

2013/4



Imprecise data envelopment analysis
for the two-stage process

Adel Hatami-Marbini, Per J. Agrell
and Nazila Aghayi



CORE

DISCUSSION PAPER

Center for Operations Research
and Econometrics

Voie du Roman Pays, 34
B-1348 Louvain-la-Neuve
Belgium

<http://www.uclouvain.be/core>

CORE DISCUSSION PAPER
2013/4

**Imprecise data envelopment analysis
for the two-stage process**

Adel HATAMI-MARBINI¹, Per J. AGRELL²
and Nazila AGHAYI³

January 2013

Abstract

The aggregate black-box approach of conventional Data Envelopment Analysis (DEA) limits its usefulness in situations where the observation is the result of independent decision making in sub-units (sub-DMUs), sequentially linked through processes or semi-finished products. The situation is commonly found in e.g supply chain management, health care provision and environmental management (waste water treatment). Alternative approaches for sublevel evaluations include two-stage or multi-stage models, where intermediate outputs or inputs are identified to span local production possibility spaces. However, the reliance upon numeric values for such intermediate inputs or outputs adds an additional difficulty that may lower the value of the assessment. In this paper, we present an approach for two-stage evaluation with interval data to resolve this problem. The results show that ignoring the interval quality of the data leads to distorted evaluations, both for the subunit and the system efficiency. The proposed method obtains an efficiency interval consisting in an upper and a lower bound for the system efficiency and the sub-DMU efficiency. In order to link two stages, we consider the interval intermediate measures that are outputs and inputs for the first stage and the second stage, respectively. The derived interval metric, along with its mean, provides a more informative basis for multi-stage evaluation in the presence of imprecise data. The ranks of DMUs and sub-DMUs are obtained based on their interval efficiencies.

Keywords: data envelopment analysis, two-stage process, interval data, interval efficiency, ranking.

JEL Classification: C14, M11, C61

¹ Université catholique de Louvain, CORE and Louvain School of Management, B-1348 Louvain-la-Neuve, Belgium. E-mail: adel.hatamimarbini@uclouvain.be.

² Université catholique de Louvain, CORE and Louvain School of Management, B-1348 Louvain-la-Neuve, Belgium. E-mail: per.agrell@uclouvain.be. This author is also member of ECORE, the association between CORE and ECARES.

³ Department of Mathematics, Science and Research Branch, Islamic Azad University, Tehran, Iran. E-mail: n.aghayi@srbiau.ac.ir

This paper presents research results of the Belgian Program on Interuniversity Poles of Attraction initiated by the Belgian State, Prime Minister's Office, Science Policy Programming. The scientific responsibility is assumed by the author.

1. Introduction

In a competitive environment, it is crucial for organizations to know how efficiently and effectively they are operating compared to similar organizations. For example a university department may wish to compare its performance with the similar departments from other universities, or a bank may want to compare the performance of its different branches throughout the country with competition. In the literature, two approaches are fundamentally advanced for measuring efficiency, namely, parametric and non-parametric frontier approaches. The parametric techniques require a priori assumptions to be made with regard to the production function while in non-parametric techniques no such assumptions are necessary. Data envelopment analysis (DEA) is a powerful non-parametric technique to measure the relative efficiency of decision making units (DMUs) which consume multiple inputs to produce multiple outputs. One of the few fundamental functional assumptions in DEA is the returns-to-scale (RTS) behavior of the production set estimated from the observed data. Charnes et al. (1978) initially introduced DEA under constant returns to scale (CRS) by extending linear programming production economics concepts of empirical efficiency proposed around twenty years earlier by Debreu (1951) and Farrell (1957).

The conventional models often consider the DMUs as black boxes although in most applications DMUs consist of sub-DMUs, internal structures, with relationships among them. That is to say, each DMU is, in practice, considered as a series of sequential activities (sub-DMUs) occurring in various sectors such as hospitals, universities, banks and etc. In the DEA literature, these models in the presence of the inner structure are called network DEA (NDEA) pioneered by Färe and Grosskopf (1996). Recently, NDEA attracts a great deal of more attention from authors and a large number of papers that deal with both theoretical issues and applications have been released. A two-stage process, a special case of Färe and Grosskopf's multi-stage framework, involves only two stages with intermediate measures between them. To the best of our knowledge, network models which are composed of a two-stage structure were initially proposed by Wang et al. (1997) and later developed by Seiford and Zhu (1999).

Based on the conventional view, DEA should strive for certainty in all its manifestations (precision, specificity, sharpness, consistency, etc.) and uncertainty (imprecision, non-specificity, vagueness, inconsistency, etc.) should be ignored in evaluating the performance. According to a more comprehensive perspective, uncertainty is considered necessary to DEA; it is not only an unavoidable plague, but also it has a great utility and relevance. Real-life problems often involve various types of uncertainty for data which may be interval, ordinal, qualitative, or fuzzy. Sengupta (1992) is a seminal paper in this change of paradigm. Accordingly, some researchers have proposed various methods for coping with the imprecise and ambiguous data in DEA (Emrouznejad et al., 2008; Hatami-Marbini et al., 2011a).

In this paper, we propose a new imprecise model for measuring efficiency scores of DMUs with two-stage structure. Indeed, we evaluate each DMU by its constituting activities. We demonstrate that the evaluation of DMUs is not straightforward when the

inputs, outputs and intermediate measure change in intervals. The proposed method calculates the upper and lower bounds for the whole process and for each stage. The final efficiency score for the whole process and two stages will be characterized by an interval bounded by the best lower bound efficiency and the best upper bound efficiency. In both optimistic and pessimistic viewpoints, the efficiency frontier is made by the best condition of all DMUs. In the optimistic viewpoint, a DMU under evaluation is in its best situation whilst in the pessimistic viewpoint, a DMU which is under evaluation put in its worst situation. We take into account the interval intermediate measures which are outputs and inputs for the first stage and the second stage, respectively, to have a linkage between two stages. We apply the upper and lower bounds of the intermediate measure as well as an average of them to obtain a linkage between two stages. We also rank DMUs and sub-DMUs according to their interval efficiencies. A numerical example is presented to demonstrate the applicability of the proposed framework.

This paper is organized as follows: The next section presents a brief literature review on the two-stage DEA models and the imprecise DEA models. Section 3, in terms of mathematics, we first provide an overview of the conventional DEA models, and then we present the efficiency measurement in a two-stage process. Section 4 presents the details of the proposed framework. In section 5 we illustrate a numerical example to exhibit the applicability of the proposed method. We close the paper in section 6 with some conclusions and future research directions.

2. Literature Review

The conventional DEA model treats the efficiency evaluation as a black box and no sub-process or stages are considered, where intermediate measures are produced by some sub-processes and used by other sub-processes. Thus, although the conventional DEA approach is appropriate for determining inefficient DMUs and evaluating the measure of their inefficiencies, it provides little insight into the inefficient sources and the locations where the inefficiency may occur. Moreover, it does not provide process-specific guidance to decision makers to help them improve the efficiencies of the DMUs. Performance evaluation of multi-plant firms in the presence of their internal structures was originated by Färe and Primont (1984) extended in Färe (1991) and Färe and Whittaker (1995). Färe and Grosskopf (1996) is a seminal study of network DEA (NDEA) in the form of a book entitled "*Intertemporal production frontiers with dynamic DEA*" as the first to go over the inner workings of the production process in more detail. According to Färe and Grosskopf (1996), NDEA can be classified into three general classes: (1) A static network model involves a finite set of sub-technologies so as to build a network, (2) A dynamic network model evaluates a sequence of production technologies where one stage (e.g. a time period) effects on later stages, (3) A technology adoption examines production on different processors where inputs are allocated among the processors to determine which technology to adopt. A special case of Färe and Grosskopf's multi-stage framework, a two-stage process, involves only two stages with intermediate measures between them.

2.1. Two-stage DEA models

Although it is evident that either reducing inputs or increasing outputs will improve the inefficiency in DEA, the main interest maybe to decompose the factor of efficiency. To deal with this concern, much effort has been done to decompose the overall efficiency into components so that the sources of inefficiency can be identified. Accordingly, some DEA studies recently focus on the two-stage structure where the first stage uses inputs to generate outputs that then become the inputs to the second stage. The second stage thus utilizes these first stage outputs to produce its own outputs.

