned} \tag{9}
$$

and the linear equivalent of the second-stage model (2):

Stage2: input oriented

$$
\begin{aligned}
& E_{j_0}^2 = \max \sum_{r=1}^{s} u_r y_{rj_0} \\
& s.t. \\
& \sum_{p=1}^{q} w_p z_{pj_0} = 1 \\
& \sum_{r=1}^{s} u_r y_{rj} - \sum_{p=1}^{q} w_p z_{pj} \leq 0, \quad j = 1,\ldots,n \\
& w_p \geq 0, u_r \geq 0 \quad p = 1,\ldots,q;\ r = 1,\ldots,s
\end{aligned} \tag{10}
$$

Appending the constraints $\sum_{r=1}^{s} u_r y_{rj} - \sum_{p=1}^{q} w_p z_{pj} \leq 0,\ j = 1,\ldots,n$ to model (9) we derive the following augmented model for the first stage:

$$
\begin{aligned}
& \min \sum_{i=1}^{m} v_i x_{ij_0} \\
& s.t. \\
& \sum_{p=1}^{q} w_p z_{pj_0} = 1 \\
& \sum_{p=1}^{q} w_p z_{pj} - \sum_{i=1}^{m} v_i x_{ij} \leq 0, \quad j = 1,\ldots,n \\
& \sum_{r=1}^{s} u_r y_{rj} - \sum_{p=1}^{q} w_p z_{pj} \leq 0, \quad j = 1,\ldots,n \\
& v_i \geq 0, w_p \geq 0, u_r \geq 0 \quad i = 1,\ldots,m;\ p = 1,\ldots,q;\ r = 1,\ldots,s
\end{aligned} \tag{11}
$$

Similarly, adding the constraints $\sum_{p=1}^{q} w_p z_{pj} - \sum_{i=1}^{m} v_i x_{ij} \leq 0,\ j = 1,\ldots,n$ to model (10) we obtain the following augmented model for the second stage:

$$\begin{aligned}
&\max \sum_{r=1}^{s} u_r y_{rj_0} \\
&s.t. \\
&\sum_{p=1}^{q} w_p z_{pj_0} = 1 \\
&\sum_{p=1}^{q} w_p z_{pj} - \sum_{i=1}^{m} v_i x_{ij} \le 0, \quad j = 1,\dots,n \\
&\sum_{r=1}^{s} u_r y_{rj} - \sum_{p=1}^{q} w_p z_{pj} \le 0 \quad j = 1,\dots,n \\
&v_i \ge 0, w_p \ge 0, u_r \ge 0 \quad i = 1,\dots,m;\ p = 1,\dots,q;\ r = 1,\dots,s
\end{aligned} \tag{12}$$

Notice that an optimal solution of model (9) is also optimal in model (11). Indeed, one can always choose small enough values for u in model (11) to make any optimal solution of model (9) feasible, yet optimal, in model (11). Analogously, an optimal solution of model (10) is also optimal in model (12), as one can choose large enough values for v in model (12) to make any optimal solution of model (10) feasible, yet optimal, in model (12). For the completeness of our developments, compact proofs of these statements are given in Appendix.

Models (11) and (12) have common constraints. The need to formulate these two models is now apparent; they enable us to jointly consider them as a bi-objective linear program. Aggregating the two objective functions additively, we derive the following single-objective linear program:

$$\begin{aligned}
&\max F = \sum_{r=1}^{s} u_r y_{rj_0} - \sum_{i=1}^{m} v_i x_{ij_0} \\
&s.t. \\
&\sum_{p=1}^{q} w_p z_{pj_0} = 1 \\
&\sum_{p=1}^{q} w_p z_{pj} - \sum_{i=1}^{m} v_i x_{ij} \le 0, \quad j = 1,\dots,n \\
&\sum_{r=1}^{s} u_r y_{rj} - \sum_{p=1}^{q} w_p z_{pj} \le 0, \quad j = 1,\dots,n \\
&v_i \ge 0, w_p \ge 0, u_r \ge 0 \quad i = 1,\dots,m\ ;\ p = 1,\dots,q\ ;\ r = 1,\dots,s
\end{aligned} \tag{13}$$

Once an optimal solution (u_r^*, v_i^*, w_p^*) of model (13) is obtained, the efficiency scores for unit j_0 in the first and the second stage are respectively:

$$e_{j_0}^1 = \frac{1}{\sum_{i=1}^{m} v_i^* x_{ij_0}} \quad \text{and} \quad e_{j_0}^2 = \sum_{r=1}^{s} u_r^* y_{rj_0}$$

