The composition and the weak-link approaches to Network Data Envelopment Analysis

A dissertation submitted in partial fulfillment of the requirements
for the degree of

Doctor of Philosophy

in the Department of Informatics
in the School of Information and Communication Technologies
at the University of Piraeus

by

Koronakos Grigorios

Under the Supervision of
Professor
Despotis Dimitrios

Department of Informatics
School of Information and Communication Technologies
University of Piraeus
March 2017

The composition and the weak-link approaches to Network Data Envelopment Analysis

A dissertation submitted in partial fulfillment of the requirements
for the degree of

Doctor of Philosophy

in the Department of Informatics
in the School of Information and Communication Technologies
at the University of Piraeus
by

Koronakos Grigorios

Advisory Committee:

Despotis Dimitrios, Professor, University of Piraeus

Apostolou Dimitrios Associate Professor, University of Piraeus

Metaxiotis Konstantinos, Associate Professor, University of Piraeus

Approved by

Date:……………………………………………………………..

…………………
Despotis Dimitrios
Professor
University of Piraeus

…………………
Apostolou Dimitrios
Associate Professor
University of Piraeus

…………………
Metaxiotis Konstantinos
Associate Professor
University of Piraeus

…………………
Chondrokoukis Gregory
Professor
University of Piraeus

…………………
Chrissikopoulos Vassilios
Professor
Ionian University

…………………
Georgiakodis Fotis
Professor Emeritus
University of Piraeus

…………………
Tsihrintzis George
Professor
University of Piraeus

Περίληψη

Η συστηματική αποτίμηση της αποδοτικότητας ενός οργανισμού και η οριοθέτηση επιτεύξιμων στόχων αποτελούν συμπληρωματικές θεμελιώδεις πτυχές για την εύρυθμη λειτουργία του και τη βιωσιμότητά του. Συνεπώς, είναι απαραίτητη η υιοθέτηση τεχνικών αξιολόγησης που λαμβάνουν υπόψη όλους τους παράγοντες από το περιβάλλον λειτουργίας του οργανισμού, ώστε να εντοπίζουν τις μη αποδοτικές παραγωγικές διαδικασίες και να προτείνουν επαρκείς τρόπους για την βελτίωσή τους. Μια τέτοια τεχνική είναι η Περιβάλλουσα Ανάλυση Δεδομένων – ΠΑΔ (Data Envelopment Analysis - DEA), η οποία αποτελεί πλέον τη δημοφιλέστερη μη παραμετρική τεχνική για την αποτίμηση της αποδοτικότητας ομοειδών μονάδων ενός συστήματος (μονάδες απόφασης) επί τη βάσει πολλαπλών εισροών και πολλαπλών εκροών.

Οι κλασσικές μεθοδολογίες της ΠΑΔ θεωρούν τις μονάδες απόφασης ως «μαύρα κουτιά» (black boxes) που χρησιμοποιούν εισροές για την παραγωγή εκροών, αγνοώντας την εσωτερική τους δομή. Αυτό έχει ως συνέπεια τα κλασσικά μοντέλα της ΠΑΔ να μπορούν μερικώς μόνο να ανταποκριθούν στην αποτίμηση της αποδοτικότητας όταν η εσωτερική δομή είναι γνωστή και κρίσιμη για τη λειτουργία της μονάδας. Το κενό αυτό έρχεται να καλύψει η Περιβάλλουσα Ανάλυση Πολυσταδιακών Διεργασιών (Network DEA), η οποία αποτελεί πρόσφατη επέκταση της κλασσικής ΠΑΔ. Πρόκειται για μια μεθοδολογία που μπορεί να εφαρμοστεί για την αξιολόγηση μονάδων απόφασης, οι οποίες απαρτίζονται από πολλά μέλη (γνωστά και ως διεργασίες, υπο-διαδικασίες, υπο-μονάδες παραγωγής ή στάδια), διότι λαμβάνει υπόψη την εσωτερική τους δομή και τις σχέσεις αλληλεπίδρασης που τη συνοδεύουν. Η κάθε μονάδα απόφασης (σύστημα) λειτουργεί ως ένα δίκτυο από διατεταγμένες διεργασίες οι οποίες συνδέονται και αλληλεπιδρούν μέσω εσωτερικών ροών υποπροϊόντων (ενδιάμεσων μεγεθών), τα οποία έχουν διττό ρόλο διότι αποτελούν ταυτόχρονα εκροές μιας υπο-διαδικασίας και εισροές μιας άλλης. Ενδεικτικό παράδειγμα αποτελεί η εφοδιαστική αλυσίδα που περιέχει πολλά μέλη και η εύρυθμη ή η μη αποδοτική λειτουργία του κάθε μέλους αντανακλάται στη συνολική λειτουργία της. Συνεπώς, η εκτίμηση της συνολικής αποδοτικότητας της εφοδιαστικής αλυσίδας (συστήματος) πρέπει να γίνεται συντονισμένα λαμβάνοντας υπόψη τις αποδοτικότητες των μελών της.