Wang et al. (1997) initially proposed a two-stage structure for measuring the performance. They ignored the intermediate measures and obtained an overall efficiency of the whole process with the inputs of the first stage and the outputs of the second stage. Seiford and Zhu (1999) explored a two-stage method to attain the profitability and marketability of the top 55 U.S. commercial banks. Seiford and Zhu (1999) used a standard DEA model separately in each stage for measuring the efficiencies. Sequentially, Zhu (2000) applied the method of Seiford and Zhu (1999) to the Fortune Global 500 companies. As demonstrated in Zhu (2003c) and Chen and Zhu (2004), such approaches that treat the stages in a two-stage process as operating independently of each another does not suitably characterize the performance of the two stages owing to the existence of the serial relationship of the two stages. Chen and Zhu (2004) suggested an alternative DEA model for the two-stage process for variable returns to scale. Their method not only measured the overall efficiency, but also computed the optimal values of the intermediate measures for each DMU. Furthermore, Chen and Zhu (2004)'s model can specify the DEA frontier of two-stage process for projecting the inefficient units onto the efficient frontier.

Liang et al. (2006, 2008) proposed a method for evaluation the efficiency of a supply chain (whole process) and its members (two stages) using the non-cooperative (leader-follower) and cooperative concept in game theory. Under the non-cooperative view, the leader is first evaluated, then the evaluation of the follower is on the basis of leader's efficiency. In the cooperative structure, the overall efficiency that is constructed as an average of the two stages' efficiencies is maximized, and accordingly the efficiencies of the two stages are evaluated.

Chen et al. (2006) evaluated the IT impact on firm performance by means of the two-stage structure. They actually modified the evaluation process of IT investment proposed in (Wang et al., 1997; Chen and Zhu, 2004) by considering shared resources (i.e., fixed assets, IT budget and employees) with the two stages. Kao and Hwang (2008) took account of the series of relationship between the whole process and the two sub-processes in measuring the efficiencies for constant returns to scale (CRS) technology. Under their framework, the overall efficiency can be decomposed into the product of the efficiencies of the two stages. Chen et al. (2009a) explored the connection between two common models proposed by Chen and Zhu (2004) and Kao and Hwang (2008). Chen et al. (2009b) proposed a weighted additive (arithmetic mean) approach to calculate the overall efficiency of the process under the VRS assumption.

Yang et al. (2011) proposed a DEA approach under the CRS assumption to measure the overall efficiency of the entire process using a predefined production possibility set. Their model can be obtained a production frontier for improving the inefficient DMUs and sub-DMUs.

Wang and Chin (2010) argued that a two-stage DEA model with a weighted harmonic mean of the two stages is equivalent to Chen et al. (2009b)'s model. Furthermore, Wang and Chin (2010) extended Kao and Hwang (2008)'s model under the CRS to the VRS technology and also generalized Chen et al. (2009b)'s model. Chen et al. (2010) proposed a method for determining the frontier to the inefficient DMUs based upon the Kao and Hwang (2008)'s model. Zha and Liang (2010) put forward an approach to measure the performance of a two-stage process in the non-cooperative and cooperative views under the game theory framework, where the shared inputs can be allocated between the two stages. They utilized the product of two stages to measure the overall efficiency of each DMU in the cooperative efficiency.

Two-stage models are classified into three approaches: independent, connected, and rational by Kao and Hwang (2010). In the independent approach, the efficiency of the whole system and all sub-DMUs are computed independently. The connected approach obtains the efficiency of the whole system based on the interactions between sub-DMUs. In the rational approach, there are some relations between the efficiencies of systems and sub-DMUs

2.2. Imprecision in DEA models

In traditional DEA, the inputs and outputs are always treated as deterministic values. However vagueness or imprecision always exists in real-world evaluation problems. Imprecise or vague data may be the result of uncertainty, unquantifiable information, incomplete and non-obtainable information, conflicting information, partiality of truth and partiality of possibility, in short; imperfect information (Zadeh 1975; 1976; 1978; 2008). In order to tackle the uncertain benchmarking problems, fuzzy, interval and stochastic approaches, originated by Sengupta (1992), Cooper et al. (1999) and Aigner et al. (1977), respectively, are commonly used to describe the imprecise characteristics. Below, we abstract these imprecise approaches.

Generally, in stochastic programming, the uncertain data is taken into account as random variables and their probability distributions are assumed to be known. In the efficiency analysis framework, particularly, the data collected can be stochastic and therefore the associated efficiency measures should be stochastic. The discussion on the stochastic approach was started by Aigner et al. (1977) and Meeusen and Van den Broeck (1977) which is classified into the parametric approach in which the functional form of the production frontier needs to be determined in advance. This leads to the so-called stochastic frontier analysis (SFA) which can be used to estimate the efficient frontier and efficiency scores. The statistical nature allows for including the stochastic errors in the analysis and the testing of the hypotheses. SFA decomposes the error term into two parts where the first part represents the inefficiency and the second part describes the statistical

noise. Readers refer to Bogetoft and Otto (2010), Cooper et al. (2000) and Bauer (1990) for further details.

In the fuzzy DEA approach, inputs and outputs are viewed as fuzzy sets and their membership functions also need to be known. The fuzzy DEA methods in the literature are categorized into five general groups by Hatami-Marbini et al. (2011a):

- (i) the tolerance approach (e.g. Sengupta, 1992),
- (ii) the α -level based approach (e.g. Kao and Liu, 2000, 2003, 2005, Saati et al., 2002; Hatami-Marbini and Saati, 2009; Hatami-Marbini et al., 2010a; Hatami-Marbini et al. 2012; Saati et al., 2011),
- (iii) the fuzzy ranking approach (e.g. Guo and Tanaka 2008; Leon et al. 2003; Hatami-Marbini et al. 2011b; Hatami-Marbini et al. 2011c),
- (iv) possibility approach (e.g. Lertworasirikul et al., 2003a, 2003b; Khodabakhshi et al. 2010), and
- (v) other developments (e.g. Hougaard 1999, 2005).

In these two kinds of approaches (stochastic and fuzzy), the membership functions and probability distributions have to be known beforehand. In recent years, the interval analysis method was developed to model the uncertainty in the imprecise DEA approach, in which the bounds of the uncertain data are only required, not necessarily knowing the probability distributions or membership functions.

Cooper et al. (1999, 2001a, 2001b) initially proposed the interval approach (so-called imprecise DEA (IDEA)) to study the interval data in DEA. They transformed a nonlinear programming problem into a linear programming problem equivalent through scale transformations and the obtained efficiency scores for each DMU were the deterministic numerical values less than or equal to unity. Consequently, a number of developments and applications have been proposed for solving nonlinear IDEA problem owing to uncertain inputs and outputs (Kim et al., 1999; Park, 2004; Zhu, 2003a, 2003b, 2004). Kim et al. (1999) discussed how to deal with interval data, strong and weak ordinal data, and ratio interval data with an application to telephone offices. Zhu (2003a) has argued that the IDEA method proposed by Cooper et al. (1999, 2001a, 2001b); significantly adds to the complexity of the DEA model because of the great number of scale transformations and variable alternations, and the scale-transformation approach on both the precise and imprecise data including preference and interval data into constraints, leading to a rapid increase in computational burden. Entani et al. (2002) developed a DEA model with interval efficiencies measured from both the optimistic and the pessimistic viewpoints. Their model (Entani et al., 2002) was first developed for crisp data and then Entani et al. (2002) extended their model to interval data and fuzzy data. Despotis and Smirlis (2002) proposed a pair of models for dealing with imprecise data in DEA by transforming a nonlinear DEA model to an LP in order to attain the upper and lower bounds of the efficiency scores for each DMU. The interval efficiency was used to classify the units into three sets: fully efficient (efficient for all cases), efficient and inefficient DMUs. Wang et al. (2005) extended a pair of interval DEA models based on Despotis and

