A fairer assessment of DMUs in a generalised two-stage DEA structure

Marios Dominikos Kremantzis*[1], Patrick Beullens†[2], and Jonathan Klein‡[3]

[1]School of Management, University of Bristol, UK

[2]Mathematical Sciences, Southampton Business School and CORMSIS, University of Southampton, UK

[3]Southampton Business School and CORMSIS, University of Southampton, UK

*Corresponding author E-mail address: marios.kremantzis@bristol.ac.uk

†p.beullens@soton.ac.uk

‡J.H.Klein@soton.ac.uk

Abstract

In Data Envelopment Analysis (DEA), a variety of approaches have been used in the context of single-stage and basic serial two-stage systems to attain fairness in the evaluation of decision-making units (DMUs). Little work, however, has been done to address this challenge in a generalised two-stage structure featuring additional inputs in the second stage and a proportion of first-stage outputs as final outputs. In this paper, we argue that in this context, fairness is enhanced by increasing measures related to the discriminatory power and the weighting scheme of the method. We describe a mechanism that gives prominence to a more contemporary concept of fairness, incorporating diversity and inclusion of minority opinions. These aspects have, to our knowledge, not yet received explicit attention in the methodological development of DEA. We propose a novel combination of an additive self-efficiency aggregation model, a minimax secondary goal model, and the CRiteria Importance Through Inter-criteria Correlation (CRITIC) method, in order to promote these aspects of fairness, and thus achieve a better degree of cooperation between the stages of a DMU and among DMUs. The additive aggregation model is chosen over the alternative multiplicative approach for a variety of reasons relating to the emphasis on the intermediate products exchanged and the simplification. The minimax model offers peer evaluation in which each DMU aims to evaluate the worst of the others in the best possible light. Application of the CRITIC method to DEA addresses the aggregation problem within the cross-efficiency concept. Practical applications of this approach could include supporting the determination of training needs in job rotation manufacturing, or evaluation of sustainable supply chains. The paper includes a description of a numerical experiment, illustrating the approach.

Keywords Data envelopment analysis; two-stage; fairness; cross-efficiency; CRITIC

1 Introduction

Data Envelopment Analysis (DEA) is a non-parametric approach for evaluating the performance of Decision-Making Units (DMUs) that use inputs to produce outputs **(Cook et al., 2014)**. DEA was developed by **Charnes et al. (1978)** (CCR) for the constant returns-to-scale assumption. Traditional DEA does not model the internal processes in a DMU. As a result, a relatively large proportion of DMUs emerge as DEA-efficient, without a means to distinguish them **(Ma et al., 2017)**. To enable the study of internal structures, research has extended DEA models to consider network structures **(Kao, 2009; Kao and Hwang, 2011; Wanke and Barros, 2014; Kao, 2014; Guo et al., 2017; Chen and Zhu, 2017; Örkcü et al., 2019)**. In a two-stage process in particular, inputs used by a DMU feed into a first stage, producing intermediate outputs that feed into a second stage, producing the final outputs of the entire system. Such a structure facilitates the measurement of both the overall system and its individual stages' efficiencies **(Mahdiloo et al., 2016)**.

Measuring the performance can be challenging when inputs and outputs are shared among different processes and are not easily distinguished **(Zha and Liang, 2010)**. **Yu and Shi (2014)** examine a two-stage structure with additional inputs in the second stage and some of the intermediate products as final outputs, towards building cooperative and leader-follower models. **Jianfeng (2015)** considers a network DEA model, in which the inputs are classified into those that are entirely integrated into one stage and those that are shared between the two stages. **Ma et al. (2017)** propose a parallel-series hybrid two-stage DEA model utilising the principles of additive and multiplicative efficiency decomposition.

While two-stage DEA models have the potential to increase managerial insight into the sources of inefficiency, two major problems similar to those in a single-stage emerge. The first one concerns the lack of discrimination power due to a high number of efficient DMUs **(Mahdiloo et al., 2016)**. The second challenge relates to an 'unrealistic' weighting scheme. Indeed, it is allowed for high relative-importance weights to be assigned to 'less important' inputs or outputs, and/or low weights to significant factors. This choice of weights could turn a DMU into an efficient unit **(Ghasemi et al., 2014)**.

In this paper we are interested in methods which aim to avoid a low degree of discrimination, unrealistic weight schemes, and to use a system of ranking that encourages cooperation by the units being evaluated. While doing so, we also wish to provide a mechanism that gives a voice for minority opinions. This aspect has, to our knowledge, not yet received explicit attention in the methodological development of DEA. In short, we say that we intend to tweak DEA methodology to improve the *fairness*[1] in the evaluation outcomes. We summarize the core literature, relevant to fairness evaluation in DEA, in Table 1.

Among those methods tested towards fairness is cross-efficiency (CE), which adds peer-evaluation to the self-evaluation principle **(Sexton et al., 1986)**. As stressed by **Anderson et al. (2002)**, CE improves the probability of obtaining a unique ranking. A critical drawback of CE is the non-uniqueness of optimal weights, which leads to the non-uniqueness of cross-efficiencies. To alleviate this, **Doyle and Green (1994)** recommended the adoption of alternative secondary goals in an aim to select unique optimal multipliers. In particular, they introduced an aggressive and a benevolent model, while the secondary objective functions in **Liang et al. (2008)** reflected the minimisation of total deviation, maximum deviation, and mean absolute deviation from an 'ideal' point. The interested reader could also check **Wang and Chin (2010a), Wang et al. (2011), Wu et al. (2012), Wu et al. (2016b)** and **Li et al. (2018)**. The non-uniqueness issue is also critical in a network system. **Kao and Liu (2019)** developed an aggressive CE model to measure the efficiency in two basic network structures. **Örkcü et al. (2019)** came up with a neutral CE model in a two-stage system, which is indifferent to the preference choice between the aggressive and benevolent formulations.

The aggregation of the cross-efficiency scores is another issue in CE. An appropriate aggregation strategy can enable the DMUs to accept their ranking. Although the average method has proven effective in ensuring a credible ranking **(Liang et al., 2008, Wang and Chin, 2010b)**, it loses sight of the weights assigned to scores **(Wang and Wang, 2013)**. To accommodate this issue, **Wu et al. (2011)** utilised the Shannon entropy, allocating a fixed but different weight to each DMU. **Wu et al. (2012a)** highlighted that this is problematic,

[1]No attempt is made to give a formal definition of fairness, but aspects which might reasonably be considered to contribute to this are discussed throughout this paper.

since it ignores the primary role of the self-evaluated efficiency of each DMU. They, thus, embedded the Shannon entropy into the CE by considering the association among the self and the peer-evaluation values. For more recent work on this, see **Wang and Chin (2011)**, **Wang and Wang (2013)**, and **Song and Liu (2018)**.

Fairness in the evaluation outcomes has been achieved even via the integration of game theoretic concepts within traditional single-stage and two-stage DEA networks. For instance, **Zhou et al. (2013)** introduced a Nash bargaining game model to obtain a unique efficiency decomposition for the two constituent sub-stages of the centralized model. Their approach leads to a fair context, in that it reflects how the two sub-stages bargain with each other for better efficiencies. **An et al. (2017)** also used Nash bargaining, but introduced a framework for setting fair target values for intermediate products of two-stage systems, so that the two stages are encouraged to collaborate with each other within a pre-agreed range of fair outcomes. **Wu et al. (2016a)** proposed a CE evaluation approach based on Pareto improvement. A merit of their approach is that it always generates a set of Pareto optimal cross-efficiencies for the DMUs. **Li (2017)** introduced a sequence of leader-follower procedures as to ensure a fair evaluation in the sense that it guarantees that the same result is obtained for the second (=follower) stage of a DMU as would be obtained applying the standard DEA model to the second stage independently. A number of studies have been reported in this direction, such as **Yu and Shi (2014)**, **Ma et al. (2014)**, and **Li et al. (2018)**.

Table 1: Related literature on fairness evaluation in DEA.

	Type of network	Cross-efficiency	Aggregation method	Game approach	Efficiency measurement
Zhou et al. (2013)	two-stage	×	×	Nash bargaining	decomposition
Yu and Shi (2014)	two-stage	×	×	cooperative & leader-follower	×
Ma et al. (2014)	two-stage	centralized	arithmetic average	non-cooperative inspired	decomposition
Wu et al. (2016a)	single-stage	Pareto improvement	arithmetic average	Pareto optimality	n/a
An et al. (2017)	two-stage	×	×	Nash bargaining	×
Li (2017)	two-stage	×	×	cooperative & leader-follower	×
Li et al. (2018)	single-stage	optimal balanced	balanced adjustment	game-like iterative algorithm	n/a
Örkcü et al. (2019)	two-stage	neutral	geometric average	×	×
This Paper	two-stage	minimax	CRITIC	×	aggregation

In summary, fairness in the evaluation of DMUs has been extensively explored via CE towards single-stage and basic network structures. Nevertheless, when the discussion shifts to more complex structures where inputs and outputs are shared among different processes, there is limited attention to how to achieve more meaningful results for the DMUs. This intricacy is due to the additional inputs in the second stage obtained from the external environment and the dual role of the intermediate products. There are several enlightening applications, especially in logistics, supply chain, and manufacturing, that could justify the necessity of exploring fairness in the performance evaluation of a generalised two-stage structure. These are discussed in more depth with an example in Section 3 and in the implications of Section 4.2.

In our paper, we firstly introduce an additive self-efficiency aggregation model that can highlight the strength of each sub-stage and obtain the most favourable efficiency for the DMU overall. Since the optimal set of weights derived from the aggregation model may not be unique, we employ a minimax secondary goal model. The reasons for the adoption of this model are twofold: *(i)* it corresponds to cooperative situations **(Liang et al., 2008b)**, since sub-stages behave benignly, and *(ii)* it is compatible with multi-stage systems where the individual sub-stages pursue mutual cooperation via the maximisation of the overall efficiency **(Yu and Shi, 2014)**. The multi-objective model is converted using the Compromise Programming methodology as a means to identify a good solution that balances the objectives.

On the aggregation of the individual CE, existing frameworks **(Wang and Chin, 2011; Wu et al., 2012a)** pay attention to the reasonable allocation of the weights by limiting the range between self and peer-assessment efficiencies. This condition may indicate consistency from the perspective of the majority opinion. However, considering that many organisations are moving towards systems of evaluation in which also the opinions of minorities are valued **(Park and DeShon, 2010)**, we introduce an aggregation method that rewards contrast. We rely upon the CRiteria Importance Through Inter-criteria Correlation (CRITIC) method **(Diakoulaki et al., 1995)**, an objective method for eliciting weights in multi-criteria problems.With the exception of **He and Ma (2015)**, our paper is the first to apply the CRITIC method in the context of DEA. Its novel function and meaning as deployed in the paper is further described in Section 3.3.2, and differences with the above study are discussed in Section 4.2. Besides, CRITIC would be compatible with the minimax model introduced herein; this is justified by the model's nature to highlight the best behaviour of the worst-performing unit, while the scores of the other better-performing units might decrease.

The remainder of the paper is organised as follows. Section 2 describes the methodological background. In Section 3, we develop the alternative modelling approach for the generalised two-stage DEA structure. Section 4 illustrates the methods with a numerical example. Section 5 presents conclusions and further research.

2 Methodological Background

In the typical input-oriented CCR DEA model **(Charnes et al., 1978)**, each DMU_j $(j = 1, 2, ..., n)$ uses m inputs $(i = 1, 2, ..., m)$ to produce s outputs $(r = 1, 2, ..., s)$. Let X_{ij} be the input value of $i \in M$ for DMU j $\in N$ and Y_{rj} be the output value of $r \in S$ for DMU $j \in N$. These values are known and non-negative. The multiplier-form model that evaluates the efficiency of the target DMU_k is the following:

$$\begin{aligned} E_{kk} = Max \quad & \sum_{r=1}^{s} \mu_{rk} Y_{rk} \\ \text{subject to} \quad & \sum_{i=1}^{m} \nu_{ik} X_{ik} = 1, \\ & \sum_{r=1}^{s} \mu_{rk} Y_{rj} - \sum_{i=1}^{m} \nu_{ik} X_{ij} \leq 0, \forall j, \\ & \mu_{rk}, \nu_{ik} \geq 0, \forall r, i, \end{aligned} \tag{1}$$

where μ_{rk}, ν_{ik} are the rth output and the ith input virtual multipliers, respectively. These are unknown decision variables and they are determined by the linear program. If the set of non-negative optimal multipliers makes the associated objective function equal to 1, then the target DMU_k is called DEA efficient; otherwise, it is called DEA non-efficient.