Notice that the unit j_0 is overall efficient, if and only if the optimal value of the objective function is zero (F*=0). Model (13) does not provide a direct measure of the overall efficiency, as it is the case in the multiplicative model (5) and the additive model (8). As noticed in [10], it is reasonable to define the overall efficiency of the two-stage process as the average (arithmetic mean) of the efficiencies of the two individual stages. In this line of thought, the overall efficiency of unit j_0 is defined as $e^o_{j_0} = (e^1_{j_0} + e^2_{j_0})/2$.

Our developments are based on the selection of the output orientation for the first stage and the input orientation for the second stage. This is the key that enables us to aggregate the two stages in an additive form, without the need to assume weights for the two stages. Hence, our approach can be considered as "neutral", as opposed to the Chen's et al. [5] one, where, for the sake of linearization, the unit under evaluation assigns its own weights to the efficiency scores of the two individual stages. Nevertheless, if it is to assign different importance to each of the two stages, one might consider as well weights a_1, a_2 (a_1+a_2=1) to compute the overall efficiency $e^o_{j_0} = a_1 e^1_{j_0} + a_2 e^2_{j_0}$. The difference between such weights and the weights t_1 and t_2 assumed in (6) is that they are specified a priori by the user and are common for all the DMUs. Going one step further, in line with the argument that the "size" of a stage reflects its importance [5], the weights could be defined as:

$$a_1 = \frac{\sum_{j=1}^{n}\sum_{i=1}^{m}\overline{x}_{ij}}{\sum_{j=1}^{n}\sum_{i=1}^{m}\overline{x}_{ij} + \sum_{j=1}^{n}\sum_{p=1}^{q}\overline{z}_{pj}}, \quad a_2 = \frac{\sum_{j=1}^{n}\sum_{p=1}^{q}\overline{z}_{pj}}{\sum_{j=1}^{n}\sum_{i=1}^{m}\overline{x}_{ij} + \sum_{j=1}^{n}\sum_{p=1}^{q}\overline{z}_{pj}}$$

where $\overline{x}_{ij}$ and $\overline{z}_{pj}$ derive by max-normalizing the raw data, column-wise, i.e.:

$$\overline{x}_{ij} = \frac{x_{ij}}{\max_j\{x_{ij}\}}, \quad \overline{z}_{pj} = \frac{z_{pj}}{\max_j\{z_{pj}\}}$$

In [5], the size of a stage is represented by the portion of the total resources used in each stage by the evaluated unit, in terms of values (virtual inputs). Hence the size is viewed differently from each DMU. Let us call this perspective a "*DMU-centric perspective*". Our approach to weighting the two stages is based on a "*stage-centric perspective*", as the size of a stage is represented by the portion of the total resources used in each stage by all the DMUs, in terms of the raw quantities. Actually, the raw quantities are max-normalized to make them units free.

Model (13) may have multiple optimal solutions and, thus, the decomposition may not by unique. To make the efficiency assessments comparable across all the units, we address this issue in a manner analogous to those proposed in Kao and Hwang [8] and Chen et al. [5]. Particularly, in a post-optimality stage, we seek the largest efficiency score in the first or the second stage (depending on the given priority), while retaining the optimal value F^* of the objective function in model (13). So, if priority is given to the first stage, the highest efficiency score e^1_{1,j_0} for unit j_o can be obtained from model (11), after appending to it the additional constraint $\sum_{r=1}^{s} u_r y_{rj_0} - \sum_{i=1}^{m} v_i x_{ij_0} = F^*$. If u^1_r, v^1_i, w^1_p is the optimal solution derived in the post-optimality stage, in favour of the first stage, the stage-1 and stage-2 efficiency scores of unit j_0 are respectively:

$$e^1_{1,j_0} = \frac{1}{\sum_{i=1}^{m} v^1_i x_{ij_0}} \text{ and } e^2_{1,j_0} = \sum_{r=1}^{s} u^1_r y_{rj_0}$$

Analogous is the derivation of the efficiency scores if priority is given to the second stage. The highest stage-2 efficiency score e^2_{2,j_0} for unit j_0 is obtained from model (12), after appending the same, as above, constraint to retain the optimal F*. Then e^1_{2,j_0} is obtained from the corresponding post-optimal solution. Apparently, if $e^1_{1,j_0} = e^1_{2,j_0}$ or $e^2_{1,j_0} = e^2_{2,j_0}$ the efficiency decomposition provided by model (13) is unique.