Smirlis (2002)'s study to construct a fixed and unified production frontier for measuring the efficiencies of DMUs. Their models determined the lower and upper bounds of the best relative efficiency of each DMU. Jahanshahloo et al. (2004) developed an analogous FDH model with interval inputs and outputs. Amirteimoori and Kordrostami (2005) extended a DEA method proposed by Zhu (2003a) for dealing with multi-component efficiency measurement with imprecise data which preserved the linearity of DEA model. Haghighat and Khorram (2005) suggested a DEA method with interval data for seeking the maximum and minimum values of efficient DMUs. Kao (2006) proposed a pair of two-level mathematical programming models and transformed them into a pair of ordinary one-level linear programs. Solving the associated pairs of linear programs produced the efficiency intervals of all the DMUs. Smirlis et al. (2006) introduced an approach based on interval DEA that allowed the evaluation of the DMUs with missing values along with the other DMUs with available crisp data. The missing values were replaced by intervals in which the unknown values were likely to belong. The constant bounds of the intervals were estimated by using statistical or experiential techniques. For the units with missing values, the proposed models were able to identify an upper and a lower bound of their efficiency scores. Park (2007) considered a classification same as Despotis and Smirlis (2002) but in a more general structure of imprecise data consisting of any combinations of bounded and ordinal data. Yu (2007) measured efficiency of DMUs in the presence of the interval data using a multiple objective programming with/without an ideal DMU. Moreover, this method improved the discriminating power of the method of Despotis and Smirlis (2002) by deriving a unique set of weights for all DMUs simultaneously. Jahanshahloo et al. (2007) considered the discriminant analysis in DEA with interval data so as to determine the existence or non-existence of an overlap between two groups, implemented with Monte Carlo algorithm. Toloo et al. (2008) proposed an imprecise DEA model for measuring the overall profit efficiency of DMU in which input and output values varied over certain ranges. They obtained upper and lower bounds of the overall profit efficiency for each DMU and with respect to efficiencies bounds, the DMUs were classified into three various groups; efficient, semi-efficient and inefficient classes. Sadjadi and Omrania (2008) proposed a robust DEA model with consideration of uncertainty on output parameters for the performance assessment of electricity distribution companies. Jahanshahloo et al. (2009) propose a generalized model for interval DEA with interval data and they proved that the model can evaluate the efficiencies of several imprecise DEA models in a unified way. Park (2010) investigated the relationships between a pair of primal and dual models based on the duality theory in imprecise DEA. Mostafaei and Saljooghi (2010) proposed a method for dealing with imprecise data in the cost efficiency analysis. They constructed a pair of two-level mathematical programming problems to obtain the lower bound and upper bound of cost efficiency in the presence of interval data. Based on a robust optimization model, Shokouhi et al. (2010) recently proposed an imprecise DEA approach (called robust data envelopment analysis) in which the input and output parameters only vary in ranges. Their method completely covers the method of Despotis and Smirlis (2002). By

applying a Monte-Carlo simulation, Shokouhi et al. (2010) also demonstrated that the classification or ranking based on the upper and lower bounds of efficiency do not necessarily have maximum conformity.

3. Efficiency Measurement

In this section, we first provide an overview of the conventional DEA models when we do not consider internal structures for DMUs. Next, we review DEA models in the two-stage process.

3.1. DEA models

DEA is a nonparametric method of measuring efficiency that uses mathematical programming rather than regression. DEA circumvents the problem of specifying an explicit form of the production function and makes only a minimum number of assumptions to estimate the underlying technology. Farrell (1957) formulated a linear programming (LP) model to measure the technical efficiency of a firm with reference to a benchmark technology characterized by constant returns to scale.

Let us consider n DMUs under evaluation where each DMU $_j$, $j = 1, 2, \dots, n$, use m inputs to produce s output. Each observation is characterized by an input vector $x_j \in \mathfrak{R}^m, (x_{1j}, \dots, x_{mj})$, and an output vector $y_j \in \mathfrak{R}^s, (y_{1j}, \dots, y_{sj})$, consisting of non-negative elements. The technology T or production possibility set (PPS) is defined as $T = \{(x, y) : x \text{ can produce } y\}$.

In DEA, a benchmark technology is constructed from the observed inputs and outputs of the DMUs. For this, we make the following general assumptions about the production technology without specifying any functional form. These are fairly weak assumptions and hold for all technologies represented by a quasi-concave and weakly monotonic production function.

- a) *Envelopment*: $(x, y) \in T$.
- b) *Free disposability*: $(x, y) \in T, y \geq y', x \leq x' \Rightarrow (x', y') \in T$.
- c) *Convexity*:
 $(x, y), (x', y') \in T, (\tilde{x}, \tilde{y}) = \lambda(x, y) + (1 - \lambda)(x', y'), 0 \leq \lambda \leq 1 \Rightarrow (\tilde{x}, \tilde{y}) \in T$.
- d) *z returns to scale*: $(x, y) \in T \Rightarrow (qx, qy) \in T, \forall q \in \Lambda(z)$, where $z = \text{crs, drs or vrs}$,
 $\Lambda(\text{crs}) = \mathfrak{R}_0, \Lambda(\text{drs}) = [0, 1]$ and $\Lambda(\text{vrs}) = \{1\}$, respectively.

Thereby, we can estimate a PPS or the technology satisfying assumptions (a–d) from the observed data for n DMUs without any explicit specification of a production function. It defines as

$$T^C = \{(x, y) : x \in \mathbb{R}_+^m, y \in \mathbb{R}_+^s \mid x \geq \lambda X, y \leq \lambda Y, \lambda \geq 0\}. \quad (1)$$

where the superscript C presents that the technology is characterized by CRS.

To evaluate the input-oriented technical efficiency of any DMU, we examine to what extent it is possible to proportionally reduce its input(s) and still produce the same output(s). This evaluation with one input is quite straightforward. However, in the presence of multiple inputs, a question would be whether reducing one input is more

important than reducing some other input. When market prices of inputs are not available, one way to deal with this problem is to seek *equiproportionate* reduction in all inputs. The input-oriented technical efficiency of DMU_o can be defined as

$$\theta_o = \theta((x_o, y_o); T) = \min \{ \theta \in \mathbb{R}_+ \mid (\theta x_o, y_o) \in T^C \} \quad (2)$$

The CCR model proposed by Charnes et al. (1978) evaluates the efficiency of a specific DMU_o under the constant returns to scale (CRS). The input oriented primal and dual models are given in (5).

Primal CCR model (input-oriented)

$$\theta_o = \min \left\{ \theta \left| \begin{array}{l} \theta x_o \geq \lambda X, \\ y_o \leq \lambda Y, \\ \lambda \geq 0. \end{array} \right. \right\} \quad (3)$$

Dual CCR model (input-oriented)

$$\theta_o = \max \left\{ u^T y_o \left| \begin{array}{l} u^T y_j - v^T x_j \leq 0, \quad \forall j \\ v^T x_o = 1, \\ u, v \geq 0. \end{array} \right. \right\} \quad (4)$$

where u and v in model (4) are the weights assigned to the outputs and the inputs, respectively. In an economic context, u and v of model (4) are the shadow prices associated with the outputs and the inputs, respectively. The shadow prices are used for aggregation as a measure of average productivity of each DMU and vary across DMUs. In such sense, two restrictions are imposed; firstly, all of these shadow prices have to be nonnegative ($u, v \geq 0$), although zero prices are acceptable for individual inputs and outputs. Secondly, the shadow prices must be such that when aggregated using these prices, no DMU's input-output data results in average productivity greater than unity, $(u^T y_j / v^T x_j) \leq 1$. Note also that the constraint $v^T x_o = 1$ in (4) is a given normalized input prices. It is obvious that there are a large number of shadow prices (u, v) satisfying these restrictions.

3.2. DEA in the two-stage process

Fig. 1 shows a two-stage process, where each DMU is composed of two sub-DMUs in series, and intermediate products by the sub-DMU in stage 1 is consumed by the sub-DMU in stage 2.

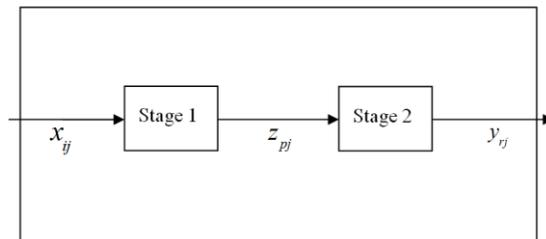


Fig 1. A two-stage process

We assume stage 1 of each DMU_j ($j=1, \dots, n$) uses m inputs (x_j) to produce q outputs (z_j). Then, the outputs of stage 1 (z_j) become the inputs to stage 2 and are called intermediate measures. Therefore, stage 2 of each DMU_j ($j=1, \dots, n$) consists of q inputs and s outputs (y_j). Note that the intermediate measure plays two roles simultaneously in the two-stage process: output and input for the stages 1 and 2, respectively.