Two significant challenges of the black-box DEA model, recall the discussion in Section 1, are to acquire a unique ranking order of the existing DMUs (dealing with the lack of discrimination power) and to obtain a more realistic weight scheme **(Örkcü et al., 2019)**. They are inter-related and concurrent **(Li and Reeves, 1999)**.

2.1 Cross-efficiency concept

A commonly used approach to overcome these inabilities is the cross-efficiency (CE) evaluation, proposed by **Sexton et al. (1986)**. Conventional DEA models provide a self-appraisal of the DMUs, using their own optimal weights **(Örkcü et al., 2019)**. Assume that for model **(1)**, μ^*_{rk}, ν^*_{ik} formulate the optimal set of multipliers. Based on this optimal solution, DMU_k is characterised as efficient if and only if $E^*_{kk} = 1$ **(Charnes et al., 1978)**. Model **(1)** needs to be resolved for each DMU (in total n times) to obtain an optimal set of weights for the corresponding DMU. Then by applying the cross-efficiency concept, in which peer-appraisal is the main idea, we evaluate each DMU, considering the weight profiles of all DMUs. In particular, $E_{kj} = \sum_{r=1}^{s} \mu^*_{rk} Y_{rj} / \sum_{i=1}^{m} \nu^*_{ik} X_{ij}$ indicates the individual cross-efficiency of the DMU_j, according to the optimal weighting scheme of DMU_k. A cross-efficiency matrix is a valuable tool for such cases. In this matrix, elements E_{kj} depict the peer-efficiency scores of DMU_j, based on the optimal weights of DMU_k. The diagonal elements of the same matrix indicate the self-efficiency scores of DMU_k. The cross-efficiency score that attributes the final rank of a DMU, is usually estimated by averaging all individual cross-efficiencies of the corresponding DMU which is being evaluated. Thus, $\hat{e}_j = \frac{1}{n} \cdot \sum_{k=1}^{n} E_{kj}$ $(j = 1, 2, ..., n)$ **(Anderson et al., 2002)**.

A key difficulty of the CE evaluation is that the optimal weights obtained by model **(1)** may not be unique, resulting in the non-uniqueness of cross-efficiency scores and rankings of DMUs. To tackle this difficulty, **Doyle and Green (1994)** proposed the use of aggressive and benevolent models, as alternative secondary goals. Model **(2)** is the aggressive. It maximises the performance of the DMU under consideration while minimising the cross-efficiencies of all other DMUs. Model **(3)** is the benevolent that ensures the maximisation of the cross-efficiencies of all other DMUs, whilst maintaining the performance of the target DMU.

$$
\begin{aligned}
Min \quad & \sum_{r=1}^{s} \mu_{rk} \Big(\sum_{j=1, j\neq k}^{n} Y_{rj} \Big) \\
\text{subject to} \quad & \sum_{i=1}^{m} \nu_{ik} \Big(\sum_{j=1, j\neq k}^{n} X_{ij} \Big) = 1, \\
& \sum_{r=1}^{s} \mu_{rk} Y_{rk} - E^{*}_{kk} \sum_{i=1}^{m} \nu_{ik} X_{ik} = 0, \\
& \sum_{r=1}^{s} \mu_{rk} Y_{rj} - \sum_{i=1}^{m} \nu_{ik} X_{ij} \leq 0, \forall j;\ j \neq k, \\
& \mu_{rk}, \nu_{ik} \geq 0, \forall r, i,
\end{aligned}
\tag{2}
$$

$$
Max \quad \sum_{r=1}^{s} \mu_{rk} \Big(\sum_{j=1, j\neq k}^{n} Y_{rj} \Big)
\tag{3}
$$

subject to the same constraints as in model (2).

Troutt (1997) and later **Liang et al. (2008)** developed a novel secondary goal, based on the minimisation of the maximum k-inefficiency (or deviation) score. By identifying an optimal set of multipliers that assigns the maximum efficiency score to the DMU with the worst performance, they achieved the reduction of deviations among all the other DMUs. Hence, they presented the following linear programming model, where $\alpha^{*}_{k} = 1 - E^{*}_{kk}$:

$$
\begin{aligned}
Min \quad & \theta_k \\
\text{subject to} \quad & \sum_{r=1}^{s} \mu_{rk} Y_{rj} - \sum_{i=1}^{m} \nu_{ik} X_{ij} + \alpha_j = 0, \forall j, \\
& \sum_{i=1}^{m} \nu_{ik} X_{ik} = 1, \\
& \sum_{r=1}^{s} \mu_{rk} Y_{rk} = 1 - \alpha^{*}_{k}, \\
& \theta_k - \alpha_j \geq 0, \forall j, \\
& \mu_{rk}, \nu_{ik}, \alpha_j, \theta_k \geq 0, \forall r, i, j.
\end{aligned}
\tag{4}
$$

Model **(4)** corresponds to a cooperative situation towards a single-stage DEA structure. In Section 3.2, it will be amended and customised to the specifications of the generalised two-stage structure to accommodate the purposes of our DEA methodology.

3 Models Development

Yu and Shi (2014) recommended a DEA structure in which each DMU consists of two sub-stages connected in series, as in **Figure 1**. The initial inputs X_{ij} (where $i = 1, 2, ..., m$) entering stage 1 are converted into intermediate products Z_{dj} (where $d = 1, 2, ..., D$). Part of intermediate products $\alpha_{dj} Z_{dj}$ is consumed during stage 2, and the remaining part $(1-\alpha_{dj})Z_{dj}$ is channeled out of the system as final output. α_{dj} is the allocation proportion, dividing this intermediate product into the aforementioned two parts, where $0 \leq \alpha_{dj} \leq 1$. In stage 2, additional inputs X^2_{hj} (where $h = 1, 2, ..., H$) are also supplied from outside. Finally, Y_{rj} (where $r = 1, 2, ..., s$) are the outputs from stage 2 produced for outside.

Note that α_{dj} is pre-specified externally by the decision maker; it is therefore an observed rather than a decision value, that is subjectively designated prior to solving the corresponding mathematical model. This conceptual idea contrasts with the handling of α_{dj} as a variable, according to **Yu and Shi (2014)**. Our decision to illustrate α_{dj} as an observed value determined by the decision maker (externally) and not the model (internally) may represent the reality better, reflecting for example: the market conditions, the contractual requirements, the produced quantity of sub-stage 1, and the alternating requirements and needs of the decision-maker.

To gain a better understanding of the reason we have selected α_{dj} as an observed value, we can refer to a real-life example that clearly describes the two-stage structure **(Figure 1)**. A stock-farmer in a cattle farm (DMU) feeds with corn, wheat, and pasture land (initial inputs in stage 1) dairy cows to produce raw milk (intermediate product at the end of stage 1). The farmer ought to decide how much quantity of the produced milk will be further processed (part of intermediate product as input of stage 2) to get butter, cheese, and yoghurt (final output), and how much quantity will be directly allocated to the outside market (remaining intermediate product as final output). Finally, the fungi for the flash pasteurisation of milk could be an additional (exogenous) input of stage 2. In this example, the decision maker i.e. the stock-farmer freely determines beforehand the way to utilise the produced quantity of milk. Evidently, his decision could be influenced by the laws of supply and demand, the production capacity of the cattle farm, and/or the state of health of the cows.

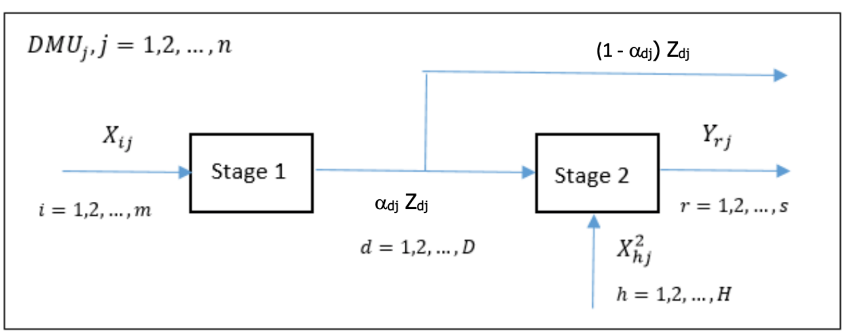

Figure 1: The *generalised* two-stage structure; **Yu and Shi (2014)**.

3.1 Additive efficiency aggregation

The constant-returns-to-scale (CRS) efficiency scores for the target DMU$_k$ can be calculated by the following two CCR models, respective to the first and second stage; they are based upon the CCR model **(Charnes et al., 1978)**:

$$
\begin{aligned}
E_{kk}^{CCR_1} = Max \quad & \frac{\sum_{d=1}^{D} \eta_{dk} Z_{dk}}{\sum_{i=1}^{m} \nu_{ik} X_{ik}} \\
\text{subject to} \quad & \frac{\sum_{d=1}^{D} \eta_{dk} Z_{dj}}{\sum_{i=1}^{m} \nu_{ik} X_{ij}} \leq 1, \forall j, \\
& \eta_{dk}, \nu_{ik} \geq 0, \forall\, d, i.
\end{aligned} \tag{5}
$$

$$
\begin{aligned}
E_{kk}^{CCR_2} = Max \quad & \frac{\sum_{r=1}^{s} \mu_{rk} Y_{rk} \; + \sum_{d=1}^{D} \eta_{dk} (1-\alpha_{dk}) Z_{dk}}{\sum_{h=1}^{H} q_{hk} X_{hk}^{2} \; + \sum_{d=1}^{D} \eta_{dk} \alpha_{dk} Z_{dk}} \\
\text{subject to} \quad & \frac{\sum_{r=1}^{s} \mu_{rk} Y_{rj} \; + \sum_{d=1}^{D} \eta_{dk} (1-\alpha_{dj}) Z_{dj}}{\sum_{h=1}^{H} q_{hk} X_{hj}^{2} \; + \sum_{d=1}^{D} \eta_{dk} \alpha_{dj} Z_{dj}} \leq 1, \forall j, \\
& \eta_{dk}, \mu_{rk}, q_{hk} \geq 0, \forall\, d, r, h.
\end{aligned} \tag{6}
$$

Yu and Shi (2014) do not measure $(1-\alpha)Z$ flows as outputs of the second stage, which merits a comment. This might make sense when part of the outflow of stage 1 is directly forwarded to an outside market without affecting the remaining outflow processed in stage 2. In this way, stage 2 does not need to consider the trade-off and the two sub-stages act rather as being independent. On the other hand, our model **(6)** measures the $(1-\alpha)Z$ flows. This conceptual difference justifies our motivation to examine how the outflows of stage 1 (intermediate products) are split into two distinctive instances which interact with one another. As a further reason of the inclusion of the $(1-\alpha)Z$ flows in our study, we draw attention to the commonly used efficiency aggregation method to build our models. As discussed in **Kao (2017)**, in such a case the efficiency of the system is defined as a function of those of the constituent sub-stages. The intermediate products (αZ flows, $(1-\alpha)Z$ flows) should be initially involved in measuring the efficiency of the corresponding sub-stage and then in calculating the overall efficiency.

The system efficiency of the DMU_k can be computed from the following CCR model **(7)**. Its objective function illustrates the ratio of the aggregate exogenous outputs to that of the aggregate exogenous inputs, considering only the operations of the entire system.

$$
\begin{aligned}
E_{kk}^{CCR} = Max \quad & \frac{\sum_{r=1}^{s} \mu_{rk} Y_{rk} \; + \sum_{d=1}^{D} \eta_{dk} (1-\alpha_{dk}) Z_{dk}}{\sum_{i=1}^{m} \nu_{ik} X_{ik} \; + \sum_{h=1}^{H} q_{hk} X_{hk}^{2}} \\
\text{subject to} \quad & \frac{\sum_{r=1}^{s} \mu_{rk} Y_{rj} \; + \sum_{d=1}^{D} \eta_{dk} (1-\alpha_{dj}) Z_{dj}}{\sum_{i=1}^{m} \nu_{ik} X_{ij} \; + \sum_{h=1}^{H} q_{hk} X_{hj}^{2}} \leq 1, \forall j, \\
& \eta_{dk}, \mu_{rk}, \nu_{ik}, q_{hk} \geq 0, \forall\, d, r, i, h,
\end{aligned} \tag{7}
$$

$\eta_{dk}, \mu_{rk}, \nu_{ik}, q_{hk}$ correspond to the weights associated with intermediate measure d, output r and inputs i and h, for the DMU_k, respectively. Note that the weights (or multipliers) of the intermediate measures are assumed to be the same for both sub-stages **(Kao and Hwang, 2008)**.