Our approach to the additive efficiency decomposition enables us to extend our developments under the VRS assumption. Indeed, the VRS variant of model (13) can be obtained from the VRS variants of (9) and (10) as follows:

$$\begin{aligned}
&\max F = \sum_{r=1}^{s} u_r y_{rj_0} - d_2 - \sum_{i=1}^{m} v_i x_{ij_0} + d_1 \\
&s.t. \\
&\sum_{p=1}^{q} w_p z_{pj_0} = 1 \\
&\sum_{p=1}^{q} w_p z_{pj} - \sum_{i=1}^{m} v_i x_{ij} + d_1 \le 0, \quad j = 1,...,n \\
&\sum_{r=1}^{s} u_r y_{rj} - \sum_{p=1}^{q} w_p z_{pj} - d_2 \le 0, \quad j = 1,...,n \\
&v_i \ge 0, w_p \ge 0, u_r \ge 0 \quad i = 1,...,m;\ p = 1,...,q;\ r = 1,...,s
\end{aligned} \tag{14}$$

4. Applications

First we apply our approach to the 24 Taiwanese non-life insurance companies originally studied in Kao and Hwang [8]. The authors consider a two-stage production process with two inputs (Operation expenses-X1 and Insurance expenses-X2), two intermediate measures (Direct written premiums-Z1 and Reinsurance premiums-Z2) and two final outputs (Underwriting profit-Y1 and Investment profit-Y2). Table 1 exhibits the data set.

>> Table 1 about here <<

Table 2 reports the efficiency scores obtained by applying model (13) on the data of Table 1 (third to fifth columns) and the corresponding results reported in [5] along with the weights used (last five columns).

>> Table 2 about here <<

The two additive approaches provide the same efficiency scores for the individual stages for all units but one; the DMU 16 (Allianz President), where one can spot the only difference when comparing e^1 and e^2 with θ^1 and θ^2 respectively. The overall efficiency scores e^o and θ^o cannot be compared directly, as the former is calculated as a simple average while the latter is derived as a weighted average, with the weights varying across the DMUs. Obviously, when equal weights w_1 and w_2 are assigned to the individual stages, the overall efficiency scores are identical. This is the case of DMUs 2, 9, 12, 15, 19 and 24.

Table 3 summarizes the results obtained from model (14) and the corresponding results given in [5] under the VRS assumption.

>> Table 3 about here <<

In the standard DEA approach, the efficiency scores obtained under the VRS assumption are not less than their counterparts under the CRS assumption. Although this is true in our additive two-stage DEA models for the overall efficiency scores, the results show that not all the intermediate efficiency scores comply with this conventional principle. This is the case for the DMUs 12 and 20, with respect to their first stage efficiency scores e^1, and for DMU 18 with respect to the second stage efficiency e^2. A similar irregularity has been spotted in Chen et al. [5].

To extend our comparisons, we apply our approach and then Chen's et al. [5] additive model to another data set, originally used in Wang et al. [12] and later in Chen and Zhu [6], in investigating the impact of information technology on productivity. There are 27 units in the study evaluated on three inputs (Fixed assets-X1, IT budget-X2 and Number of employees-X3), a single intermediate measure (Deposits-Z1) and two final outputs (Profit-Y1 and Fraction of loans recovered-Y2). The data set is given in Table 4.

>> Table 4 about here <<

Table 5 reports the efficiency scores obtained by applying model (13) on the IT data of Table 4 (second to fourth columns) and the corresponding scores along with the weights obtained by our calculations based on the model of Chen et al. [5] (last five columns). As concerns the efficiency scores for the two individual stages, the results obtained from the two models are identical. However, the overall efficiency scores e^o

and θ^{o} differentiate. Indeed, there are numerous units (thirteen of the twenty-seven DMUs), for which $e^{o} > \theta^{o}$. DMU 18 has been commonly identified by both models as overall efficient.

>> Table 5 about here <<

The post-optimality stage applied to both examples showed that the efficiency decompositions obtained from models (13) and (14) are unique.