Under CRS technology, the overall efficiency scores of the two-stage process and the two individual stages can be obtained as:

$$E_o = \max \left\{ \frac{u^T y_o}{v^T x_o} \left| \begin{array}{l} \frac{u^T y_j}{v^T x_j} \leq 1, \quad \forall j, \\ u, v \geq 0. \end{array} \right. \right\} \quad (5)$$

$$\text{Stage 1) } E_o^1 = \max \left\{ \frac{w^T z_o}{v^T x_o} \left| \begin{array}{l} \frac{w^T z_j}{v^T x_j} \leq 1, \quad \forall j, \\ w, v \geq 0. \end{array} \right. \right\} \quad (6a)$$

$$\text{Stage 2) } E_o^2 = \max \left\{ \frac{u^T y_o}{\bar{w}^T z_o} \left| \begin{array}{l} \frac{u^T y_j}{\bar{w}^T z_j} \leq 1, \quad \forall j, \\ u, \bar{w} \geq 0. \end{array} \right. \right\} \quad (6b)$$

where v and w are the input and output (intermediate measure) weights for stage 1, respectively, while \bar{w} and u are the input (intermediate measure) and output weights for stage 2, respectively.

Kao and Hwang (2008) proposed the overall efficiency as the product of the efficiencies of the two sub-processes when $\bar{w} = w$:

$$E_o = E_o^1 \times E_o^2 \quad (7)$$

Consequently, the overall efficiency is defined as follows:

$$E_o = \max \left\{ \frac{u^T y_o}{v^T x_o} \left| \begin{array}{l} \frac{u^T y_j}{v^T x_j} \leq 1, \quad \forall j, \\ \frac{w^T z_j}{v^T x_j} \leq 1, \quad \forall j, \\ \frac{u^T y_j}{w^T z_j} \leq 1, \quad \forall j, \\ u, v, w \geq 0. \end{array} \right. \right\} \quad (8)$$

Note that the constraint set of (8) is built from models (5), (6a) and (6b) although constraints $u^T y_j / v^T x_j \leq 1$ are redundant in model (8). Model (8) is equivalent to the following linear program

$$E_o = \max \left\{ u^T y_o \left\{ \begin{array}{l} w^T z_j - v^T x_j \leq 0, \quad \forall j, \\ u^T y_j - w^T z_j \leq 1, \quad \forall j, \\ v^T x_o = 1, \\ u, v, w \geq 0. \end{array} \right. \right\} \quad (9)$$

After obtaining the optimal values v^* , u^* and w^* , the efficiencies of stages 1 and 2, respectively, can be calculated as $E_o^1 = w^{*T} z_o / v^{*T} x_o$ and $E_o^2 = u^{*T} y_o / w^{*T} z_o$.

4. Two-stage DEA Model under Uncertainty

All data in the conventional two-stage model are known precisely or given as crisp values. However, in an environment of imprecision, uncertainty, incompleteness of information, crisp data are inadequate or insufficient to model real-life evaluation problems. In some applications, a number of the data sets may be known only within specified bounds. Cooper et al. (1999, 2001a, 2001b) were pioneers to use the interval approach to study the uncertainty in DEA.

Let us consider the uncertainty in model (5) by applying interval inputs and outputs so as to arrive the following model for measuring the overall efficiency of DMU_o :

$$E_o = \max \left\{ \frac{u^T \hat{y}_o}{v^T \hat{x}_o} \left\{ \begin{array}{l} \frac{u^T \hat{y}_j}{v^T \hat{x}_j} \leq 1, \quad \forall j, \\ \frac{w^T \hat{z}_j}{v^T \hat{x}_j} \leq 1, \quad \forall j, \\ \frac{u^T \hat{y}_j}{w^T \hat{z}_j} \leq 1, \quad \forall j, \\ u, v, w \geq 0. \end{array} \right. \right\} \quad (10)$$

“ $\hat{}$ ” represents the uncertainty in the inputs and outputs. We assume that \hat{x}_j , \hat{y}_j and \hat{z}_j are only known to lie within the upper and lower bounds represented by the ranges as $[x_j^l, x_j^u]$, $[y_j^l, y_j^u]$ and $[z_j^l, z_j^u]$, respectively. Note that the lower and upper bounds of the interval inputs and outputs are assumed to be constants and strictly positive. In order to cope with such uncertainty, based on Wang et al. (2005), a new mathematical program is proposed to obtain the lower and upper bound of efficiency of each process and sub-process. These models utilize a common production frontier because the use of same constraint sets results in a more fair and meaningful efficiency assessment of the DMUs and the sub-DMUs. Many present approaches used the different constraint sets for calculating the upper and lower bounds of the efficiency (some references). Wang et al. (2005, p.351) accordingly state: “The main drawback of the use of different constraint sets to measure the efficiencies of DMUs is the lack of comparability among the

efficiencies because different production frontiers were adopted in the process of efficiency measure". In an environment of uncertain data, the interval intermediate measures play a key role of a linkage between two stages for calculating the overall efficiency. Therefore, we have

$$E_o = E_o^1 \times E_o^2 = \frac{w^T \bar{z}_o}{v^T \bar{x}_o} \times \frac{u^T \hat{y}_o}{w^T \bar{z}_o} = \frac{u^T \hat{y}_o}{v^T \bar{x}_o} \quad (11)$$

The upper and lower bounds of the interval overall efficiency of each DMU_o under consideration, denoted by $[E_o^l, E_o^u]$, are obtained from the pessimistic and optimistic viewpoints, respectively, using the following pair of LP models

$$E_o^u = \max \left\{ \begin{array}{l} \frac{u^T y_o^u}{v^T x_o^l} \left| \begin{array}{l} \frac{w^T \bar{z}_j}{v^T x_j^l} \leq 1, \quad \forall j, \\ \frac{u^T y_j^u}{w^T \bar{z}_j} \leq 1, \quad \forall j, \end{array} \right. \\ u, v, w \geq 0. \end{array} \right. \quad (12)$$

$$E_o^l = \max \left\{ \begin{array}{l} \frac{u^T y_o^l}{v^T x_o^u} \left| \begin{array}{l} \frac{w^T \bar{z}_j}{v^T x_j^l} \leq 1, \quad \forall j, \\ \frac{u^T y_j^u}{w^T \bar{z}_j} \leq 1, \quad \forall j, \end{array} \right. \\ u, v, w \geq 0. \end{array} \right. \quad (13)$$

In the optimistic and pessimistic viewpoints, the best situation for all sub-DMUs is considered to build the common production set. In other words, models (12) and (13) consist of the same constraints, however, the best situation for the DMU_o under evaluation is used in the optimistic viewpoint and the worst situation for the DMU_o is used in the pessimistic viewpoint. To achieve this purpose, we take into account three cases for the interval intermediate measures in evaluation analysis: (a) $\bar{z}_j = z_j^u$, (b) $\bar{z}_j = z_j^l$, (c) $\bar{z}_j = 0.5(z_j^l + z_j^u)$. Therefore, we have three pairs of models (12) and (13) with respect to the upper bound, the lower bound and the arithmetic mean of the upper and lower bounds for the intermediate measures. Models (12) and (13) are fractional programs which can be transformed into the following linear program:

$$E_o^u = \max \left\{ u^T y_o^u \left| \begin{array}{l} v^T x_o^l = 1, \\ w^T \bar{z}_j - v^T x_j^l \leq 0, \quad \forall j, \\ u^T y_j^u - w^T \bar{z}_j \leq 0, \quad \forall j, \\ u, v, w \geq 0. \end{array} \right. \right. \quad (14)$$

$$E_o^l = \max \left\{ u^T y_o^l \left| \begin{array}{l} v^T x_o^u = 1, \\ w^T \bar{z}_j - v^T x_j^l \leq 0, \quad \forall j, \\ u^T y_j^u - w^T \bar{z}_j \leq 0, \quad \forall j, \\ u, v, w \geq 0. \end{array} \right. \right\} \quad (15)$$

Once the optimal weights v^* , u^* and w^* of model (14) are determined, the upper efficiencies of stages 1 and 2, respectively, will be calculated as $E_o^{1(u)} = w^{*T} \bar{z}_o / v^{*T} x_o^l$ and $E_o^{2(u)} = u^{*T} y_o^u / w^{*T} \bar{z}_o$, whilst after obtaining the optimal values v^* , u^* and w^* of model (15), the lower efficiencies of stages 1 and 2, respectively, will be computed as: $E_o^{1(l)} = w^{*T} \bar{z}_o / v^{*T} x_o^u$ and $E_o^{2(l)} = u^{*T} y_o^l / w^{*T} \bar{z}_o$. Note that in this paper \bar{z}_j can be the upper bound, the lower bound and the arithmetic mean of the upper and lower bounds for the intermediate measures.