Model **(7)** disregards the internal operations of DMUs and treats each DMU as a black box that uses exogenous inputs to produce exogenous outputs. Neglecting the internal operations of DMUs could spur us to results that are not accurate. For instance, while the overall system could be characterised as efficient, one or both of its individual stages may be inefficient. This is one of the main reasons why we need to examine and model the operations of the internal structures for each DMU.

To accommodate the aforementioned issue, we explore the efficiency aggregation method as previously discussed. It is known that it might take either an additive or a multiplicative form depending on the nature of the problem. To **Chen et al. (2009)**, additive efficiency aggregation models are a way of aggregating components in a two-stage structure. This type of aggregation requires the allocation of a relative importance weight to each sub-stage. The weights can be user-specified. They can alternatively be DMU-specific to recognise the strength of each stage as well as the discrepancies between them, and to facilitate the transformation of the non-linear model to a linear one **(Guo et al., 2017)**. As discussed in **Kao (2016)**, the DMU-specific weights will obtain the most favourable efficiency for the system under evaluation. We believe that this might be a reason towards ensuring a fairer and more cooperative environment for the competing DMUs. This approach can also estimate how much more the inputs of the system can be reduced, while ensuring the same level of output production. Finally, it is applicable to both constant and variable returns-to-scale assumptions.

On the other hand, the multiplicative efficiency aggregation method does not require predetermined weights for building the model. Nevertheless, it can put less emphasis on the intermediate products that are being exchanged between the sub-stages of a DMU, whereas a weighted aggregation method does and thus better reflects the level of cooperation between the stages of a DMU. In addition, when it handles a generalised two-stage network structure with exogenous outputs leaving from stage 1 and/or exogenous inputs entering to stage 2, it is extremely nonlinear and cannot be easily converted into a linear model using the Charnes-Cooper transformation. Even the utilisation of a heuristic search method cannot guarantee a global optimal solution **(Chen and Zhu, 2017)**. For the above reasons, this study selects to define the system efficiency as the weighted (arithmetic mean) approach **(Chen et al., 2009)** of its two sub-stage efficiencies.

$$(w_k^1 \frac{\sum_{d=1}^{D} \eta_{dk} Z_{dk}}{\sum_{i=1}^{m} \nu_{ik} X_{ik}} + w_k^2 \frac{\sum_{r=1}^{s} \mu_{rk} Y_{rk} + \sum_{d=1}^{D} \eta_{dk}(1-\alpha_{dk}) Z_{dk}}{\sum_{h=1}^{H} q_{hk} X_{hk}^2 + \sum_{d=1}^{D} \eta_{dk} \alpha_{dk} Z_{dk}}), \tag{8}$$

where w_k^1 and w_k^2 are weights determined by the decision-maker, so that $w_k^1 + w_k^2 = 1$. These weights are not unknown variables, but functions of the optimisation variables. We can, thus, estimate the overall efficiency of the DMU_k by solving model **(9)**.

$$
\begin{aligned}
Max \quad & (w_k^1 \cdot \frac{\sum_{d=1}^{D} \eta_{dk} Z_{dk}}{\sum_{i=1}^{m} \nu_{ik} X_{ik}} + w_k^2 \cdot \frac{\sum_{r=1}^{s} \mu_{rk} Y_{rk} + \sum_{d=1}^{D} \eta_{dk}(1-\alpha_{dk}) Z_{dk}}{\sum_{h=1}^{H} q_{hk} X_{hk}^2 + \sum_{d=1}^{D} \eta_{dk} \alpha_{dk} Z_{dk}}) \\
\text{subject to} \quad & \frac{\sum_{d=1}^{D} \eta_{dk} Z_{dj}}{\sum_{i=1}^{m} \nu_{ik} X_{ij}} \leq 1, \forall j, \\
& \frac{\sum_{r=1}^{s} \mu_{rk} Y_{rj} + \sum_{d=1}^{D} \eta_{dk}(1-\alpha_{dj}) Z_{dj}}{\sum_{h=1}^{H} q_{hk} X_{hj}^2 + \sum_{d=1}^{D} \eta_{dk} \alpha_{dj} Z_{dj}} \leq 1, \forall j, \\
& w_k^1 + w_k^2 = 1, \\
& w_k^1, w_k^2, \eta_{dk}, \mu_{rk}, \nu_{ik}, q_{hk} \geq 0, \forall\, d, r, i, h.
\end{aligned} \tag{9}
$$

Weights w_k^1 and w_k^2 represent the relative importance of the performances of stages 1 and 2 respectively, divided by the overall performance of the evaluated DMU. A larger weight indicates the corresponding stage's stronger effect on the entire performance of the system. To **Chen et al. (2009)** and **Kao (2016)**, the portions of total resources devoted to each stage could correspond to the relative size of a stage. This is also due to the nature of the models which are input-oriented. Therefore, we define:

$$
w_k^1 = \frac{\sum_{i=1}^{m} \nu_{ik} X_{ik}}{\sum_{i=1}^{m} \nu_{ik} X_{ik} + \sum_{h=1}^{H} q_{hk} X_{hk}^2 + \sum_{d=1}^{D} \eta_{dk} \alpha_{dk} Z_{dk}} \tag{10}
$$

and

$$
w_k^2 = \frac{\sum_{h=1}^{H} q_{hk} X_{hk}^2 + \sum_{d=1}^{D} \eta_{dk} \alpha_{dk} Z_{dk}}{\sum_{i=1}^{m} \nu_{ik} X_{ik} + \sum_{h=1}^{H} q_{hk} X_{hk}^2 + \sum_{d=1}^{D} \eta_{dk} \alpha_{dk} Z_{dk}}. \tag{11}
$$

Substituting **(10)** and **(11)** into the objective function of model **(9)**, we obtain the following linear fractional programming model:

$$
\begin{aligned}
E_{kk}^{CCR} = Max \quad & \frac{\sum_{d=1}^{D} \eta_{dk} Z_{dk} + \sum_{r=1}^{s} \mu_{rk} Y_{rk} + \sum_{d=1}^{D} \eta_{dk}(1-\alpha_{dk}) Z_{dk}}{\sum_{i=1}^{m} \nu_{ik} X_{ik} + \sum_{h=1}^{H} q_{hk} X_{hk}^2 + \sum_{d=1}^{D} \eta_{dk} \alpha_{dk} Z_{dk}} \\
\text{subject to} \quad & \frac{\sum_{d=1}^{D} \eta_{dk} Z_{dj}}{\sum_{i=1}^{m} \nu_{ik} X_{ij}} \leq 1, \forall j, \\
& \frac{\sum_{r=1}^{s} \mu_{rk} Y_{rj} + \sum_{d=1}^{D} \eta_{dk}(1-\alpha_{dj}) Z_{dj}}{\sum_{h=1}^{H} q_{hk} X_{hj}^2 + \sum_{d=1}^{D} \eta_{dk} \alpha_{dj} Z_{dj}} \leq 1, \forall j, \\
& \eta_{dk}, \mu_{rk}, \nu_{ik}, q_{hk} \geq 0, \forall\, d, r, i, h.
\end{aligned} \tag{12}
$$

By applying the variable substitution technique in **Charnes** and **Cooper (1962)** and by replacing $\eta_{dk}\alpha_{dk} = \phi_{dk}^1$ and $\eta_{dk}(1-\alpha_{dk}) = \phi_{dk}^2$, we introduce the self-evaluation CCR performance score model **(13)**, which is equivalent to model **(12)**. According to the following (implicitly) linear model, it is possible to measure the performance for each DMU, whose internal structure is illustrated by the two-stage DEA process of **Figure 1**. This relational model estimates the aggregated system efficiency while considering the internal mechanisms of its individual stages.

$$
\begin{aligned}
E_{kk}^{CCR} = Max \quad & \sum_{d=1}^{D} \eta_{dk} Z_{dk} \ + \sum_{r=1}^{s} \mu_{rk} Y_{rk} \ + \sum_{d=1}^{D} \phi_{dk}^{2} Z_{dk} \\
\text{subject to} \quad & \sum_{i=1}^{m} \nu_{ik} X_{ik} \ + \sum_{h=1}^{H} q_{hk} X_{hk}^{2} \ + \sum_{d=1}^{D} \phi_{dk}^{1} Z_{dk} = 1, \\
& \sum_{d=1}^{D} \eta_{dk} Z_{dj} \ - \sum_{i=1}^{m} \nu_{ik} X_{ij} \leq 0, \forall j, \\
& \sum_{r=1}^{s} \mu_{rk} Y_{rj} \ + \sum_{d=1}^{D} \phi_{dk}^{2} Z_{dj} \ - \sum_{h=1}^{H} q_{hk} X_{hj}^{2} \ - \sum_{d=1}^{D} \phi_{dk}^{1} Z_{dj} \leq 0, \forall j, \\
& \eta_{dk}, \mu_{rk}, \nu_{ik}, q_{hk} \geq 0, \ \ \phi_{dk}^{1} + \phi_{dk}^{2} = \eta_{dk}, \ \ \phi_{dk}^{1} \geq 0, \ \ \phi_{dk}^{2} \geq 0, \ \forall \ d, r, i, h.
\end{aligned}
\tag{13}
$$

At the optimality of model **(13)**, the system efficiency is computed as $E_{kk}^{CCR} = (\sum_{d=1}^{D} \eta_{dk}^{*} Z_{dk} \ + \sum_{r=1}^{s} \mu_{rk}^{*} Y_{rk} \ + \sum_{d=1}^{D} \phi_{dk}^{2*} Z_{dk}) / (\sum_{i=1}^{m} \nu_{ik}^{*} X_{ik} \ + \sum_{h=1}^{H} q_{hk}^{*} X_{hk}^{2} \ + \sum_{d=1}^{D} \phi_{dk}^{1*} Z_{dk})$, the efficiency of sub-stage 1 as $E_{kk}^{1} = (\sum_{d=1}^{D} \eta_{dk}^{*} Z_{dk}) / (\sum_{i=1}^{m} \nu_{ik}^{*} X_{ik})$, and the efficiency of sub-stage 2 as $E_{kk}^{2} = (\sum_{r=1}^{s} \mu_{rk}^{*} Y_{rk} + \sum_{d=1}^{D} \phi_{dk}^{2*} Z_{dk}) / (\sum_{h=1}^{H} q_{hk}^{*} X_{hk}^{2} + \sum_{d=1}^{D} \phi_{dk}^{1*} Z_{dk})$.

3.2 Proposed cross-efficiency model

Model **(13)** searches for the optimal most favourable weights $\eta_{dk}, \mu_{rk}, \nu_{ik}, q_{hk}, \phi_{dk}^{1}, \phi_{dk}^{2}$ to yield an optimistic self-efficiency score for DMU_k. However, this DEA flexibility of the DMU_k in choosing its own weights could sometimes lead to an unrealistically high efficiency score of the corresponding DMU. This results in a lack of discrimination power and therefore in unrealistic weight distribution. Besides, the optimal solution for model **(13)** may not be unique, reducing the theoretical value of the potential results **(Mahdiloo et al., 2016)**. A point to focus on in this paper is the best possible treatment of the limited discriminatory power and the unrealistic weight distribution, for the two-stage structure **(Figure 1)**.

To overcome these weaknesses, we apply the cross-efficiency concept in the two-stage DEA structure that we examine. We initially propose an alternative secondary goal model to mainly encounter the shortcoming of the non-unique optimal set of multipliers of model **(13)**. This model contains two objective functions (i.e. criteria); each of them represents one of the two stages of the whole system. These two criteria need to be optimised simultaneously.