6. Conclusion

We presented in this paper an alternative model for two-stage DEA under the assumption of series relationship between the two stages. Our modeling approach is based on the selection of an output orientation for the first stage and an input orientation for the second stage. In this manner, the intermediate measures are used as pivot that links the efficiency assessment models for the two stages in a single linear program. The proposed CRS model is straightforwardly extended to fit VRS situations. The additive efficiency decomposition approach coined in this paper is straightforward and, thus, free of the weighting assumption made in the original additive model [5]. Testing our models with data sets taken from previous studies, shows that the results obtained are comparable to those reported in the literature.

Appendix

An optimal solution of model (9) is also optimal in model (11).

Proof:

Let $v_i^*, i=1,...,m$ and $w_p^*, p=1,...,q$ be an optimal solution of (9). First we will show that this solution is feasible in (11). Indeed, it satisfies the first two constraints of (11), as they are identical to the constraints in (9). Notice that the first two constraints of (11) are independent of the variables $u_r, r=1,...,s$, which appear only in the third constraint. Then,

(a) If $s \leq q$, the third constraint of (11) is satisfied for

$$u_r = \frac{w_r^* z_r^{\min}}{y_r^{\max}} \geq 0, \quad r=1,...,s$$

where $z_r^{\min} = \min_j \{z_{rj}\}$ is the smallest observed value of the intermediate measure z_r and $y_r^{\max} = \max_j \{y_{rj}\}$ is the largest observed value of output y_r.

(b) If $s>q$, the third constraint of (11) is satisfied for

$$u_p = \frac{w_p^* z_p^{\min}}{y_p^{\max}} \geq 0,\ p = 1,...,q,\ u_r = 0,\ r = q+1,...,s$$

Thus, the optimal solution $v_i^*, i = 1,...,m$ and $w_p^*, p = 1,...,q$ of (9) is a feasible solution of (11). Moreover, as the objective functions in both the (9) and (11) are independent of u_r, the above solution is optimal in (11) as well. □

An optimal solution of model (10) is also optimal in model (12).

Proof:

Let $u_r^*, r = 1,...,s$ and $w_p^*, p = 1,...,q$ be an optimal solution of (10). First we will show that this solution is feasible in (12). Indeed, it satisfies the first and the third constraint of (12), as they are identical to the constraints in (10). Notice that the first and the third constraint of (12) are independent of the variables $v_i, i = 1,...,m$, which appear only in the second constraint. Then,

(a) If $q \leq m$, the second constraint of (12) is satisfied for

$$v_p = \frac{w_p^* z_p^{\max}}{x_p^{\min}},\ p = 1,...q\ ,\ v_i \geq 0, i = q+1,...,m$$

where $z_p^{\max} = \max_j \{z_{pj}\}$ is the largest observed value of the intermediate measure z_p and $x_p^{\min} = \min_j \{x_{pj}\}$ is the smallest observed value of the input x_p.

(b) If $q>m$, the second constraint of (12) is satisfied for

$$v_i = \frac{w_i^* z_i^{\max}}{x_i^{\min}},\ i = 1,...m-1$$

$$v_m = \frac{w_m^* z_m^{\max}}{x_m^{\min}} + \sum_{p=m+1}^{q} \frac{w_p^* z_p^{\max}}{x_m^{\min}}$$

Thus, the optimal solution $u_r^*, r = 1,...,s$ and $w_p^*, p = 1,...,q$ of (10) is a feasible solution of (12). Moreover, as the objective functions in both the (10) and (12) are independent of v_i, the above solution is optimal in (12) as well.□

References

[1] R.D. Banker, A. Charnes and W.W. Cooper, *Some models for estimating technical and scale inefficiencies in data envelopment analysis*, Manage. Sci. 30 (1984), pp. 1078-1092.

[2] L. Castelli, R. Pesenti and W. Ukovich, *A classification of DEA models when the internal structure of the Decision Making Units is considered,* Ann. Oper. Res. 173 (2010), pp. 207-235.

[3] A. Charnes and W.W. Cooper, *Programming with linear fractional functionals*, Nav. Res. Log. 9 (1962), pp.181-185.

[4] A. Charnes, W.W. Cooper and E. Rhodes, *Measuring the efficiency of decision making units*, Eur. J. Oper. Res. 2 (1978), pp. 429-444.

[5] Y. Chen, Y., W.D., Cook, N. Li and J. *Zhu, Additive efficiency decomposition in two-stage DEA*, Eur. J. Oper. Res. 196 (2009), pp. 1170-1176.