We can advance the following definitions to classify DMUs or sub-DMUs into different classes with respect to the interval efficiency scores:

Definition 1. Assume that E_j^{l*} , $E_j^{1(l)*}$ and $E_j^{2(l)*}$ present the lower efficiency of a DMU, sub-DMUs 1 and 2, respectively. A DMU and the first (second) sub-DMU are called DEA fully efficient if $E_j^{l*} = 1$ and $E_j^{1(l)*} = 1$ ($E_j^{2(l)*} = 1$).

Definition 2. Assume that E_j^{u*} , $E_j^{1(u)*}$ and $E_j^{2(u)*}$ present the upper efficiency of a DMU, sub-DMUs 1 and 2, respectively. A DMU and the first (second) sub-DMU are called DEA efficient if $E_j^{u*} = 1$ and $E_j^{1(u)*} = 1$ ($E_j^{2(u)*} = 1$). A DMU (or a sub-DMU) is called DEA inefficient if its upper efficiency, E_j^{u*} ($E_j^{1(u)*}$ or $E_j^{2(u)*}$), is smaller than unity.

Note that a DMU which is placed in the fully-efficient class remains efficient for all cases (any combination of the input/output data). In addition, as a result of the above definitions and Remark 2, when the lower bound of the overall performance is equivalence to unity, the DEA fully efficiency for the DMU and its sub-DMUs simultaneously occurs in the performance analysis.

The relationship between E_o^u and E_o^l in models (14) and (15) can be defined with Theorem 1.

Theorem 1. If E_o^{u*} and E_o^{l*} are the optimum objective function values of models (14) and (15), respectively, then $E_o^{l*} \leq E_o^{u*}$.

Proof. Let v^* , u^* and w^* be the optimal solution to model (15). We introduce the new variables

$$\beta_o = v^{*T} x_o^l, \quad \tilde{u} = \frac{u^*}{\beta_o}, \quad \tilde{v} = \frac{v^*}{\beta_o} \quad \text{and} \quad \tilde{w} = \frac{w^*}{\beta_o}.$$

Therefore

$$\beta_o = v^{*T} x_o^l \leq v^{*T} x_o^u = 1,$$

$$\tilde{v}^T x_o^l = \frac{v^*}{\beta_o} x_o^l = \frac{1}{\beta_o} v^{*T} x_o^l = 1,$$

Moreover, we have

$$\tilde{u}^T y_j^u - \tilde{w}^T \bar{z}_j = \frac{1}{\beta_o} (u^{*T} y_j^u - w^{*T} \bar{z}_j) \leq 0, \quad \forall j$$

$$\tilde{w}^T \bar{z}_j - \tilde{v}^T x_j^l = \frac{1}{\beta_o} (w^{*T} \bar{z}_j - v^{*T} x_j^l) \leq 0, \quad \forall j$$

It is obvious that all new variables are also constrained to be non-negative as

$$\tilde{u} = \frac{u^*}{\beta_o} \geq 0,$$

$$\tilde{v} = \frac{v^*}{\beta_o} \geq 0,$$

$$\tilde{w} = \frac{w^*}{\beta_o} \geq 0,$$

Thus, \tilde{u} , \tilde{v} and \tilde{w} is a feasible solution to model (14). Therefore, $\tilde{u}^T y_o^u$ is always smaller than or equal to the optimum objective function of (14), $E_o^{u^*}$.

$$E_o^{l^*} = u^{*T} y_o^l \leq u^{*T} y_o^u = (\beta_o \tilde{u})^T y_o^u = \beta_o \tilde{u}^T y_o^u \leq \beta_o E_o^{u^*} \leq E_o^{u^*}.$$

This completes the proof. \square

In addition, the relationship between the lower and upper bounds of the first and second stages can be considered as the following corollaries directly from Theorem 1, stated without their proofs.

Corollary 1. *If $E_o^{1(l)}$ and $E_o^{1(u)}$ are the efficiency scores for the first stage, then $E_o^{1(l)} \leq E_o^{1(u)}$.*

Corollary 2. *If $E_o^{2(l)}$ and $E_o^{2(u)}$ are the efficiency scores for the second stage, then $E_o^{2(l)} \leq E_o^{2(u)}$.*

It is clear that the feasible region of sub-DMUs is greater than the feasible region of the whole process (see formula (7)). As a result, we state the following remark concerning the interval efficiency score of stages and the whole process.

Remark 1. *If $[E_o^{l^*}, E_o^{u^*}]$, $[E_o^{1(l)^*}, E_o^{1(u)^*}]$ and $[E_o^{2(l)^*}, E_o^{2(u)^*}]$ are the interval efficiency scores of the whole process, stage 1 and stage 2, respectively, then $E_o^{l^*} \leq E_o^{u^*} \leq E_o^{1(l)^*} \leq E_o^{1(u)^*}$ and $E_o^{l^*} \leq E_o^{u^*} \leq E_o^{2(l)^*} \leq E_o^{2(u)^*}$.*

Models (14) and (15) may have multiple optimal solutions. Therefore, it is possible that the uniqueness of the decomposition of the overall efficiency does not occur in the

optimistic and pessimistic scenarios. According to Kao and Hwang's (2008) approach, we find a set of multipliers which produces the largest upper and lower efficiency score of the first (or second) stage while maintaining the overall efficiency score at E_o^{u*} and E_o^{l*} calculated from (14) and (15). Hence, for this extension, we give priority to the first stage to obtain its maximum upper and lower efficiency via models (16) and (17) when the overall efficiency is, respectively, equal to E_o^{u*} and E_o^{l*} .

$$E_o^{1(u)} = \max \left\{ \begin{array}{l} \frac{\sum_{p=1}^q w_p \bar{z}_{p0}}{\sum_{i=1}^m v_i x_{i0}^l} \left\{ \begin{array}{l} \frac{\sum_{r=1}^s u_r y_{r0}^u}{\sum_{i=1}^m v_i x_{i0}^l} = E_o^{u*}, \quad \frac{\sum_{p=1}^q w_p \bar{z}_{pj}}{\sum_{i=1}^m v_i x_{ij}^l} \leq 1, \quad \forall j, \\ \frac{\sum_{r=1}^s u_r y_{rj}^u}{\sum_{p=1}^q w_p \bar{z}_{pj}} \leq 1, \quad \forall j, \quad u_r, v_i, w_p \geq 0, \quad \forall r, i, p. \end{array} \right. \end{array} \right\} \quad (16)$$

$$E_o^{1(l)} = \max \left\{ \begin{array}{l} \frac{\sum_{p=1}^q w_p \bar{z}_{p0}}{\sum_{i=1}^m v_i x_{i0}^l} \left\{ \begin{array}{l} \frac{\sum_{r=1}^s u_r y_{r0}^l}{\sum_{i=1}^m v_i x_{i0}^u} = E_o^{l*}, \quad \frac{\sum_{p=1}^q w_p \bar{z}_{pj}}{\sum_{i=1}^m v_i x_{ij}^l} \leq 1, \quad \forall j, \\ \frac{\sum_{r=1}^s u_r y_{rj}^u}{\sum_{p=1}^q w_p \bar{z}_{pj}} \leq 1, \quad \forall j, \quad u_r, v_i, w_p \geq 0, \quad \forall r, i, p. \end{array} \right. \end{array} \right\} \quad (17)$$

or equivalently,

$$E_o^{1(u)} = \max \left\{ \begin{array}{l} \frac{\sum_{p=1}^q w_p \bar{z}_{p0}}{\sum_{i=1}^m v_i x_{i0}^l} \left\{ \begin{array}{l} \sum_{r=1}^s u_r y_{r0}^u = E_o^{u*}, \quad \sum_{i=1}^m v_i x_{i0}^l = 1, \\ \sum_{p=1}^q w_p \bar{z}_{pj} - \sum_{i=1}^m v_i x_{ij}^l \leq 0, \quad \forall j, \\ \sum_{r=1}^s u_r y_{rj}^u - \sum_{p=1}^q w_p \bar{z}_{pj} \leq 0, \quad \forall j, \quad u_r, v_i, w_p \geq 0, \quad \forall r, i, p. \end{array} \right. \end{array} \right\} \quad (18)$$

$$E_o^{1(l)} = \max \left\{ \begin{array}{l} \frac{\sum_{p=1}^q w_p \bar{z}_{p0}}{\sum_{i=1}^m v_i x_{i0}^l} \left\{ \begin{array}{l} \sum_{r=1}^s u_r y_{r0}^l - E_o^{l*} \sum_{i=1}^m v_i x_{i0}^u = 0, \quad \sum_{i=1}^m v_i x_{i0}^l = 1, \\ \sum_{p=1}^q w_p \bar{z}_{pj} - \sum_{i=1}^m v_i x_{ij}^l \leq 0, \quad \forall j, \\ \sum_{r=1}^s u_r y_{rj}^u - \sum_{p=1}^q w_p \bar{z}_{pj} \leq 0, \quad \forall j, \quad u_r, v_i, w_p \geq 0, \quad \forall r, i, p. \end{array} \right. \end{array} \right\} \quad (19)$$

The upper and lower efficiency of the second stage is then calculated as $E_o^{2(u)} = E_o^u / E_o^{1(u)}$ and $E_o^{2(l)} = E_o^l / E_o^{1(l)}$, respectively. Note that we can similarly follow the above models when giving priority to the second stage.