To advance our multiple criteria-based secondary goal, we have been influenced by the concept of the "minimisation of the maximum k-inefficiency" **(Troutt, 1997; Liang et al., 2008)**. In the minimax model **(14)**, there are two independent objective functions. The first objective ($\theta1$) represents the situation in which we have to minimise the maximum deviation of stage 1 among all DMUs. The second objective ($\theta2$) illustrates the minimisation of the maximum deviation of stage 2 among all DMUs. There is no preference order between these criteria. Considering the theoretical framework of the concept, this approach might be proved useful in cooperative situations **(Liang et al., 2008)**. In our case this is vital, as the two stages that constitute the entire system should have the same bargaining power and should cooperate in order to maximise the overall efficiency **(Halkos et al., 2014)**.

$$
\begin{aligned}
&Min \quad && \theta 1 \\
&Min \quad && \theta 2 \\
&\text{subject to} \quad && \sum_{i=1}^{m} \nu_{ik} X_{ik} \ + \sum_{h=1}^{H} q_{hk} X_{hk}^2 \ + \sum_{d=1}^{D} \phi_{dk}^1 Z_{dk} = 1, \\
& && \sum_{d=1}^{D} \eta_{dk} Z_{dk} \ + \sum_{r=1}^{s} \mu_{rk} Y_{rk} \ + \sum_{d=1}^{D} \phi_{dk}^2 Z_{dk} = E_{kk}^{CCR^*}, \\
& && \sum_{d=1}^{D} \eta_{dk} Z_{dj} \ - \sum_{i=1}^{m} \nu_{ik} X_{ij} \ + b_j^1 = 0, \forall j, \\
& && \sum_{r=1}^{s} \mu_{rk} Y_{rj} \ + \sum_{d=1}^{D} \phi_{dk}^2 Z_{dj} \ - \sum_{h=1}^{H} q_{hk} X_{hj}^2 \ - \sum_{d=1}^{D} \phi_{dk}^1 Z_{dj} \ + b_j^2 = 0, \forall j, \\
& && \theta_1 \geq b_j^1, \forall j, \\
& && \theta_2 \geq b_j^2, \forall j, \\
& && b_j^1, b_j^2, \eta_{dk}, \mu_{rk}, \nu_{ik}, q_{hk} \geq 0, \forall \, j, d, r, i, h, \\
& && \phi_{dk}^1 + \phi_{dk}^2 = \eta_{dk}, \ \ \phi_{dk}^1 \geq 0, \ \ \phi_{dk}^2 \geq 0, \forall \, d.
\end{aligned} \tag{14}
$$

In the above model, $E_{kk}^{CCR^*}$ denotes the optimal objective function value of model **(13)**. The reason why we are using the restrictions "$\theta_1 \geq b_j^1$" and "$\theta_2 \geq b_j^2$" (where $j = 1, 2, ..., n$) is to set θ_1 as the maximum deviation of stage 1, and θ_2 as the maximum deviation of stage 2. This model results in an optimal set of weights that will highlight the best behaviour of the worst-performing DMU, underpinning the fairness in the decision-making process.

Model **(14)** is a bi-objective programming model that can hardly obtain a global optimal solution. A multi-objective program usually provides a set of non-dominated solutions (see **Li and Reeves, 1999**). The researcher could either apply the objectives interactively (**Mahdiloo et al., 2016**) or identify an alternative way of satisfying the conditions simultaneously. Goal programming has been proposed for optimising all criteria at the same time **(Ghasemi et al., 2014; dos Santos Rubem et al., 2017)**.

We apply the concept of **dos Santos Rubem et al. (2017)** to convert the MOLP model **(14)** into a goal programming model. However, given the utopian values assigned to each of the two objective functions (goals), the model should be aligned more closely to Compromise Programming. Moreover, since $b_j^1, b_j^2 \geq 0, \forall \, j$, it follows that $\theta_1, \theta_2 \geq 0$ and thus there is no need to use the negative deviations, d_1^- and d_2^-, in such a model. Actually $\theta_1 = d_1^+$ and $\theta_2 = d_2^+$. Hence, the model is just formulated as follows:

$$
\begin{aligned}
&Min \quad \theta 1 + \theta 2 \\
&\text{subject to the same constraints as in model (14).}
\end{aligned} \tag{15}
$$

Model **(15)** is the proposed minimax secondary goal model for the two-stage structure **(Figure 1)** in this paper and is run under the CRS assumption. It is seeking a particular solution on the Pareto frontier of model **(14)** i.e. one with equally weighted deviations. This model can significantly reduce the number of zero weights assigned to the known factors and better discriminate the DEA-efficient DMUs.

3.3 Alternative aggregation approach

Recalling the discussion in Section 2.1, we are going to calculate the individual cross-efficiencies, based on the representative optimal weights from model **(15)**. In addition, we will determine the cross-efficiencies to get the final ranks of the considered DMUs, based on the CRITIC method.

3.3.1 Individual & ultimate cross-efficiencies

Like all DEA models for cross-efficiency evaluation, the proposed secondary model **(15)** needs to be solved n times, once for every DMU. There will be n sets of input, intermediate measure and output weights available for cross-efficiency evaluation. According to **Kao and Liu (2019)**, in a series DEA structure as the one we probe in this paper, the discriminatory power is stronger due to the increasing number of restrictions; thus, there are less chances that the optimal set of multipliers derived from the first secondary goal model for each DMU is non-unique. Therefore, we can adopt **Kao and Liu's (2019, p.73)** belief that this optimal set is "representative enough" for our analysis. At the optimality of model **(15)**, for each DMU_j $(j \neq k)$, $E_{kj} = (\sum_{d=1}^{D} \eta_{dk}^{*} Z_{dj} + \sum_{r=1}^{s} \mu_{rk}^{*} Y_{rj} + \sum_{d=1}^{D} \phi_{dk}^{2*} Z_{dj})/(\sum_{i=1}^{m} \nu_{ik}^{*} X_{ij} + \sum_{h=1}^{H} q_{hk}^{*} X_{hj}^{2} + \sum_{d=1}^{D} \phi_{dk}^{1*} Z_{dj})$, $E_{kj}^{1} = (\sum_{d=1}^{D} \eta_{dk}^{*} Z_{dj})/(\sum_{i=1}^{m} \nu_{ik}^{*} X_{ij})$, $E_{kj}^{2} = (\sum_{r=1}^{s} \mu_{rk}^{*} Y_{rj} + \sum_{d=1}^{D} \phi_{dk}^{2*} Z_{dj})/(\sum_{h=1}^{H} q_{hk}^{*} X_{hj}^{2} + \sum_{d=1}^{D} \phi_{dk}^{1*} Z_{dj})$.
These are referred to as the cross-efficiency values of the DMU_j of the overall system, of stage 1 and of stage 2, according to the optimal weight scheme of DMU_k respectively, and reflect the peer-evaluation of DMU_j.

For each DMU_j, the weighted average cross-efficiency score, produced by the weighted cross-efficiency aggregation is the following:

$$\hat{e}_j = \frac{\sum_{k=1}^{n} w_k \cdot E_{kj}}{\sum_{k=1}^{n} w_k}, \quad \hat{e}_j^{1} = \frac{\sum_{k=1}^{n} w_k \cdot E_{kj}^{1}}{\sum_{k=1}^{n} w_k}, \quad \hat{e}_j^{2} = \frac{\sum_{k=1}^{n} w_k \cdot E_{kj}^{2}}{\sum_{k=1}^{n} w_k}. \tag{16}$$

They are called the cross-efficiencies for the overall system, stage 1 and stage 2, respectively. $w_1, ..., w_n$ are the relative importance weights for cross-efficiency aggregation and satisfy the conditions: $w_k \geq 0$ $(k = 1, ..., n)$ and $\sum_{k=1}^{n} w_k = 1$.

3.3.2 CRITIC method in DEA

To estimate the weights in **(16)** and solve the aggregation problem, we apply the CRITIC method, an objective way to determine the relative importance in multi-criteria decision-making (MCDM) situations. This objectivity stems from its formal mathematical procedure and the fact that it is less prone to subjective modifications by a decision-maker. CRITIC considers the evaluation decision-making matrix (in our case the cross-efficiency matrix) to elicit information involved in the evaluation criteria. The elicited information is capable of altering the decision situation and the order of preference. This information delves into two dimensions: the contrast intensity and the conflict among the evaluation criteria **(Diakoulaki et al., 1995)**. Below, we will provide an overview of their method as we would apply it to single-stage DEA structures; we further explain why it is a sensible tool for promoting fairness, and why it is compatible with the proposed minimax secondary model.

For a finite set A with $j = 1, 2, ..., n$ alternatives and $k = 1, 2, ..., n$ evaluation criteria E_k, the multi-criteria decision making problem is as follows: $Max\ \{E_1(\alpha), E_2(\alpha), ..., E_n(\alpha) \mid \alpha \in A\}$. Initially, we obtain the generalised cross-efficiency matrix (Table 2), considering the E_{kj} values for $k, j = 1, 2, ..., n$. See more details of that in Section 2.1.

Table 2: Cross-efficiency matrix; **Doyle and Green (1994)**.

	Target DMU_j			
Evaluator DMU_k	1	2	...	n
1	E_{11}	E_{12}	...	E_{1n}
2	E_{21}	E_{22}	...	E_{2n}
...	...	...	...	...
n	E_{n1}	E_{n2}	...	E_{nn}

We proceed to converting the initial cross-efficiency matrix (Table 2) into a matrix of relative scores (Table 3) with the generic element X_{kj}, where $X_{kj} = (E_k(j) - E_k^{min})/(E_k^{max} - E_k^{min})$. In this mathematical formula, E_k^{max} is equivalent to $max\{E_{k1}, E_{k2}, ..., E_{kn}\}$ and E_k^{min} is equivalent to $min\{E_{k1}, E_{k2}, ..., E_{kn}\}$.

Table 3: Matrix of relative scores.

	Target DMU_j			
Evaluator DMU_k	1	2	...	n
1	X_{11}	X_{12}	...	X_{1n}
2	X_{21}	X_{22}	...	X_{2n}
...	...	...	...	...
n	X_{n1}	X_{n2}	...	X_{nn}

We generate a vector X_k signifying the scores of all n alternatives $X_k = (X_k(1), X_k(2), ..., X_k(n))$. This vector is characterised by the standard deviation σ_k, which quantifies the contrast intensity of criterion k. Define $\sigma_k = \sqrt{\frac{\sum_{j=1}^{n}(X_k(j)-\hat{X_k})^2}{n}}$, where $\hat{X_k} = \sum_{j=1}^{n} X_k(j)/n$. Then, a symmetric matrix of $n \otimes n$ criteria with R_{kj} elements (Spearman rank correlation coefficients) is constructed (Table 4), connecting the rank orders of the elements included in the vector X_k and X_j. Note that, in contrast to the previous two tables, Table 4's columns do not list the 'target' DMUs. Instead, each element R_{kj} is a measure of how the degree by which the viewpoint of DMU k as evaluator corresponds to the viewpoint of DMU j as evaluator.

Table 4: Symmetric matrix.

	Evaluator DMU_j			
Evaluator DMU_k	1	2	...	n
1	R_{11}	R_{12}	...	R_{1n}
2	R_{21}	R_{22}	...	R_{2n}
...	...	...	...	...
n	R_{n1}	R_{n2}	...	R_{nn}

The amount of information C_k emitted by the kth criterion can be determined by multiplying the two measures σ_k (i.e. contrast intensity) and $\sum_{j=1}^{n}(1-R_{kj})$ (i.e. conflict):

$$C_k = \sigma_k \cdot \sum_{j=1}^{n}(1-R_{kj}). \tag{17}$$

The higher the C_k, the more information we receive from criterion k and the higher its relative importance. Thereby, they define the formula for the weight of criterion k as:

$$w_k = \frac{C_k}{\sum_{l=1}^{n} c_l}. \tag{18}$$

The value of the weights w_k $(k = 1, 2, ..., n)$ in formula **(18)**, can be used to determine the cross-efficiency of DMU_j for the overall system ($\hat{e}_j$), the stage 1 ($\hat{e}_j^1$), and the stage 2 ($\hat{e}_j^2$) in **(16)**. CRITIC should, in effect, run three times, based on the investigation of the cross-efficiency matrix of the respective system/stage.