[6] Y. Chen and J. Zhu, *Measuring information technology's indirect impact on firm performance*. Inf. Technol. Manage. 5 (2004), pp.9–22.

[7] Y. Chen, Y., W.D., Cook and J. *Zhu, Deriving the DEA frontier for two-stage DEA processes*, Eur. J. Oper. Res. 202 (2010), pp. 138-142.
[8] C. Kao and S-N. Hwang, *Efficiency decomposition in two-stage data envelopment analysis: An application to non-life insurance companies in Taiwan,* Eur. J. Oper. Res. 185 (2008), pp. 418–429.
[9] C. Kao and S-N. Hwang, *Decomposition of technical and scale efficiencies in two-stage production systems*, Eur. J. Oper. Res. 211 (2011), pp. 515-519.
[10] L. Liang, W.D. Cook and J. Zhu, *DEA models for two-stage processes: Game approach and efficiency decomposition*, Nav. Res. Log. 55 (2008), pp. 643-653.
[11] L.M. Seiford and J. Zhu, 1999. *Profitability and marketability of the top 55 US commercial banks*, Manage. Sci. 45 (1999), pp. 1270-1288.
[12] C.H. Wang, R. Gopal and S. Zionts, *Use of data envelopment analysis in assessing information technology impact on firm performance*, Ann. Oper. Res. 73 (1997), pp. 191–213.

Figures

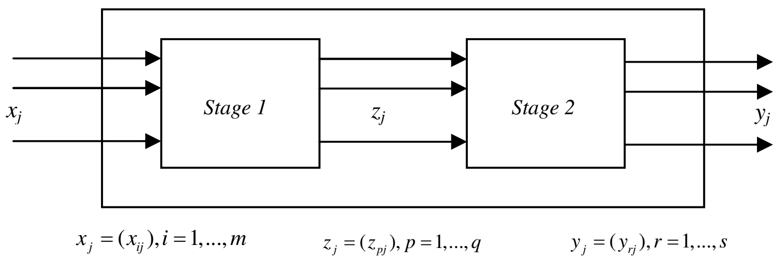

Figure 1: The architecture of a generic two-stage process

Tables

Table 1: Taiwanese non-life insurance companies data set (source: Kao and Hwang [8]).

#	DMU	X1	X2	Z1	Z2	Y1	Y2
1	Taiwan Fire	1178744	673512	7451757	856735	984143	681687
2	Chung Kuo	1381822	1352755	10020274	1812894	1228502	834754
3	Tai Ping	1177494	592790	4776548	560244	293613	658428
4	China Mariners	601320	594259	3174851	371863	248709	177331
5	Fubon	6699063	3531614	37392862	1753794	7851229	3925272
6	Zurich	2627707	668363	9747908	952326	1713598	415058
7	Taian	1942833	1443100	10685457	643412	2239593	439039
8	Ming Tai	3789001	1873530	17267266	1134600	3899530	622868
9	Central	1567746	950432	11473162	546337	1043778	264098
10	The First	1303249	1298470	8210389	504528	1697941	554806
11	Kuo Hua	1962448	672414	7222378	643178	1486014	18259
12	Union	2592790	650952	9434406	1118489	1574191	909295
13	Shing kong	2609941	1368802	13921464	811343	3609236	223047
14	South China	1396002	988888	7396396	465509	1401200	332283
15	Cathay Century	2184944	651063	10422297	749893	3355197	555482
16	Allianz President	1211716	415071	5606013	402881	854054	197947
17	Newa	1453797	1085019	7695461	342489	3144484	371984
18	AIU	757515	547997	3631484	995620	692731	163927
19	North America	159422	182338	1141950	483291	519121	46857
20	Federal	145442	53518	316829	131920	355624	26537
21	Royal & Sunalliance	84171	26224	225888	40542	51950	6491
22	Aisa	15993	10502	52063	14574	82141	4181
23	AXA	54693	28408	245910	49864	0.1	18980
24	Mitsui Sumitomo	163297	235094	476419	644816	142370	16976

Table 2: Results from model (13) compared to Chen et al. [5]