Στην παρούσα διδακτορική διατριβή διεξάγουμε μια λεπτομερή ανασκόπηση των μεθόδων που έχουν προταθεί στη βιβλιογραφία στο πλαίσιο της Περιβάλλουσας Ανάλυσης Πολυσταδιακών Διεργασιών. Μελετούμε τις ιδιότητες-χαρακτηριστικά τους, τις τεχνικές επίλυσης που χρησιμοποιούν καθώς και τις ομοιότητες και διαφορές τους ώστε να τις κατατάξουμε σε κατηγορίες. Αναδεικνύουμε τα μειονεκτήματα των πιο διαδεδομένων προσεγγίσεων, τα οποία αφορούν την κλίμακα αποδόσεων και την αδυναμία να αποδώσουν επαρκή πληροφορία ώστε να καταστήσουν αποδοτικές τις μη αποδοτικές μονάδες. Επίσης, οι προσεγγίσεις που προτείνονται στη βιβλιογραφία δεν διασφαλίζουν τη μοναδικότητα των τιμών αποδοτικότητας των υπο-διαδικασιών, συνεπώς θέτουν σε αμφισβήτηση την εγκυρότητα των παραγόμενων αποτελεσμάτων. Εμφανίζονται δηλαδή περιπτώσεις όπου το ίδιο επίπεδο συνολικής αποδοτικότητας του συστήματος μπορεί να προκύπτει από διαφορετικούς συνδυασμούς τιμών αποδοτικότητας των επιμέρους διαδικασιών. Επίσης, υπάρχουν προσεγγίσεις που συχνά μεροληπτούν κατά την αποτίμηση της συνολικής αποδοτικότητας του συστήματος. Αποδεικνύουμε ότι η αθροιστική μέθοδος μεροληπτεί κατά την αποτίμηση υπέρ κάποιων συγκεκριμένων σταδίων. Επιπροσθέτως, δείχνουμε ότι οι προτεινόμενες προσεγγίσεις δεν μπορούν να εφαρμοστούν σε γενικές δικτυακές δομές παραγωγικών μονάδων.