Using the traditional average method, we would assign equal weights ($1/n$) to everyone's opinion, thus conforming to the majority vote. It would also not matter how diversified or not each of these opinions are. CRITIC, however, emphasises the value of those opinions that are more diversified and less mainstream. In particular, criterion k (here, evaluator DMU_k) will receive more weight if it achieves a wider gap between the best and the worst alternative (here, the target DMUs) in the process of evaluation. This explicitly leads to a higher standard deviation (contrast intensity), implying that its opinion is taken more into account. In other words, the opinion of someone who ranks everyone the same is given less importance, which agrees with the widely accepted viewpoint of **Zeleny (1982)**. This may be justified in the context of DEA, or peer evaluation in general, if the lack of discriminatory signals in the evaluation report of one particular evaluator is believed to represent less reliable information. The only way in which such a viewpoint is able to receive importance would be through the number of evaluators sharing this opinion.

The second feature of the CRITIC method, known as the conflict measure, assigns more weight to the criterion (opinion of evaluator DMU) that puts emphasis on the minority opinion with respect to peer evaluation. The less someone corresponds to a mainstream evaluation profile, the more their opinion is opposed to the majority, the higher their conflict score. This indicates that their opinion will be more valued under these circumstances.

One way to give the application of CRITIC to DEA an interpretation is to say that the CRITIC method infuses a flavour of the 'scientific' approach into a 'political' voting system. Politics is usually in compliance with the majority vote, but in matters of science we often value the most transparent and well-documented opinion. The 'conflict' measure of CRITIC is quite in accordance with the latter viewpoint. However, this analogy is certainly not exact since in science it suffices to have one new opinion that is proven to be correct that can overturn all other opinions (the status quo). CRITIC does not go that far as it does still account for everyone's opinion; the ultimate efficiency measure a DMU receives is still a weighted average.

Another, and perhaps more fruitful interpretation we believe, is that the CRITIC method avoids assigning too large a weight to the majority vote which, by definition, excludes the minority opinion. In this way, it does not let the mass influence too much the public opinion, and in addition, promotes diversity and inclusion. This reflects a contemporary understanding of fairness as an accommodative attitude which is inclusive of a broad variety of legitimate opinion rather than simply mirroring the viewpoint of the majority.

Finally, CRITIC could be compatible with the proposed minimax secondary goal model (see Section 3.2), since it rewards contrast intensity. Hence, it is more likely that while the worst performing DMU attempts to assess itself in its best possible light, the efficiency scores of the other better performing DMUs might decrease **(Liang et al., 2008)**. Since this situation increases the contrast intensity, our proposed model seems to be an acceptable option to coexist with the CRITIC method.

4 Numerical Experiments

This section illustrates the use of the mathematical concepts developed/presented in Section 3 to examine the issue of fairness in DEA context. Our study applies the figures drawn from **Yu and Shi (2014)** for the evaluation of the efficiency of 10 generalised two-stage supply chains of different milk and dairy farm communities. The cattle farms compete with each other, aiming to decide on a sensible allocation of the available raw milk produced. The generalised two-stage DEA structure is considered for this example (see Figure 1), with part of intermediate measures as final outputs and additional inputs in the second stage.

The input resources corn (X_1), wheat (X_2), and pasture land (X_3) are the food of dairy cows consumed by stage 1 to produce raw milk. The raw milk illustrates the intermediate product at the end of stage 1. We distinguish the raw milk between high-fat ($3.5 - 4.5\%$) content and low-fat ($\leq 2.5\%$) content. The former represents the intermediate measure Z_1 and the latter the Z_2. The farmer (i.e. the decision maker) in each community needs to pre-specify how much of these quantities will be further processed in stage 2 and how much will be forwarded to the external environment (i.e. the end-market), as final output. α_{dj} is a proportion, freely determined by the decision-maker, that acts as a regulator of the amount of the dth intermediate measure assigned for processing to stage 2. In this example, we can assume that the stock-farmer has set each α_{dj}

equal to 0.7 for simplicity, reflecting market conditions, customer requirements, and updated research surveys; they desire a major proportion of the produced outputs of stage 1 to be further processed in stage 2, while nevertheless channelling a significant quantity as final output. This proportion might consider, for example, the degree to which raw milk contains amino acids, vitamins, minerals, and fatty acids as well as to what extent it is a proper option for those with lactose intolerance, asthma, and allergic conditions. The current observed values of α_{dj} could have been any continuous value between 0 and 1, leading to equally meaningful results. Once the quantity of the respective type of raw milk is processed, the working time for the flash pasteurisation of milk (X_1^2) and the working time for its homogenisation through fine nozzles (X_2^2) will be taken into account. The final (exogenous) outputs will be pasteurised milk (Y_1) and cheese (Y_2). The dataset with the 10 farming communities (DMUs) is summarised in Table 5. For modelling, running, and analysing our data, we have utilised the programming language Python 3.7.6 and in particular the version 2.1 of PuLP as the free linear programming library. The experiment ran on a computer with 16GB RAM.

Table 5: The numerical example of **Yu and Shi (2014)**.

DMUs	X_1	X_2	X_3	Z_1	Z_2	X_1^2	X_2^2	Y_1	Y_2
1	9	50	1	20	10	5	8	100	25
2	10	18	10	10	15	7	10	70	20
3	9	30	3	8	20	2	8	96	30
4	8	25	1	20	20	10	10	80	20
5	10	40	5	15	20	5	15	85	15
6	7	35	2	35	10	5	5	90	35
7	7	30	3	10	25	8	10	100	30
8	12	40	4	20	25	4	8	120	10
9	9	25	2	10	10	5	15	110	15
10	10	50	1	20	15	9	10	80	20

4.1 Findings

We first consider solving the problem of evaluation and ranking with the classic self-evaluation DEA approach. This serves as a benchmark for comparison with our proposed approach. Table 6 exhibits the optimal multipliers from solving the proposed additive self-evaluation two-stage DEA model **(13)**, i.e. the basic model without the further model improvements we have introduced in Section 3.2. There are 35 zero weights in total, assigned to the respective known factors. The existence of a zero weight indicates that the information of the corresponding known factor is not considered. The larger this number of zeros, the more uneven the weight distribution becomes.

Table 7 shows the CCR self-efficiency scores and their corresponding rankings of the 10 cattle farms for the overall system (E_{kk}^{CCR}), the stage 1 (E_{kk}^{1}), and the stage 2 (E_{kk}^{2}), respectively. Recall that the efficiency scores have been calculated via the optimal weights of model **(13)**. DMUs 3,6 and 8 are characterised as DEA-efficient for the overall system, DMUs 1,4,6 and 7 are DEA-efficient for stage 1, and DMUs 3,6,8 and 9 are DEA-efficient for stage 2. Only DMU 6 can be deemed as entirely efficient, since the efficiency of their sub-stages is one.

It is evident from the results in Table 7 that the cattle farms cannot be easily ranked via the self-evaluation method, and from the results in Table 6, that this is also based on many flows receiving zero weights and thus not being accounted for.

Table 6: Optimal multipliers for the proposed self-evaluation model (13).

DMUs	ν_{1k}	ν_{2k}	ν_{3k}	η_{1k}	η_{2k}	q_{1k}	q_{2k}	μ_{1k}	μ_{2k}
1	0	0	0.3030	0.0152	0	0	0.0606	0.0049	0.0022
2	0	0.0337	0	0.0143	0.0279	0	0	0.0001	0.0087
3	0.0000	0	0	0.0000	0	0.1444	0.0889	0	0.0333
4	0.0513	0.0071	0	0.0109	0.0186	0	0	0.0004	0.0048
5	0.0413	0.0040	0	0.0086	0.0130	0.0064	0.0080	0.0022	0
6	0	0.0140	0	0.0140	0	0	0.0332	0	0.0104
7	0.0592	0.0058	0	0.0123	0.0186	0	0	0.0005	0.0047
8	0	0.0000	0	0.0000	0.0000	0.1000	0.0750	0.0083	0
9	0	0.0217	0.0008	0.0017	0.0254	0.0357	0.0058	0.0034	0
10	0	0	0.5919	0.0278	0.0018	0	0	0.0011	0

Table 7: CCR self-efficiencies for the overall system, stage 1, and stage 2, derived via model (13).

DMUs	Overall Efficiency E_{kk}^{CCR}	Rank Overall System	Efficiency Stage1 E_{kk}^{1}	Rank Stage1	Efficiency Stage2 E_{kk}^{2}	Rank Stage2
1	0.936	5	**1**	**1**	0.908	5
2	0.909	6	0.924	7	0.886	7
3	**1**	**1**	0.889	8	**1**	**1**
4	0.894	7	**1**	**1**	0.741	8
5	0.690	10	0.676	9	0.709	9
6	**1**	**1**	**1**	**1**	**1**	**1**
7	0.955	4	**1**	**1**	0.890	6
8	**1**	**1**	0.935	6	**1**	**1**
9	0.728	9	0.499	10	**1**	**1**
10	0.844	8	0.985	5	0.640	10

We now consider the proposed minimax secondary goal model **(15)**. In this manner, we will be able to find flow weights in a cross evaluation approach that exhibit some desirable characteristics. In particular, as discussed in Section 3.2, this model will keep the DMU's optimal overall self-efficiency score unchanged, but seeks to minimise the maximum k-inefficiency for each of the stages across all DMUs. Table 8 lists the optimal weights from solving model **(15)**. The reduction of zero weights compared to the foregoing results of Table 6 is noteworthy. In total, there are now only 19 zero weights (compare with 35 in the previous model), improving the weight distribution and providing more balanced results for the evaluated DMUs. The optimal weights from Table 8 are subsequently used to calculate the elements of the cross-efficiency matrices for the overall system, stage 1, and stage 2, respectively (see Appendix A, Tables A.1.1, A.2.1, and A.3.1). The latter are the decision-making matrices, whose elements (peer-efficiency scores for each DMU) are found according to the discussion in Section 3.3.1.

Table 8: Optimal multipliers for the proposed minimax secondary model (15).

DMUs	ν_{1k}	ν_{2k}	ν_{3k}	η_{1k}	η_{2k}	q_{1k}	q_{2k}	μ_{1k}	μ_{2k}
1	0.0000	0	0.3030	0.0152	0	0	0.0606	0.0049	0.0022
2	0	0.0337	0	0.0143	0.0279	0	0	0.0001	0.0087
3	0.0000	0.0000	0.0000	0	0.0000	0.0882	0.1029	0.0096	0.0027
4	0.0575	0.0048	0.0089	0.0109	0.0186	0	0.0000	0.0004	0.0048
5	0.0413	0.0040	0	0.0086	0.0130	0.0064	0.0080	0.0022	0.0000
6	0.0561	0.0046	0.0041	0.0111	0.0173	0	0.0087	0.0007	0.0058
7	0.0592	0.0058	0.0000	0.0123	0.0186	0	0	0.0005	0.0047
8	0.0000	0.0000	0	0.0000	0.0000	0.0791	0.0854	0.0082	0.0019
9	0.0000	0.0217	0.0008	0.0017	0.0254	0.0357	0.0058	0.0034	0
10	0.0000	0	0.5919	0.0278	0.0018	0	0	0.0011	0

The next step in our proposed approach is to apply the CRITIC method, see Section 3.3.2, to help determine an appropriate weight set for combining the individual cross-efficiency scores into a final cross-efficiency score for each DMU and stage. This technique initially converts the cross-efficiency matrix into a matrix of relative scores for the respective sub-stage, identifying the standard deviation; this indicates the contrast in the viewpoints of the same evaluator DMU_k (see in Appendix A, Tables A.1.2, A.2.2, and A.3.2). It then displays the symmetric matrix for the respective sub-stage, identifying the conflict, the information, and the final weight, for each DMU (see in Appendix A, Tables A.1.3, A.2.3, and A.3.3). Conflict particularly gives voice to the less mainstream opinions of the different evaluators regarding a certain evaluated DMU. An evaluator will be assigned a greater relative importance (final weight) if it provides more valuable information. This information should reward contrast, diversity, and inclusion, in the case of the CRITIC multi-criteria method. As an example, in stage 1, the evaluator (cattle farm) 1 is assigned the highest final weight (0.119) due to its standard deviation (0.442) and conflict (11.024) measures, which are the highest among their peers. Similarly, in stage 2, the evaluator with the highest final weight (0.128) is cattle farm 3.