		Our CRS model (13)			Chen et al. [5] – CRS model				
#	DMU	e^1	e^2	$e^o=(e^1+e^2)/2$	θ^1	θ^2	θ^o	w_1	w_2
1	Taiwan Fire	0.993	0.704	0.849	0.993	0.704	0.849	0.502	0.498
2	Chung Kuo	0.998	0.626	0.812	0.998	0.626	0.812	0.500	0.500
3	Tai Ping	0.690	1	0.845	0.690	1	0.817	0.592	0.408
4	China Mariners	0.724	0.420	0.572	0.724	0.420	0.596	0.580	0.420
5	Fubon	0.831	0.923	0.877	0.831	0.923	0.873	0.546	0.454
6	Zurich	0.961	0.406	0.683	0.961	0.406	0.689	0.510	0.490
7	Taian	0.752	0.352	0.552	0.752	0.352	0.580	0.571	0.429
8	Ming Tai	0.726	0.378	0.552	0.726	0.378	0.579	0.580	0.420
9	Central	1	0.223	0.612	1	0.223	0.612	0.500	0.500
10	The First	0.862	0.541	0.701	0.862	0.541	0.713	0.537	0.463
11	Kuo Hua	0.729	0.207	0.468	0.729	0.207	0.509	0.578	0.422
12	Union	1	0.760	0.880	1	0.760	0.880	0.500	0.500
13	Shing kong	0.811	0.243	0.527	0.811	0.243	0.557	0.552	0.448
14	South China	0.725	0.374	0.549	0.725	0.374	0.577	0.580	0.420
15	Cathay Century	1	0.614	0.807	1	0.614	0.807	0.500	0.500
16	Allianz President	0.907	0.336	0.621	0.886	0.362	0.639	0.530	0.470
17	Newa	0.723	0.460	0.591	0.723	0.460	0.613	0.580	0.420
18	AIU	0.794	0.326	0.560	0.794	0.326	0.587	0.558	0.442
19	North America	1	0.411	0.706	1	0.411	0.706	0.500	0.500
20	Federal	0.933	0.586	0.759	0.933	0.586	0.765	0.517	0.483
21	Royal & Sunalliance	0.751	0.262	0.506	0.751	0.262	0.541	0.571	0.429
22	Aisa	0.590	1	0.795	0.590	1	0.742	0.629	0.371
23	AXA	0.843	0.499	0.671	0.843	0.499	0.685	0.543	0.457
24	Mitsui Sumitomo	1	0.087	0.544	1	0.087	0.544	0.500	0.500

Table 3: Results from model (14) compared to Chen et al. [5] under the VRS assumption

	Our VRS model (14)			Chen et al. [5] – VRS model				
DMU	e^1	e^2	$e^o=(e^1+e^2)/2$	θ^1	θ^2	θ^o	w_1	w_2
1	1	0.736	0.868	0.990	0.743	0.867	0.503	0.497
2	1	0.711	0.856	1	0.711	0.856	0.500	0.500
3	0.700	1	0.850	0.690	1	0.818	0.587	0.413
4	0.724	0.425	0.575	0.726	0.424	0.599	0.581	0.419
5	1	1	1	1	1	1	0.483	0.517
6	0.975	0.490	0.733	0.964	0.490	0.732	0.511	0.489
7	0.803	0.592	0.698	0.752	0.593	0.684	0.571	0.429
8	0.838	0.687	0.762	0.783	0.722	0.754	0.523	0.477
9	1	0.285	0.643	1	0.276	0.639	0.501	0.499
10	0.862	0.727	0.794	0.862	0.727	0.780	0.538	0.462
11	0.750	0.432	0.591	0.741	0.443	0.614	0.576	0.424
12	0.968	0.803	0.885	0.968	0.803	0.887	0.511	0.489
13	0.869	0.763	0.816	0.846	0.763	0.804	0.494	0.506
14	0.725	0.555	0.640	0.725	0.555	0.654	0.581	0.419
15	1	0.880	0.940	1	0.880	0.940	0.503	0.497
16	0.910	0.417	0.663	0.911	0.417	0.676	0.526	0.474
17	0.723	1	0.862	0.724	1	0.840	0.581	0.419
18	0.974	0.278	0.626	0.850	0.369	0.618	0.517	0.483
19	1	0.657	0.828	1	0.657	0.833	0.515	0.485
20	0.894	1	0.947	0.902	1	0.946	0.548	0.452
21	0.895	0.362	0.628	0.913	0.362	0.679	0.575	0.425
22	1	1	1	1	1	1	0.634	0.366
23	0.972	0.620	0.796	0.976	0.620	0.815	0.547	0.453
24	1	0.101	0.551	1	0.098	0.564	0.517	0.483

Table 4: IT data (source: Wang et al. [12])