Για την αντιμετώπιση των παραπάνω αδυναμιών, εισάγουμε νέες μεθοδολογίες που βασίζονται στην ενσωμάτωση τεχνικών πολυκριτήριου προγραμματισμού στην Περιβάλλουσα Ανάλυση Δεδομένων. Επικεντρώνουμε την έρευνά μας σε μονάδες απόφασης που περιλαμβάνουν δύο υπο-διαδικασίες διατεταγμένες σε σειρά και μοντελοποιούμε το πρόβλημα της μέτρησης της αποδοτικότητάς τους ως πρόβλημα πολυκριτήριου προγραμματισμού. Χρησιμοποιούμε πραγματικές συναρτήσεις επίτευξης (achievement scalarizing functions) ώστε να ενσωματώσουμε τις ιδέες μας αλλά και τις ιδιότητες της αμεροληψίας και της μοναδικότητας των αποτελεσμάτων που θα πρέπει να διέπουν οι μέθοδοι της Περιβάλλουσας Ανάλυσης Πολυσταδιακών Διεργασιών. Εισάγουμε τη συνθετική προσέγγιση (composition approach), αντιμετωπίζοντας με ουδετερότητα τις υπο-διαδικασίες και κατασκευάζουμε αρχικά ένα μοντέλο με μια προσθετική συνάρτηση επίτευξης βασιζόμενοι στην L_1 μετρική. Αυτό το μοντέλο αποδίδει αμερόληπτα αποτελέσματα, τα οποία απεικονίζονται ως ακραία σημεία (κορυφές) στο σύνορο Pareto. Σχηματίζουμε επιπλέον μοντέλα χρησιμοποιώντας συναρτήσεις επίτευξης, για την κατασκευή των οποίων εφαρμόζουμε μεθοδολογίες πολυκριτήριας βελτιστοποίησης που βασίζονται σε σημεία αναφοράς (reference points). Ειδικότερα, χρησιμοποιούμε την μετρική Tchebycheff (L_∞) για τον εντοπισμό μιας μοναδικής Pareto βέλτιστης λύσης ελαχιστοποιώντας τη μέγιστη απόκλιση από το ιδεώδες σημείο (ideal point). Ήτοι στοχεύει στον υπολογισμό των επιμέρους αποδοτικοτήτων των υπο-διαδικασιών όσο δύναται πλησιέστερα στα υψηλότερα επίπεδα αποδοτικοτήτων που μπορούν να επιτύχουν οι υπο-διαδικασίες ξεχωριστά. Το μοντέλο αυτό αποδίδει αμερόληπτα αποτελέσματα και διασφαλίζει τη μοναδικότητά τους. Έπειτα, αναπτύσσουμε δύο μεθόδους που παρέχουν την απαραίτητη πληροφορία για τον σχηματισμό των προβολών των μη αποδοτικών μονάδων στο σύνορο αποδοτικότητας. Η πρώτη προκύπτει απ’ ευθείας από την συνθετική προσέγγιση ενώ η δεύτερη είναι προσανατολισμένη στο να καταστήσει αποδοτικές τη μη αποδοτικές μονάδες επιφέροντας όσο το δυνατόν ελάχιστες αλλαγές στα αρχικά επίπεδα των ενδιάμεσων μεγεθών.

Στη συνέχεια, διατυπώνουμε έναν νέο ορισμό της συνολικής αποδοτικότητας των μονάδων απόφασης που περιέχουν δύο υπο-διαδικασίες διατεταγμένες σε σειρά, εμπνευσμένοι από τον ρόλο του αδύναμου κρίκου στις εφοδιαστικές αλυσίδες και από το θεώρημα της μέγιστης ροής-ελάχιστης κοπής (max flow-min cut) στα δίκτυα. Για την αξιολόγηση των μονάδων που περιέχουν δύο υπο-διαδικασίες με ποικίλη σειριακή διάταξη εισάγουμε την προσέγγιση του «αδύναμου κρίκου» (weak-link approach). Αναπτύσσουμε μια νέα μέθοδο βελτιστοποίησης max-min δύο φάσεων με την οποία διασφαλίζεται ότι η προκύπτουσα λύση θα είναι μοναδική και βέλτιστη κατά Pareto. Πρωτίστως, μεγιστοποιούμε την ελάχιστη αποδοτικότητα (αποδοτικότητα του αδύναμου κρίκου) μεταξύ των υπο-διαδικασιών και στη συνέχεια διασφαλίζουμε ότι η προκύπτουσα λύση είναι μοναδική και βέλτιστη κατά Pareto. Για την καθοδήγηση της διαδικασίας βελτιστοποίησης χρησιμοποιήσαμε τις ιδεώδεις αποδοτικότητες των υπο-διαδικασιών, ωστόσο, διαφορετικές προτιμήσεις δύναται να ενσωματωθούν για τον εντοπισμό εναλλακτικών βέλτιστων κατά Pareto λύσεων.

Τέλος, επανεξετάζουμε τη μεθοδολογία των Aviles-Sacoto et al (2015) και αποδεικνύουμε ότι είναι προβληματική. Προτείνουμε μια εναλλακτική μοντελοποίηση η οποία διορθώνει τα μεθοδολογικά προβλήματα που παρατηρούμε.