Recalling that the weights derived by formula **(18)** are used to estimate the final cross-efficiencies in **(16)**, see Sections 3.3.1 and 3.3.2, we display the CRITIC cross-efficiencies for each DMU and stage, in Tables 9-11. In particular, the CRITIC cross-efficiencies ($\hat{e}_j$) with their respective ranks for the overall system are summarised in the fourth and fifth columns of Table 9. The proposed minimax secondary model **(15)** evaluated that cattle farm 6 is the most efficient (0.888) and cattle farm 5 has the worst performance (0.532) compared to others; thus, a unique ranking order is achieved. In addition, it can be statistically concluded that the rankings derived from the CRITIC and the traditional average method (third column of table 8) are not significantly different based on a Spearman rank correlation coefficient test **(Daniel, 1978)**, with $r_s = 0.94$. This is significant at the 0.01 level (two-tailed).

The CRITIC cross-efficiencies ($\hat{e}_j^1$) with their respective ranks for the stage 1 are exhibited in the fourth and fifth columns of Table 10. Model **(15)** deemed cattle farm 4 as the most promising (1.000), attaining a unique ranking order once again. The differences between the ranks of CRITIC and average cross-efficiencies (third column of Table 10) are also statistically insignificant. With respect to the fourth and fifth columns of

Table 11 (CRITIC cross-efficiencies and their corresponding ranks for stage 2), cattle farm 3 is located in the first place, with a perfect efficiency score. The dissimilarities with the average cross-efficiency rankings are also negligible based on the Spearman rank correlation test ($r_s = 0.988$). Note that the average cross-efficiencies have been computed following the same reasoning as in CRITIC cross-efficiencies with the sole exception of the method to aggregate the individual cross-efficiencies (see Section 2.1). Although their difference is negligible, we consider that the averaging method, privileges the majority vote, and downplays minority opinion by failing to fully respect diversity and the principle of inclusion. CRITIC method fills this gap, assigning more weight to "mavericks" and promoting the modern concept of fairness, as discussed in Section 3.3.2.

The CRITIC cross-efficiency scores obtained with our proposed minimax secondary model **(15)** are also compared with the geometric average cross-efficiency scores obtained with **Kao and Liu's (2019)** aggressive-based approach. Note that prior to executing our analysis, we have easily adjusted their model to the specifications of our generalised two-stage DEA structure. The geometric average cross-efficiency scores along with their ranks of the overall system, the stage 1, and the stage 2, are respectively depicted in the sixth and seventh columns of Tables 9, 10, and 11. Correlation analysis suggests that there is a highly strong association between the ranks of these two approaches, as indicated by the correlation values 0.927 (overall system), 0.976 (stage 1), and 0.988 (stage 2), which are significant at the 0.01 level (two-tailed). This can be demonstrated even by the fact that both methods achieve total agreement towards the most desirable unit in all three tables. However, there are a number of points that need to be considered, highlighting the preferability of our method over the other in terms of attaining fairer evaluation results.

Firstly, by solving **Kao and Liu's (2019)** model, we obtained an optimal set of multipliers containing 23 zero weights (as compared with the 19 zero weights of our proposed model). This may indicate a less realistic weight scheme for their method. Secondly, in our minimax model both sub-stages of the generalised two-stage structure have the same bargaining power and improve the overall efficiency. This is conducive to the development of a cooperative situation, where the sub-stages behave altruistically even without having reasons to assume that their cooperation will be returned. This stands in sharp contrast with the aggressive method proposed by **Kao and Liu (2019)**. Although they guaranteed unique cross-efficiencies, they selected a non-cooperative approach, in which DMUs act egoistically with a view to maximising their self-evaluation and downplaying the peer-evaluation. Thirdly, we have managed to acquire a higher absolute cross-efficiency score for each DMU and stage (compared to the respective score in **Kao and Liu's (2019)** results), associated with some performance reward; this is connected with the cooperative role of our model **(15)**.

4.2 Implications

This example has illustrated the approach proposed in this paper, which is a novel combination of the use of an additive self-efficiency aggregation model, a minimax secondary goal model, and the CRITIC method in order to improve fairness and objectivity in a cross evaluation context for a generalised two-stage DEA system. Firstly, a more sensible weight distribution has been obtained via the proposed minimax model **(15)** than the basic self-evaluation model **(13)** and the aggressive-based model of **Kao and Liu (2019)**, highlighting our successful efforts in obtaining more meaningful rankings. Secondly, the minimax model developed has in addition been combined with the CRITIC approach to obtain a greater discrimination power than model **(13)** (see Tables 9-11). Thirdly, on the aggregation of the individual cross-efficiencies, we have compared the traditional average method with the weighted average method, in which the weights are computed via the CRITIC approach. In the former, the opinions of the evaluators are centred around the average (majority) viewpoint. In the latter, more credence and higher inclusion is given to these evaluators that exhibit diversity. These may be desirable characteristics, supporting the more modern mindset of many organisations. Fourthly, it has been proven that our minimax model results in higher absolute efficiency scores (than **Kao and Liu's (2019)** scores) connected with some performance prize; this is due to the cooperative nature of the model.

In addition, it is noteworthy that the final rankings obtained are very similar between the three different methods displayed in Tables 9-11. In practice, however, we think that it is very important for DMUs, when subject to peer evaluation leading ultimately to a ranking, that the methods by which this is achieved are agreeable to modern standards of inclusiveness and diversity and provide an acceptable level of objectivity. We can expect that results are more easily accepted, indeed, if these characteristics are more prominently present in the theoretical foundations of the methods deployed.

As stated in the introduction, the only study having used CRITIC in a DEA context before, seems to be **He and Ma's (2015)**. In that article, CRITIC was used to objectively determine weights used within a DEA collaborative development evaluation model for comparing the internal mechanisms of the regional economy and regional logistics within a 10-year period. Our approach differs in that we use it in the context of peer-evaluation, as an alternative method to address the aggregation problem, in addition to the considering this in the generalised two-stage DEA structure. But more importantly, our study highlights how CRITIC's main components of conflict and contrast intensity can contribute towards a fairer and more diversified cross-efficiency perspective.

As for the possible areas where our study could be applicable, we begin by referring to the manufacturing job shop or to line configurations like clothes manufacturing. In such contexts, the Just-in-Time philosophy takes significantly into account the worker rotations. This practice can eliminate employees' fatigue, encourage their development, and help identifying where they can work best. In this example, it is doable to take day-to-day snapshots (DMUs) of the same factory floor, where the workers are being rotated. In this way, it is possible to measure which of the working stations and settings are (in)efficient and on which days. This will facilitate management towards fairly identifying all those workers that need additional training for certain tasks.

Another promising area could be, for instance, the process of the refinement of the selected cocoa beans into chocolate within a specialized factory. From the first stage, where the cocoa beans are roasted and the cocoa nibs are ground, we mainly obtain cocoa powder. The production manager, in collaboration with the marketing and sales department as well as the outbound logistics manager, will eventually decide on a sensible allocation of the available cocoa powder. On this basis, a proportion of this quantity will be directly forwarded to the outside market for sale, and the remaining will be further blended back with the butter, milk, and liquor in varying quantities, in the second stage, to make different types of chocolate. The main target is to fairly compare the efficiency of several generalised two-stage supply chains of different factory branches or farming communities that make use of cocoa beans from different species of cocoa trees.

As a general ascertainment, it is imperative to improve the processes of efficiency measurement and decision-making under a multi-criteria context and within more advanced network DEA structures; an organisational environment that will promote cooperation, leniency, diversity, and inclusion can result in more effective benchmarking strategies.

Table 9: Average cross-efficiencies, CRITIC cross-efficiencies, Geometric average cross-efficiencies **(Kao and Liu, 2019)**, and their respective ranks for the overall system.

DMUs	Average CE	Ranking	CRITIC CE $\hat{e}_j$	Ranking	Geometric average CE **(Kao and Liu, 2019)**	Ranking
1	0.705	6	0.701	6	0.688	5
2	0.531	9	0.533	9	0.444	10
3	0.760	2	0.747	3	0.699	4
4	0.742	4	0.755	2	0.726	2
5	0.531	10	0.532	10	0.495	9
6	**0.895**	**1**	**0.888**	**1**	**0.878**	**1**
7	0.732	5	0.734	4	0.684	6
8	0.746	3	0.732	5	0.716	3
9	0.567	8	0.567	8	0.566	8
10	0.599	7	0.605	7	0.590	7

Table 10: Average cross-efficiencies, CRITIC cross-efficiencies, Geometric average cross-efficiencies **(Kao and Liu, 2019)**, and their respective ranks for the stage 1.

DMUs	Average CE	Ranking	CRITIC CE $\hat{e}^1_j$	Ranking	Geometric average CE **(Kao and Liu, 2019)**	Ranking
1	0.577	6	0.553	7	0.555	6
2	0.492	9	0.523	9	0.384	10
3	0.575	7	0.596	6	0.509	7
4	**1.000**	**1**	**1.000**	**1**	**1.000**	**1**
5	0.526	8	0.540	8	0.492	8
6	0.828	2	0.814	3	0.889	2
7	0.794	3	0.818	2	0.694	3
8	0.648	4	0.664	4	0.618	5
9	0.412	10	0.421	10	0.411	9
10	0.646	5	0.624	5	0.627	4

Table 11: Average cross-efficiencies, CRITIC cross-efficiencies, Geometric average cross-efficiencies **(Kao and Liu, 2019)**, and their respective ranks for the stage 2.

DMUs	Average CE	Ranking	CRITIC CE $\hat{e}_j^2$	Ranking	Geometric average CE **(Kao and Liu, 2019)**	Ranking
1	0.910	3	0.916	3	0.914	2
2	0.676	7	0.706	7	0.673	7
3	**1.000**	**1**	**1.000**	**1**	**0.997**	**1**
4	0.607	9	0.627	9	0.614	9
5	0.606	10	0.620	10	0.605	10
6	0.924	2	0.921	2	0.912	3
7	0.774	6	0.796	6	0.781	6
8	0.840	4	0.809	5	0.826	4
9	0.818	5	0.837	4	0.815	5
10	0.639	8	0.662	8	0.645	8

5 Conclusions & Future Research

Single-stage and the basic serial two-stage DEA systems have fruitfully used various quantitative methods to attain fairness in the evaluation outcomes. Little work, however, has been done addressing the challenge of attaining fairness in a network with more complex interactions among its internal elements. This paper provides new insight to the generalised two-stage DEA structure of **Yu and Shi (2014)**. We have here proposed a modelling approach for this structure, which promotes fairness among the evaluated DMUs.

In this study, we argue that fairness, or the acceptance of an evaluation and ranking by the different DMUs and their stages, is improved by increasing measures related to the degree of discriminatory power, the weight scheme, and the minority vote. We particularly propose a combination of an additive self-efficiency aggregation model, a multi-objective minimax secondary model, and the CRITIC method in an aim to achieve these aspects of fairness and thus a better degree of cooperation between stages of a DMU and among DMUs. This combination is novel in the DEA literature.

The proposed minimax secondary goal model helps tackle the non-unique optimal multipliers derived from the additive self-evaluation model. The minimax model has the capacity to better discriminate the efficient DMUs than the additive self-evaluation model. In addition, it has significantly reduced the number of the zero weights assigned to the respective known factors by the additive self-evaluation model and the aggressive-based approach of **Kao and Liu (2019)**.

We have shown in this paper that the CRITIC method can be applied in DEA to alternatively address the aggregation problem within the DEA cross-efficiency concept. This approach will objectively determine the weights assigned to individual cross-efficiencies to obtain the final cross-efficiencies. By taking into consideration both the contrast intensity and the conflict measures among the DMUs, it manifests the general message of this paper towards satisfying a more contemporary concept of fairness about diversity and inclusion of minority opinions. Moreover, the proposed minimax model seeks for peer evaluation whereby each peer aims to evaluate the worst of the other players in the best possible light. Its benign and cooperative nature, in conjunction with CRITIC, has the benefit to obtain higher absolute efficiency scores for each DMU and stage than the geometric average efficiencies based on the aggressive method of **Kao and Liu (2019)**. This might be connected with some performance reward, encouraging in a way the DMUs to join the efficiency evaluation and ranking.