DMU	X1 Fixed assets ($billions)	X2 IT budget ($billions)	X3 Number of employees (thousand)	Z1 Deposits ($billions)	Y1 Profit ($billions)	Y2 Fraction of loans recovered
1	0.713	0.15	13.3	14.478	0.232	0.986
2	1.071	0.17	16.9	19.502	0.34	0.986
3	1.224	0.235	24	20.952	0.363	0.986
4	0.363	0.211	15.6	13.902	0.211	0.982
5	0.409	0.133	18.485	15.206	0.237	0.984
6	5.846	0.497	56.42	81.186	1.103	0.955
7	0.918	0.06	56.42	81.186	1.103	0.986
8	1.235	0.071	12	11.441	0.199	0.985
9	18.12	1.5	89.51	124.072	1.858	0.972
10	1.821	0.12	19.8	17.425	0.274	0.983
11	1.915	0.12	19.8	17.425	0.274	0.983
12	0.874	0.05	13.1	14.342	0.177	0.985
13	6.918	0.37	12.5	32.491	0.648	0.945
14	4.432	0.44	41.9	47.653	0.639	0.979
15	4.504	0.431	41.1	52.63	0.741	0.981
16	1.241	0.11	14.4	17.493	0.243	0.988
17	0.45	0.053	7.6	9.512	0.067	0.98
18	5.892	0.345	15.5	42.469	1.002	0.948
19	0.973	0.128	12.6	18.987	0.243	0.985
20	0.444	0.055	5.9	7.546	0.153	0.987
21	0.508	0.057	5.7	7.595	0.123	0.987
22	0.37	0.098	14.1	16.906	0.233	0.981
23	0.395	0.104	14.6	17.264	0.263	0.983
24	2.68	0.206	19.6	36.43	0.601	0.982
25	0.781	0.067	10.5	11.581	0.12	0.987
26	0.872	0.1	12.1	22.207	0.248	0.972
27	1.757	0.0106	12.7	20.67	0.253	0.988

Table 5: Results for IT data

	Our CRS model (13)			Chen et al. [5]- CRS model				
	e^1	e^2	e^o	θ^1	θ^2	θ^o	w_1	w_2
1	0.639	0.746	0.692	0.639	0.746	0.681	0.610	0.390
2	0.651	0.782	0.716	0.651	0.782	0.702	0.606	0.394
3	0.518	0.773	0.645	0.518	0.773	0.605	0.659	0.341
4	0.599	0.714	0.656	0.599	0.714	0.642	0.626	0.374
5	0.556	0.724	0.640	0.556	0.724	0.616	0.643	0.357
6	0.760	0.576	0.668	0.760	0.576	0.680	0.568	0.432
7	1	0.576	0.788	1	0.576	0.788	0.500	0.500
8	0.535	0.825	0.680	0.535	0.825	0.636	0.651	0.349
9	0.625	0.635	0.630	0.625	0.635	0.629	0.615	0.385
10	0.496	0.719	0.607	0.496	0.719	0.570	0.668	0.332
11	0.495	0.719	0.607	0.495	0.719	0.569	0.669	0.331
12	0.668	0.595	0.632	0.668	0.595	0.639	0.599	0.401
13	0.949	0.858	0.903	0.949	0.858	0.905	0.513	0.487
14	0.588	0.578	0.583	0.588	0.578	0.584	0.630	0.370
15	0.658	0.603	0.631	0.658	0.603	0.636	0.603	0.397
16	0.665	0.643	0.654	0.665	0.643	0.656	0.601	0.399
17	0.718	0.788	0.753	0.718	0.788	0.747	0.582	0.418
18	1	1	1	1	1	1	0.500	0.500
19	0.814	0.593	0.703	0.814	0.593	0.715	0.551	0.449
20	0.693	1	0.847	0.693	1	0.819	0.591	0.409
21	0.707	0.994	0.850	0.707	0.994	0.825	0.586	0.414
22	0.794	0.641	0.717	0.794	0.641	0.726	0.557	0.443
23	0.780	0.699	0.740	0.780	0.699	0.745	0.562	0.438
24	0.930	0.714	0.822	0.930	0.714	0.826	0.518	0.482
25	0.627	0.652	0.639	0.627	0.652	0.636	0.615	0.385
26	1	0.515	0.758	1	0.515	0.758	0.500	0.500
27	1	0.564	0.782	1	0.564	0.782	0.500	0.500