Likewise, we may give priority to the second stage to obtain its maximum upper and lower efficiency while keeping the upper and lower overall efficiency, E_o^{l*} and E_o^{u*} , respectively. Then, similarly the upper and lower efficiency of the first stage can be computed as $E_o^{1(u)} = E_o^u / E_o^{2(u)}$ and $E_o^{1(l)} = E_o^l / E_o^{2(l)}$, respectively.

Corollary 3. Let $[\bar{E}_o^{1(l)*}, \bar{E}_o^{1(u)*}]$ and $[\bar{E}_o^{2(l)*}, \bar{E}_o^{2(u)*}]$ be the interval efficiencies when giving priority to the first stage and let also $[\underline{E}_o^{1(l)*}, \underline{E}_o^{1(u)*}]$ and $[\underline{E}_o^{2(l)*}, \underline{E}_o^{2(u)*}]$ be the interval efficiencies when giving priority to the second stage. Then, $\bar{E}_o^{2(u)*} = \bar{E}_o^{2(l)*}$ and $\bar{E}_o^{1(u)*} = \bar{E}_o^{1(l)*}$.

Remark 2. If $[E_o^{l*}, E_o^{u*}]$, $[\bar{E}_o^{1(l)*}, \bar{E}_o^{1(u)*}]$ and $[\bar{E}_o^{2(l)*}, \bar{E}_o^{2(u)*}]$ are the interval efficiency scores of the whole process, stage 1 and stage 2, respectively, while giving priority to the first stage, then $E_o^{l*} \leq E_o^{u*} \leq \bar{E}_o^{1(l)*} \leq \bar{E}_o^{1(u)*}$.

Remark 3. If $[E_o^{l*}, E_o^{u*}]$, $[\underline{E}_o^{1(l)*}, \underline{E}_o^{1(u)*}]$ and $[\underline{E}_o^{2(l)*}, \underline{E}_o^{2(u)*}]$ are the interval efficiency scores of the whole process, stage 1 and stage 2, respectively, while giving priority to the second stage, then $E_o^{l*} \leq E_o^{u*} \leq \underline{E}_o^{2(l)*} \leq \underline{E}_o^{2(u)*}$.

5. Conclusions and further research

Performance assessment in real applications relies critically upon the minimal imposition of *a priori* assumptions on the structural relationships, the representativeness of the reference set and the adequate consideration of the uncertainty in the underlying data generation process. The two-stage process modeling in DEA is a reflection of the first criterion, since the non-parametric approach removes the otherwise restrictive assumptions regarding the functional form and due to the explicit consideration of the interdependencies of the process through the two-stage model. The second criterion is also addressed with the two-stage modeling, since data in certain applications indeed emanate from decentralized organizations that should be compared relative to both their individual and team performance, just as in the analysis of industrial supply chains. Finally, the third criterion is the most neglected in the current literature and the main contribution of this paper. Departing from the restrictive assumption of access to perfect and deterministic data for all DMUs, our approach allows for impreciseness in the sense of interval data for inputs and outputs. This approach naturally results in the calculation of distance function results in intervals as well, mapping the domain of the production space into the domain of its projections for aggregate and intermediate processes. However, our approach for the two-stage process has the strength of a consistent production possibility set for the two evaluations, with links back to the first and second criteria evoked.

Further research is underway to apply this approach to real two-stage systems with larger datasets and to explore the interpretation of the obtained measures with respect to the individual and system performance in these systems.

References

- Aigner, D.J., Lovell, C.A.K. & Schmidt, P., 1977. Formulation and estimation of stochastic frontier production function models. *Journal of Econometrics* 6, 21–37.
- Amirteimoori, A. & Kordrostami, S., 2005. Multi-component efficiency measurement with imprecise data. *Applied Mathematics and Computation*, 162, 1265–1277.
- Bauer, P.W., 1990. Recent developments in the econometric estimation of frontiers. *Journal of Econometrics*, 46, 39–56.
- Bogetoft, P., & Otto, L., 2010. Benchmarking with DEA, SFA, and R. Springer.
- Charnes, A., Cooper, W.W. & Rhodes, E.L., 1978. Measuring the efficiency of decision making units. *European Journal of Operational Research*, 2(6), 429–444.
- Chen, Y., & Zhu, J. 2004. Measuring information technology's indirect impact on firm performance. *Information Technology and Management*, 5(12), 9–22.
- Chen, Y., Cook, W. D., Li, N. & Zhu, J. 2009b. Additive efficiency decomposition in two-stage DEA. *European Journal of Operational Research*, 196, 1170–1176.
- Chen, Y., Cook, W.D. & Zhu, J., 2010. Deriving the DEA frontier for two-stage processes. *European Journal of Operational Research*, 202, 138–142.
- Chen, Y., Liang, L. & Zhu, J. 2009a. Equivalence in two-stage DEA approaches. *European Journal of Operational Research*, 193, 600–604.
- Chen, Y., Liang, L., Yang, F. & Zhu, J., 2006. Evaluation of information technology investment: a data envelopment analysis approach. *Computers & Operations Research*, 33, 1368–1379.
- Cooper, W.W., Park, K.S. & Yu, G., 2001a. An illustrative application of IDEA (imprecise data envelopment analysis) to a Korean mobile telecommunication company. *Operations Research*, 49, 807–820.
- Cooper, W.W., Park, K.S. & Yu, G., 2001b. IDEA (imprecise data envelopment analysis) with CMDs (column maximum decision making units). *Journal of the Operational Research Society*, 52, 176–181.
- Cooper, W.W., Park, K.S., & Yu, G., 1999. IDEA and AR-IDEA: models for dealing with imprecise data in DEA. *Management Science*, 45, 597–607.
- Cooper, W.W., Seiford, L.M. & Tone, K. 2000. *Data Envelopment Analysis: A Comprehensive Text with Models, Applications, References and DEA-Solver Software*, Boston: Kluwer Academic Publishers.
- Debreu, G. 1951. The coefficient of resource utilization. *Econometrica*, 19(3), 273–292.
- Despotis, D.K. & Smirlis, Y.G., 2002. Data envelopment analysis with imprecise data. *European Journal of Operational Research*, 140, 24–36.
- Emrouznejad, A., Barnett, R. P. & Gabriel T., 2008. Evaluation of research in efficiency and productivity: A survey and analysis of the first 30 years of scholarly literature in DEA. *Socio-Economic Planning Sciences*, 42(3), 151–157.
- Entani, T., Maeda, Y. & Tanaka, H., 2002. Dual models of interval DEA and its extension to interval data. *European Journal of Operational Research*, 136, 32–45.
- Färe, R. & Primont, D., 1984. Efficiency measures for multiplant firms. *Operations Research Letters*, 3, 257–260.
- Färe, R. & Whittaker, G., 1995. An intermediate input model of dairy production using