In this study, we have proposed an additive self-efficiency aggregation model in the spirit of **Chen et al. (2009)**. This is the basic self-evaluation model without the further improvements introduced in later sections. In such a model, the system efficiency is defined as the weighted arithmetic average of its sub-stages. As for its decomposition weights, **Ang and Chen (2016)** proved that they are non-increasing in the order of sub-stages. Put simply, they highlighted that earlier stages would be assigned higher relative importance, affecting the system's efficiency to a greater extent. Based on that, they also demonstrated that the overall and sub-stages' efficiency scores are prone to the impact of the decomposition weights. We acknowledge this as a limitation of our study, and we believe that a re-definition of the weights, reflecting **Ang and Chen's (2016)** research, could accommodate such an issue. In addition, this study could also focus more on the testing of the proposed models and frameworks with empirical data, that is testing their practical value. It would be desirable, for instance, to evaluate the performance of these methods in one of the potential areas described in Section 4.2, or other (fair-trade) supply chains.

The models in this study were developed under the assumption of the constant returns-to-scale. A direction for future research could be their advancement to variable returns-to-scale input-oriented DEA models. Another potential path could be the intention to tweak the CRITIC method by focusing perhaps on the level of acceptance of the participants on the final evaluation and ranking scheme obtained. To this end, the conflict measure could be adapted, for example, to fine-tune the impact of opinions with large contrast intensity in relation to their distance to majority opinions. Finally, current research studies the evaluation of the performance of DMUs with a generalised two-stage structure, only when the data are positive real numbers, and the DEA models are based on this condition. In particular in the envisaged areas of application such as sustainable supply chains, datasets can be expected to be incomplete or less accurately described. Future research could thus relax this assumption by allowing the data points to be imprecise and lie in an interval, for example. Other cases to be investigated concern missing data or intervals, where some values are more likely to occur over other values. In the latter case, since there is no information of the probability distributions, fuzzy numbers and mathematical operations **(Zimmermann, 2011)** could be used as an ideal alternative option.

Acknowledgments

The work was supported by the Engineering and Physical Sciences Research Council (Industrial CASE award) under grant 5162031105656, and BAE Systems.

References

1. Anderson, T.R., Hollingsworth, K. and Inman, L. (2002). The fixed weighting nature of a cross-evaluation model. *Journal of Productivity Analysis*, 17(3), 249-255.

2. Ang, S. and Chen, C.M. (2016). Pitfalls of decomposition weights in the additive multi-stage DEA model. *Omega*, 58, 139-153.

3. An, Q., Chen, H., Xiong, B., Wu, J. and Liang, L. (2017). Target intermediate products setting in a two-stage system with fairness concern. *Omega*, 73, 49-59.

4. Charnes, A., and Cooper, W. W. (1962). Programming with linear fractional functionals. *Naval Research Logistics Quarterly*, 9(3-4), 181-186.

5. Charnes, A., Cooper, W.W. and Rhodes, E. (1978). Measuring the efficiency of decision-making units. *European Journal of Operational Research*, 2(6), 429-444.

6. Chen, K. and Zhu, J. (2017). Second order cone programming approach to two-stage network data envelopment analysis. *European Journal of Operational Research*, 262(1), 231-238.

7. Chen, Y., Cook, W.D., Li, N. and Zhu, J. (2009). Additive efficiency decomposition in two-stage DEA. *European Journal of Operational Research*, 196(3), 1170-1176.

8. Cook, W.D., Tone, K. and Zhu, J. (2014). Data envelopment analysis: Prior to choosing a model. *Omega*, 44, 1-4.

9. Daniel, W. W. (1978). *Applied nonparametric statistics.* Houghton Mifflin.

10. Diakoulaki, D., Mavrotas, G. and Papayannakis, L. (1995). Determining objective weights in multiple criteria problems: The critic method. *Computers and Operations Research*, 22(7), 763-770.

11. Doyle, J. and Green, R. (1994). Efficiency and cross-efficiency in DEA: Derivations, meanings and uses. *Journal of the Operational Research Society*, 45(5), 567-578.

12. Dyson, R.G. and Thanassoulis, E. (1988). Reducing weight flexibility in data envelopment analysis. *Journal of the Operational Research Society*, 39(6), 563-576.

13. Ghasemi, M.R., Ignatius, J. and Emrouznejad, A. (2014). A bi-objective weighted model for improving the discrimination power in MCDEA. *European Journal of Operational Research*, 233(3), 640-650.

14. Guo, C., Shureshjani, R. A., Foroughi, A. A. and Zhu, J. (2017). Decomposition weights and overall efficiency in two-stage additive network DEA. *European Journal of Operational Research*, 257(3), 896-906.

15. Halkos, G.E., Tzeremes, N.G. and Kourtzidis, S.A. (2014). A unified classification of two-stage DEA models. *Surveys in Operations Research and Management Science*, 19(1), 1-16.

16. Halme, M., Joro, T., Korhonen, P., Salo, S. and Wallenius, J. (1999). A value efficiency approach to incorporating preference information in data envelopment analysis. *Management Science*, 45(1), 103-115.

17. He, Y. and Ma, Z. (2015). Study on CRITIC-DEA for Regional Logistics and Regional Economic Coordinated Development Model and Evaluation—A Case Study of Sichuan Province. *Soft Science*, 03.

18. Jianfeng, M.A. (2015). A two-stage DEA model considering shared inputs and free intermediate measures. *Expert Systems with Applications*, 42(9), 4339-4347.

19. Kao, C. and Hwang, S. N. (2011). Decomposition of technical and scale efficiencies in two-stage production systems. *European Journal of Operational Research*, 211(3), 515-519.

20. Kao, C. and Hwang, S.N. (2008). Efficiency decomposition in two-stage data envelopment analysis: An application to non-life insurance companies in Taiwan. *European Journal of Operational Research*, 185(1), 418-429.

21. Kao, C. (2009). Efficiency decomposition in network data envelopment analysis: A relational model. *European Journal of Operational Research*, 192(3), 949-962.

22. Kao, C. (2016). Efficiency decomposition and aggregation in network data envelopment analysis. *European Journal of Operational Research*, 255(3), 778-786.

23. Kao, C. (2014). Network data envelopment analysis: A review. *European Journal of Operational Research*, 239(1), 1-16.

24. Kao, C. (2017). Network Data Envelopment Analysis: Foundations and Extensions. 2nd ed. Springer.

25. Kao, C. and Liu, S.T. (2019). Cross efficiency measurement and decomposition in two basic network systems. *Omega*, 83, 70-79.

26. Li, F., Zhu, Q., Chen, Z. and Xue, H. (2018). A balanced data envelopment analysis cross-efficiency evaluation approach. *Expert Systems with Applications*, 106, 154-168.

27. Li, X. (2017). A fair evaluation of certain stage in a two-stage structure: revisiting the typical two-stage DEA approaches. *Omega*, 68, 155-167.

28. Li, X.B. and Reeves, G.R. (1999). A multiple criteria approach to data envelopment analysis. *European Journal of Operational Research*, 115(3), 507-517.

29. Liang, L., Wu, J., Cook, W.D. and Zhu, J. (2008). Alternative secondary goals in DEA cross-efficiency evaluation. *International Journal of Production Economics*, 113(2), 1025-1030.

30. Ma, C., Liu, D., Zhou, Z., Zhao, W. and Liu, W. (2014). Game cross efficiency for systems with two-stage structures. *Journal of Applied Mathematics*, 2014, 8.

31. Ma, J., Qi, L. and Deng, L. (2017). Efficiency measurement and decomposition in hybrid two-stage DEA with additional inputs. *Expert Systems with Applications*, 79, 348-357.

32. Mahdiloo, M., Jafarzadeh, A.H., Saen, R.F., Tatham, P. and Fisher, R. (2016). A multiple criteria approach to two-stage data envelopment analysis. *Transportation Research Part D: Transport and Environment*, 46, 317-327.

33. Örkcü, H. H., Özsoy, V. S., Örkcü, M., and Bal, H. (2019). A neutral cross efficiency approach for basic two stage production systems. *Expert Systems with Applications*, 125, 333-344.

34. Park, G. and DeShon, R.P. (2010). A multilevel model of minority opinion expression and team decision-making effectiveness. *Journal of Applied Psychology*, 95(5), 824-833.

35. Roll, Y., Cook, W.D. and Golany, B. (1991). Controlling factor weights in data envelopment analysis. *IIE transactions*, 23(1), 2-9.

36. dos Santos Rubem, A.P., de Mello, J.C.C.S. and Meza, L.A. (2017). A goal programming approach to solve the multiple criteria DEA model. *European Journal of Operational Research*, 260(1), 134-139.

37. Sexton, T.R., Silkman, R.H. and Hogan, A.J. (1986). Data envelopment analysis: Critique and extensions. *New Directions for Program Evaluation*, 1986(32), 73-105.

38. Song, L. and Liu, F. (2018). An improvement in DEA cross-efficiency aggregation based on the Shannon entropy. *International Transactions in Operational Research*, 25(2), 705-714.

39. Troutt, M.D. (1997). Derivation of the maximin efficiency ratio model from the maximum decisional efficiency principle. *Annals of Operations Research*, 73, 323-338.

40. Wang, L., Chu, J. and Wu, J. (2007). Selection of optimum maintenance strategies based on a fuzzy analytic hierarchy process. *International Journal of Production Economics*, 107(1), 151-163.

41. Wang, Y.M. and Chin, K.S. (2010a). A neutral DEA model for cross-efficiency evaluation and its extension. *Expert Systems with Applications*, 37(5), 3666-3675.

42. Wang, Y.M. and Chin, K.S. (2010b). Some alternative models for DEA cross-efficiency evaluation. *International Journal of Production Economics*, 128(1), 332-338.

43. Wang, Y.M. and Chin, K.S. (2011). The use of OWA operator weights for cross-efficiency aggregation. *Omega*, 39(5), 493-503.

44. Wang, Y.M., Chin, K.S. and Luo, Y. (2011). Cross-efficiency evaluation based on ideal and anti-ideal decision-making units. *Expert Systems with Applications*, 38(8), 10312-10319.

45. Wang, Y.M. and Wang, S. (2013). Approaches to determining the relative importance weights for cross-efficiency aggregation in data envelopment analysis. *Journal of the Operational Research Society*, 64(1), 60-69.

46. Wanke, P. and Barros, C. (2014). Two-stage DEA: An application to major Brazilian banks. *Expert Systems with Applications*, 41(5), 2337-2344.

47. Wu, J., Chu, J., Sun, J. and Zhu, Q. (2016a). DEA cross-efficiency evaluation based on Pareto improvement. *European Journal of Operational Research*, 248(2), 571-579.

48. Wu, J., Chu, J., Sun, J., Zhu, Q. and Liang, L. (2016b). Extended secondary goal models for weights selection in DEA cross-efficiency evaluation. *Computers and Industrial Engineering*, 93, 143-151.

49. Wu, J., Sun, J. and Liang, L. (2012a). DEA cross-efficiency aggregation method based upon Shannon entropy. *International Journal of Production Research*, 50(23), 6726-6736.

50. Wu, J., Sun, J. and Liang, L. (2012b). Cross efficiency evaluation method based on weight-balanced data envelopment analysis model. *Computers and Industrial Engineering*, 63(2), 513-519.

51. Wu, J., Sun, J., Liang, L. and Zha, Y. (2011). Determination of weights for ultimate cross efficiency using Shannon entropy. *Expert Systems with Applications*, 38(5), 5162-5165.

52. Yu, Y. and Shi, Q. (2014). Two-stage DEA model with additional input in the second stage and part of intermediate products as final output. *Expert Systems with Applications*, 41(15), 6570-6574.

53. Zeleny, M. (1982). Multiple Criteria Decision Making. McGraw-Hill, New York.

54. Zha, Y. and Liang, L. (2010). Two-stage cooperation model with input freely distributed among the stages. *European Journal of Operational Research*, 205(2), 332-338.

55. Zhou, Z., Sun, L., Yang, W., Liu, W. and Ma, C. (2013). A bargaining game model for efficiency decomposition in the centralized model of two-stage systems. *Computers and Industrial Engineering*, 64(1), 103-108.

56. Zhu, J. (1996). Data envelopment analysis with preference structure. *Journal of the Operational Research Society*, 47(1), 136-150.