- complex survey data. *Journal of Agricultural Economics*, 46, 201–213.
- Färe, R., & Grosskopf, S., (1996). *Intertemporal Production Frontiers: With Dynamic DEA*. Kluwer Academic Publishers.
- Färe, R., 1991. Measuring Farrell efficiency for a firm with intermediate inputs, *Academia Economic Papers*, 19 (2), 329–340.
- Farrell, M.J. 1957, The measurement of productive efficiency. *Journal of the Royal Statistical Society*, Series A, 120, 253-281.
- Guo, P. & Tanaka, H., 2001. Fuzzy DEA: a perceptual evaluation method. *Fuzzy Sets and Systems*, 119 (1), 149–160.
- Haghighat, M.S. & Khorram, E., 2005. The maximum and minimum number of efficient units in DEA with interval data. *Applied Mathematics and Computation* 163, 919–930.
- Hatami-Marbini, A. & Saati, S. 2009. Stability of RTS of efficient DMUs in DEA with fuzzy under fuzzy data. *Applied Mathematical Sciences* 3(44):2157–2166.
- Hatami-Marbini, A., Emrouznejad, A. & Tavana, M., 2011a. A taxonomy and review of the fuzzy data envelopment analysis literature: Two decades in the making. *European Journal of Operational Research*, 214, 457–472.
- Hatami-Marbini, A., Saati, S. & Tavana, M. 2010a. An ideal-seeking fuzzy data envelopment analysis framework. *Applied Soft Computing*, 10 (4):1062–1070.
- Hatami-Marbini, A., Saati, S. & Tavana, M., 2011c. Data envelopment analysis with fuzzy parameters: an interactive approach. *International Journal of Operations Research and Information Systems*, 2(3), 39-53.
- Hatami-Marbini, A., Tavana, M. & Ebrahimi, A., 2011b. A fully fuzzified data envelopment analysis model. *International Journal of Information and Decision Sciences* 3 (3), 252–264.
- Hatami-Marbini, A., Tavana, M., Emrouznejad, A. & Saati, S., 2012. Efficiency measurement in fuzzy additive data envelopment analysis. *International Journal of Industrial and Systems Engineering*, 10(1), 1-20.
- Hougaard, J. L. 1999. Fuzzy scores of technical efficiency. *European Journal of Operational Research* 115(3), 529-541.
- Hougaard, J.L. 2005. A simple approximation of productivity scores of fuzzy production plans. *Fuzzy Sets and Systems*, 152(3), 455-465.
- Jahanshahloo, G.R., Hosseinzadeh Lotfi, F., Rezaei Balf, F. & Zhiani Rezaei, H., 2007. Discriminant analysis of interval data using Monte Carlo method in assessment of overlap. *Applied Mathematics and Computation*, 191, 521–532.
- Jahanshahloo, G.R., Hosseinzadeh Lotfi, F., Rostamy Malkhalifeh, M. & Ahadzadeh Namin, M., 2009. A generalized model for data envelopment analysis with interval data. *Applied Mathematical Modelling*, 33, 3237–3244.
- Jahanshahloo, G.R., Matin, R.K. & Vencheh, A.H., 2004. On FDH efficiency analysis with interval data. *Applied Mathematics and Computation*, 159, 47–55.
- Kao, C. & Hwang, S.-N. 2008. Efficiency decomposition in two-stage data envelopment analysis: An application to non-life insurance companies in Taiwan. *European Journal of Operational Research*, 185, 418–429.
- Kao, C. & Hwang, S.N., 2010. Efficiency measurement for network systems: IT impact on firm performance. *Decision Support Systems*, 48, 437-446.
- Kao, C. & Liu, S.T. 2000. Fuzzy efficiency measures in data envelopment analysis. *Fuzzy Sets and Systems*, 113(3), 427-37.

- Kao, C. & Liu, S.T. 2003. A mathematical programming approach to fuzzy efficiency ranking. *International Journal of Production Economics*, 86(2), 145-154.
- Kao, C. & Liu, S.T. 2005. Data envelopment analysis with imprecise data: An application of Taiwan machinery firms. *International Journal of Uncertainty, Fuzziness and Knowledge-Based Systems*, 13(2), 225-240.
- Kao, C., 2006. Interval efficiency measures in data envelopment analysis with imprecise data. *European Journal of Operational Research*, 174, 1087-1099.
- Khodabakhshi, M., Gholami, Y. & Kheirollahi, H., 2010. An additive model approach for estimating returns to scale in imprecise data envelopment analysis. *Applied Mathematical Modelling*, 34 (5), 1247-1257.
- Kim, S.H., Park, C.K. & Park, K.S., 1999. An application of data envelopment analysis in telephone offices evaluation with partial data. *Computers & Operations Research*, 26, 59-72.
- Leon, T., Liern, V., Ruiz, J.L. & Sirvent, I., 2003. A fuzzy mathematical programming approach to the assessment of efficiency with DEA models. *Fuzzy Sets and Systems* 139 (2), 407-419.
- Lertworasirikul, S., Fang S.C., Joines, J.A. & Nuttle, H.L.W. 2003a. Fuzzy data envelopment analysis (DEA): A possibility approach. *Fuzzy Sets and Systems* 139(2), 379-394.
- Lertworasirikul, S., Fang, S.C., Nuttle, H.L.W. & Joines J.A. 2003b. Fuzzy BCC model for data envelopment analysis. *Fuzzy Optimization and Decision Making*, 2(4), 337-358.
- Liang, L., Cook, W. D. & Zhu, J., 2008. DEA models for two-stage processes: Game approach and efficiency decomposition. *Naval Research Logistics*, 55, 643-653.
- Liang, L., Yang, F., Cook, W. D. & Zhu, J. 2006 DEA models for supply chain efficiency evaluation. *Annals of Operations Research*, 145, 35-49.
- Meeusen, W. & Van den Broeck, J. 1977. Efficiency estimation from Cobb-Douglas production functions with composite error. *International Economic Review*, 18, 435-444.
- Mostafae, A. & Saljooghi, F.H., 2010. Cost efficiency measures in data envelopment analysis with data uncertainty. *European Journal of Operational Research*, 202, 595-603.
- Park, K.S., 2004. Simplification of the transformations and redundancy of assurance regions in IDEA (imprecise DEA). *Journal of the Operational Research Society* 55, 1363-1366.
- Park, K.S., 2007. Efficiency bounds and efficiency classifications in DEA with imprecise data. *Journal of the Operational Research Society*, 58, 533-540.
- Park, K.S., 2010. Duality, efficiency computations and interpretations in imprecise DEA. *European Journal of Operational Research*, 200, 289-296.
- Saati, S., Hatami-Marbini, A. & Tavana, M., 2011. A data envelopment analysis model with discretionary and non-discretionary factors in fuzzy environments. *International Journal of Productivity and Quality Management*, 8 (1), 45-63.
- Saati, S., Memariani, A. & Jahanshahloo, G.R., 2002. Efficiency analysis and ranking of DMUs with fuzzy data. *Fuzzy Optimization and Decision Making* 1, 255-267.
- Sadjadi, S. J. & Omrani, H., 2008. Data envelopment analysis with uncertain data: An application for Iranian electricity distribution companies. *Energy Policy*, 36, 4247-4254.

- Seiford, L. & Zhu, J. 1999. Profitability and marketability of the top 55 U.S. commercial banks. *Management Science*, 45(9), 1270–1288.
- Sengupta, J. K. 1992. A fuzzy systems approach in data envelopment analysis. *Computers and Mathematics with Applications*, 24(8-9), 259-266.
- Shokouhi, A.H., Hatami-Marbini, A., Tavana, M., & Saati, S., 2010. A robust optimization approach for imprecise data envelopment analysis. *Computers and Industrial Engineering*, 59, 387–397.
- Smirlis, Y.G., Maragos, E.K. & Despotis, D.K., 2006. Data envelopment analysis with missing values: An interval DEA approach. *Applied Mathematics and Computation*, 177(1), 1-10.
- Toloo, M., Aghayi, N. & Rostamy-malkhalifeh, M., 2008,. Measuring overall profit efficiency with interval data. *Applied Mathematics and Computation*, 201, 640–649.
- Wang C.H., Gopal, R. & Zionts, S., 1997. Use of data envelopment analysis in assessing information technology impact on firm performance. *Annals of Operations Research*, 73, 191–213.
- Wang, Y.M. & Chin, K.S., 2010. Some alternative DEA models for two-stage process. *Expert Systems with Applications*, 37 (12), 8799-8808.
- Wang, Y.M., Greatbanks, R. & Yang, J. B., 2005. Interval efficiency assessment using data envelopment analysis. *Fuzzy Sets and Systems*, 153, 347–370.
- Yang, F., Wu, D., Liang, L., Bi, G., & Wu, D.D., 2011. Supply chain DEA: production possibility set and performance evaluation model. *Annals of Operations Research*, 185, 195–211.
- Yu, J.R., 2007. IDEA by multiple objective programming. *International Journal of Innovative Computing, Information and Control*, 3(1), 21-29.
- Zadeh, L.A., 1975. Fuzzy logic and approximate reasoning. *Synthese*, 30, 407–428.
- Zadeh, L.A., 1976. A fuzzy-algorithmic approach to the definition of complex or imprecise concepts. *International Journal of Man–Machine Studies*, 8, 249–291.
- Zadeh, L.A., 1978. Fuzzy sets as a basis for a theory of possibility. *Fuzzy Sets and Systems*, 1, 3–28.
- Zadeh, L.A., 2008. Is there a need for fuzzy logic?. *Information Sciences*, 178, 2751–2779.
- Zha, Y. & Liang, L., 2010. Two-stage cooperation model with input freely distributed among the stage. *European Journal of Operational Research*, 205, 332–338
- Zhu, J. 2003c. *Quantitative Models for Performance Evaluation and Benchmarking: Data Envelopment Analysis with Spreadsheets*. Boston: Kluwer Academic Publishers.
- Zhu, J., 2000. Multi-factor performance measure model with an application to Fortune 500 companies. *European Journal of Operational Research*, 123 (1), 105–124.
- Zhu, J., 2003a. Imprecise data envelopment analysis (IDEA): A review and improvement with an application. *European Journal of Operational Research*, 144, 513–529.
- Zhu, J., 2003b. Efficiency evaluation with strong ordinal input and output measures. *European Journal of Operational Research*, 146, 477–485.
- Zhu, J., 2004. Imprecise DEA via standard linear DEA models with a revisit to a Korean mobile telecommunication company. *Operations Research*, 52, 323–329.