57. Zimmermann, H. J. (2011). Fuzzy set theory—and its applications. Springer Science & Business Media.

A Appendix

A.1 Overall system

Table A.1.1: Cross-efficiency Matrix of the overall system for the proposed model (15).

DMUs	1	2	3	4	5	6	7	8	9	10
1	0.936	0.461	0.810	0.654	0.691	0.673	0.647	0.801	0.473	0.859
2	0.155	0.909	0.440	0.615	0.640	0.629	0.666	0.433	0.767	0.077
3	0.467	0.771	1.000	0.772	0.794	0.785	0.772	1.000	0.885	0.224
4	0.738	0.891	0.429	0.894	0.837	0.866	0.891	0.420	0.731	0.850
5	0.287	0.604	0.430	0.677	0.690	0.649	0.689	0.431	0.651	0.208
6	0.941	0.767	1.000	0.955	0.939	1.000	0.964	0.975	0.602	0.738
7	0.462	0.851	0.598	0.955	0.905	0.943	0.955	0.588	0.818	0.263
8	0.524	0.675	1.000	0.706	0.827	0.710	0.715	1.000	0.837	0.326
9	0.472	0.605	0.551	0.581	0.669	0.559	0.588	0.553	0.728	0.362
10	0.738	0.503	0.450	0.661	0.644	0.656	0.653	0.442	0.463	0.844

Table A.1.2: Matrix of relative scores for the overall system for the proposed model (15).

DMUs	1	2	3	4	5	6	7	8	9	10
1	0.993	0.000	0.668	0.194	0.171	0.259	0.158	0.657	0.024	1.000
2	0.000	1.000	0.019	0.091	0.000	0.158	0.208	0.022	0.721	0.000
3	0.397	0.691	1.000	0.511	0.516	0.514	0.490	1.000	1.000	0.189
4	0.742	0.959	0.000	0.835	0.658	0.697	0.807	0.000	0.635	0.989
5	0.168	0.319	0.003	0.255	0.168	0.204	0.268	0.019	0.446	0.167
6	1.000	0.683	1.000	1.000	1.000	1.000	1.000	0.957	0.329	0.845
7	0.391	0.869	0.297	0.998	0.887	0.870	0.976	0.289	0.841	0.238
8	0.469	0.478	1.000	0.333	0.627	0.342	0.338	1.000	0.886	0.319
9	0.403	0.320	0.214	0.000	0.097	0.000	0.000	0.229	0.627	0.364
10	0.742	0.093	0.036	0.213	0.013	0.220	0.173	0.037	0.000	0.980
Std Deviation	0.332	0.354	0.446	0.374	0.370	0.331	0.361	0.436	0.347	0.397

Table A.1.3: Symmetric Matrix for the overall system for the proposed model (15).

DMUs	1	2	3	4	5	6	7	8	9	10
1	0.538	-0.655	0.383	-0.185	-0.163	-0.129	-0.262	0.386	-0.473	0.508
2	-0.536	0.384	-0.360	-0.368	-0.362	-0.353	-0.282	-0.359	0.250	-0.467
3	-0.438	0.238	0.429	-0.219	0.077	-0.211	-0.192	0.449	0.771	-0.608
4	0.109	0.019	-0.745	0.007	-0.290	-0.018	0.026	-0.763	-0.584	0.350
5	-0.160	0.160	-0.621	-0.138	-0.248	-0.183	-0.105	-0.628	-0.118	-0.020
6	0.339	0.107	0.345	0.607	0.541	0.620	0.582	0.335	-0.050	0.247
7	-0.245	0.577	-0.477	0.382	0.258	0.317	0.437	-0.488	0.187	-0.183
8	-0.380	-0.143	0.454	-0.423	-0.124	-0.417	-0.429	0.480	0.562	-0.530
9	-0.103	-0.567	-0.107	-0.861	-0.760	-0.834	-0.877	-0.092	-0.228	-0.040
10	0.610	-0.578	-0.137	-0.128	-0.303	-0.093	-0.189	-0.149	-0.827	0.737
Conflict	10.265	10.459	10.836	11.326	11.375	11.301	11.290	10.829	10.509	10.005
Information	3.412	3.707	4.828	4.236	4.213	3.743	4.081	4.725	3.642	3.967
Final Weight	0.084	0.091	0.119	0.104	0.104	0.092	0.101	0.117	0.090	0.098

A.2 Stage 1

Table A.2.1: Cross-efficiency Matrix of the stage 1 for the proposed model (15).

DMUs	1	2	3	4	5	6	7	8	9	10
1	1.000	0.335	0.319	0.526	0.526	0.535	0.526	0.526	0.266	0.970
2	0.050	0.924	0.412	0.516	0.578	0.541	0.578	0.578	1.000	0.052
3	0.133	0.664	0.889	0.666	0.666	0.663	0.666	0.666	0.800	0.145
4	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
5	0.150	0.572	0.557	0.659	0.676	0.669	0.676	0.676	0.613	0.153
6	0.875	0.660	0.398	0.962	1.000	1.000	1.000	1.000	0.413	0.837
7	0.167	0.830	1.000	1.000	1.000	1.000	1.000	1.000	1.000	0.182
8	0.250	0.728	0.697	0.743	0.755	0.749	0.755	0.935	0.770	0.254
9	0.250	0.500	0.437	0.449	0.456	0.452	0.456	0.456	0.499	0.250
10	1.000	0.417	0.464	0.602	0.596	0.605	0.596	0.596	0.383	0.985

Table A.2.2: Matrix of relative scores for the stage 1 for the proposed model (15).

DMUs	1	2	3	4	5	6	7	8	9	10
1	1.000	0.000	0.000	0.139	0.128	0.150	0.128	0.128	0.000	0.968
2	0.000	0.886	0.137	0.121	0.223	0.162	0.223	0.223	1.000	0.000
3	0.088	0.495	0.837	0.394	0.386	0.385	0.386	0.386	0.728	0.099
4	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
5	0.105	0.357	0.350	0.380	0.404	0.396	0.404	0.404	0.473	0.107
6	0.868	0.488	0.116	0.932	1.000	1.000	1.000	1.000	0.200	0.828
7	0.123	0.744	1.000	1.000	1.000	1.000	1.000	1.000	1.000	0.137
8	0.211	0.592	0.555	0.533	0.550	0.542	0.550	0.880	0.686	0.213
9	0.211	0.248	0.174	0.000	0.000	0.000	0.000	0.000	0.318	0.209
10	1.000	0.124	0.213	0.277	0.257	0.280	0.257	0.257	0.160	0.984
Std Deviation	0.442	0.323	0.383	0.377	0.380	0.381	0.380	0.399	0.379	0.429

Table A.2.3: Symmetric Matrix for the stage 1 for the proposed model (15).

DMUs	1	2	3	4	5	6	7	8	9	10
1	0.705	-0.673	-0.416	-0.253	-0.297	-0.257	-0.297	-0.316	-0.663	0.703
2	-0.565	0.206	-0.272	-0.471	-0.418	-0.457	-0.418	-0.418	0.236	-0.571
3	-0.661	0.318	0.318	-0.098	-0.083	-0.107	-0.083	-0.092	0.410	-0.658
4	0.285	-0.078	-0.224	-0.491	-0.502	-0.502	-0.502	-0.413	-0.082	0.290
5	-0.621	0.568	0.370	0.309	0.344	0.313	0.344	0.375	0.531	-0.619
6	0.411	0.087	0.059	0.581	0.580	0.592	0.580	0.620	-0.065	0.411
7	-0.570	0.617	0.525	0.456	0.482	0.455	0.482	0.509	0.580	-0.566
8	-0.652	0.517	0.307	0.160	0.193	0.161	0.193	0.346	0.513	-0.650
9	-0.063	-0.434	-0.521	-0.850	-0.850	-0.851	-0.850	-0.892	-0.313	-0.069
10	0.706	-0.688	-0.389	-0.252	-0.300	-0.258	-0.300	-0.321	-0.669	0.704
Conflict	11.024	9.559	10.242	10.908	10.853	10.910	10.853	10.601	9.520	11.025
Information	4.869	3.092	3.921	4.115	4.123	4.162	4.123	4.231	3.610	4.727
Final Weight	0.119	0.075	0.096	0.100	0.101	0.102	0.101	0.103	0.088	0.115

A.3 Stage 2

Table A.3.1: Cross-efficiency Matrix of the stage 2 for the proposed model (15).

DMUs	1	2	3	4	5	6	7	8	9	10
1	0.908	1.000	0.810	1.000	1.000	0.969	0.978	0.801	1.000	0.696
2	0.604	0.886	0.440	0.889	0.734	0.802	0.884	0.433	0.610	0.782
3	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
4	0.641	0.735	0.429	0.741	0.679	0.710	0.736	0.420	0.547	0.636
5	0.482	0.684	0.430	0.715	0.709	0.617	0.715	0.431	0.703	0.717
6	1.000	1.000	1.000	0.945	0.868	1.000	0.913	0.975	0.937	0.568
7	0.840	0.886	0.598	0.890	0.812	0.876	0.890	0.588	0.670	0.905
8	1.000	0.571	1.000	0.634	0.936	0.645	0.639	1.000	0.926	0.736
9	0.605	0.903	0.551	1.000	1.000	0.762	1.000	0.553	1.000	1.000
10	0.641	0.795	0.450	0.800	0.718	0.750	0.791	0.442	0.592	0.640

Table A.3.2: Matrix of relative scores for the stage 2 for the proposed model (15).

DMUs	1	2	3	4	5	6	7	8	9	10
1	0.822	1.000	0.668	1.000	1.000	0.920	0.940	0.657	1.000	0.297
2	0.236	0.735	0.019	0.696	0.172	0.483	0.680	0.022	0.137	0.494
3	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
4	0.307	0.382	0.000	0.293	0.000	0.243	0.269	0.000	0.000	0.158
5	0.000	0.263	0.003	0.221	0.096	0.000	0.212	0.019	0.343	0.345
6	1.000	1.000	1.000	0.850	0.590	1.000	0.758	0.957	0.862	0.000
7	0.691	0.735	0.297	0.699	0.415	0.676	0.696	0.289	0.270	0.779
8	1.000	0.000	1.000	0.000	0.801	0.073	0.000	1.000	0.837	0.388
9	0.237	0.775	0.214	1.000	1.000	0.379	1.000	0.229	1.000	1.000
10	0.307	0.523	0.036	0.453	0.124	0.348	0.420	0.037	0.099	0.165
Std Deviation	0.383	0.343	0.446	0.361	0.410	0.372	0.354	0.436	0.419	0.353

Table A.3.3: Symmetric Matrix for the stage 2 for the proposed model (15).

DMUs	1	2	3	4	5	6	7	8	9	10
1	-0.263	0.121	-0.215	0.153	-0.073	-0.023	0.158	-0.219	-0.012	0.142
2	-0.307	0.111	-0.540	-0.030	-0.698	0.048	-0.052	-0.562	-0.768	-0.357
3	0.338	-0.063	0.343	-0.040	0.324	0.068	-0.049	0.344	0.360	-0.022
4	-0.029	0.348	-0.260	0.198	-0.376	0.336	0.160	-0.282	-0.451	-0.417
5	-0.723	-0.125	-0.804	-0.029	-0.444	-0.399	0.000	-0.803	-0.576	0.173
6	0.479	0.310	0.526	0.332	0.516	0.395	0.333	0.527	0.527	0.246
7	-0.145	0.235	-0.380	0.051	-0.573	0.225	0.004	-0.405	-0.615	-0.592
8	0.226	-0.095	0.422	0.056	0.682	-0.061	0.081	0.445	0.704	0.345
9	-0.853	-0.268	-0.911	-0.222	-0.679	-0.528	-0.200	-0.913	-0.719	0.036
10	-0.172	0.254	-0.381	0.119	-0.498	0.205	0.090	-0.403	-0.552	-0.338
Conflict	11.450	9.172	12.201	9.410	11.820	9.733	9.476	12.271	12.103	10.785
Information	4.382	3.143	5.437	3.401	4.848	3.616	3.356	5.354	5.076	3.807
Final Weight	0.103	0.074	0.128	0.080	0.114	0.085	0.079	0.126	0.120	0.090