Recent titles

CORE Discussion Papers

- 2012/29. Pierre DEHEZ. Incomplete-markets economies: The seminar work of Diamond, Drèze and Radner.
- 2012/30. Helmuth CREMER, Pierre PESTIEAU and Grégory PONTHIÈRE. The economics of long-term care: a survey.
- 2012/31. Mathieu VAN VYVE, Laurence A. WOLSEY and Hande YAMAN. Relaxations for two-level multi-item lot-sizing problem.
- 2012/32. Jean-François MERTENS and Anna RUBINCHIK. Equilibria in an overlapping generations model with transfer policies and exogenous growth.
- 2012/33. Jean-François MERTENS and Anna RUBINCHIK. Pareto optimality of the golden rule equilibrium in an overlapping generations model with production and transfers.
- 2012/34. Hélène LATZER and Florian MAYNERIS. Income distribution and vertical comparative advantage. Theory and evidence.
- 2012/35. Pascal MOSSAY and Takatoshi TABUCHI. Preferential trade agreements harm third countries.
- 2012/36. Aitor CALO-BLANCO. The compensation problem with fresh starts.
- 2012/37. Jean-François CARPANTIER and Arnaud DUFAYS. Commodities volatility and the theory of storage.
- 2012/38. Jean-François CARPANTIER and Christelle SAPATA. Unfair inequalities in France: A regional comparison.
- 2012/39. Shui-Ki WAN, Shin-Huei WANG and Chi-Keung WOO. Total tourist arrival forecast: aggregation vs. disaggregation.
- 2012/40. Thierry BRECHET, Yann MENIERE and Pierre M. PICARD. The clean development mechanism in a global carbon market.
- 2012/41. Nikolaos GEORGANTZIS, Rafael MONER-COLONQUES, Vicente ORTS and José J. SEMPERE-MONERRIS. Theoretical and experimental insights on firms' internationalization decisions under uncertainty.
- 2012/42. Berardino CESI and Dimitri PAOLINI. Peer group and distance: when widening university participation is better.
- 2012/43. Arnaud DUFAYS. Infinite-state Markov-switching for dynamic volatility and correlation models.
- 2012/44. Miguel JARA, Dimitri PAOLINI and J.D. TENA. Management efficiency in football: An empirical analysis of two extreme cases.
- 2012/45. Guillaume ROELS, Philippe CHEVALIER and Ying WEI. United we stand? Coordinating capacity investment and allocation in joint ventures.
- 2012/46. Mikel BEDAYO, Ana MAULEON and Vincent VANNETELBOSCH. Bargaining and delay in trading networks.
- 2012/47. Pierre M. PICARD and Ridwan D. RUSLI. State owned firms: Private debt, cost revelation and welfare.
- 2012/48. Shin-Huei WANG, Luc BAUWENS and Cheng HSIAO. Forecasting long memory processes subject to structural breaks.
- 2012/49. Thierry BRECHET, Carmen CAMACHO and Vladimir M. VELIOV. Adaptive model-predictive climate policies in a multi-country setting.
- 2012/50. Vladyslav NORA and Hiroshi UNO. Saddle functions and robust sets of equilibria.
- 2012/51. Thomas BAUDIN, David DE LA CROIX and Paula GOBBI. DINKs, DEWkS & Co. Marriage, fertility and childlessness in the United States.
- 2012/52. David DE LA CROIX and Omar LICANDRO. The longevity of famous people from Hammurabi to Einstein.
- 2012/53. Marijn VERSCHELDE, Jean HINDRIKS, Glenn RAYP and Koen SCHOORS. School staff autonomy and educational performance: within school type evidence.
- 2012/54. Thierry BRECHET and Susana PERALTA. Markets for tradable emission permits with fiscal competition.

Recent titles

CORE Discussion Papers - continued

- 2012/55 Sudipto BHATTACHARYA, Claude D'ASPREMONT, Sergei GURIEV, Debapriya SEN and Yair TAUMAN. Cooperation in R&D: patenting, licensing and contracting.
- 2012/56 Guillaume WUNSCH, Michel MOUCHART and Federica RUSSO. Functions and mechanisms in structural-modelling explanations.
- 2012/57 Simone MORICONI, Pierre M. PICARD and Skerdilajda ZANAJ. Commodity taxation and regulatory competition.
- 2012/58 Yurii NESTEROV and Arkadi NEMIROVSKI. Finding the stationary states of Markov chains by iterative methods.
- 2012/59 Tanguy ISAAC and Paolo PIACQUADIO. Equity and efficiency in an overlapping generation model.
- 2012/60 Luc BAUWENS, Giuseppe STORTI and Francesco VIOLANTE. Dynamic conditional correlation models for realized covariance matrices.
- 2012/61 Mikhail ESKAKOV and Alexey ISKAKOV. Equilibrium in secure strategies.
- 2012/62 Francis BLOCH and Axel GAUTIER. Strategic bypass deterrence.
- 2012/63 Olivier DURAND-LASSERVE, Axel PIERRU and Yves SMEERS. Sensitivity of policy simulation to benchmark scenarios in CGE models: illustration with carbon leakage.
- 2013/1 Pierre PESTIEAU and Maria RACIONERO. Harsh occupations, health status and social security.
- 2013/2 Thierry BRECHET and Henry TULKENS. Climate policies: a burden or a gain?
- 2013/3 Per J. AGRELL, Mehdi FARSI, Massimo FILIPPINI and Martin KOLLER. Unobserved heterogeneous effects in the cost efficiency analysis of electricity distribution systems.
- 2013/4 Adel HATAMI-MARBINI, Per J. AGRELL and Nazila AGHAYI. Imprecise data envelopment analysis for the two-stage process.

Books

- G. DURANTON, Ph. MARTIN, Th. MAYER and F. MAYNERIS (2010), *The economics of clusters – Lessons from the French experience*. Oxford University Press.
- J. HINDRIKS and I. VAN DE CLOOT (2011), *Notre pension en heritage*. Itinera Institute.
- M. FLEURBAEY and F. MANIQUET (2011), *A theory of fairness and social welfare*. Cambridge University Press.
- V. GINSBURGH and S. WEBER (2011), *How many languages make sense? The economics of linguistic diversity*. Princeton University Press.
- I. THOMAS, D. VANNESTE and X. QUERRIAU (2011), *Atlas de Belgique – Tome 4 Habitat*. Academia Press.
- W. GAERTNER and E. SCHOKKAERT (2012), *Empirical social choice*. Cambridge University Press.
- L. BAUWENS, Ch. HAFNER and S. LAURENT (2012), *Handbook of volatility models and their applications*. Wiley.
- J-C. PRAGER and J. THISSE (2012), *Economic geography and the unequal development of regions*. Routledge.
- M. FLEURBAEY and F. MANIQUET (2012), *Equality of opportunity: the economics of responsibility*. World Scientific.
- J. HINDRIKS (2012), *Gestion publique*. De Boeck.

CORE Lecture Series

- R. AMIR (2002), Supermodularity and complementarity in economics.
- R. WEISMANTEL (2006), Lectures on mixed nonlinear programming.
- A. SHAPIRO (2010), Stochastic programming: modeling and theory.