Λέξεις Κλειδιά: *Περιβάλλουσα Ανάλυση Δεδομένων, Περιβάλλουσα Ανάλυση Πολυσταδιακών Διεργασιών, συνθετική προσέγγιση, μέθοδος του αδύναμου κρίκου, πολυκρίτηριος προγραμματισμός*.

Abstract

The systematic performance evaluation of the organizations as well as the target setting are key aspects for its proper operation and viability. Thus, the adoption of evaluation methods is necessary, which are capable of taking into account all the environmental factors of the organization, identifying the inefficient production processes and suggesting adequate ways to improve them. Such a method is Data Envelopment Analysis (DEA), which is the most popular non-parametric technique for assessing the efficiency of homogeneous decision making units (DMUs) that use multiple inputs to produce multiple outputs.

The DMUs may consist of several sub-processes (also known as stages, sub-units, divisions etc.) that interact and perform various operations. However, the classical DEA models treat the DMU as a "black box", i.e. a single stage production process that transforms some external inputs to final outputs. In such a setting, the internal structure of the DMU is not taken into consideration. Thus, the conventional DEA models fail to mathematically represent the internal characteristics of the DMUs, as well as they fall short to provide precise results and useful information regarding the sources that cause inefficiency. In order to take into account for the internal structure of the DMUs, recent methodological advancements are developed, which extend the standard DEA and constitute a new field, namely the network DEA. The network DEA methods are capable of reflecting accurately the DMUs' internal operations as well as to incorporate their relationships and interdependences. In network DEA, the DMU is considered as a network of interconnected sub-units, with the connections indicating the flow of intermediate products (commonly called intermediate measures or links). An indicative example of such a DMU is a supply chain, which has a network structure and is composed of several members whose performances affect the overall performance of the supply chain. Therefore, the overall efficiency of the supply chain (DMU) should be evaluated by taking into account the individual efficiencies of its members in a coordinated manner.

In this thesis, we conduct a critical survey and categorization of the state-of-the art network DEA methods and we classify a great volume of network DEA studies based on the assessment method they follow. We unveil the relations and the differences of

the existing network DEA methods. Also, we uncover their defects concerning the returns to scale, the inconsistency between the multiplier and the envelopment models as well as the inadequate information that provide for the calculation of efficient projections. The most important network DEA methods do not secure the uniqueness of the efficiency scores, i.e. the same level of overall efficiency is obtained from different combinations of the efficiencies of the sub-processes. Also, we prove that the additive efficiency decomposition method unduly and implicitly assigns different priority to the sub-processes, hence provides biased efficiency assessments. Finally, we discuss about the inability of the existing approaches to be universally applied on every type of network structure.

We develop two new approaches in network DEA that overcome effectively the deficiencies and provide unique and unbiased efficiency scores, based on a multiple objective framework. We focus our research to serial two-stage network structures and we formulate the problem of their efficiency assessment as a multi-objective mathematical programming problem. Initially, we introduce the *composition approach* to two-stage network DEA, which is based on a bi-objective mathematical program for the efficiency assessments. We employ two scalarization techniques, firstly based on the L_1 norm we aggregate the two objective functions additively without giving any priority between them. The application of this scalarizing function yields an extreme (vertex) Pareto-optimal solution. Then, we employ a min-max scalarization technique, i.e. the Tchebycheff norm (L_∞), which minimizes the distance between the ideal point and the feasible objective functions space so as to locate a point on the Pareto front not necessarily extreme. This model provides unique and unbiased efficiency scores. In the composition approach, we estimate first the stage efficiencies and then we aggregate them either additively or multiplicatively to obtain the overall efficiency. Next, we develop two methods to derive the efficient frontier in two-stage DEA and provide efficient projections. The first naturally stems from our composition approach, while the second seeks to provide efficient projections by altering the original levels of the intermediate measures at a minimum distortion.

We build upon the composition approach and we introduce the "*weak-link*" approach to two-stage network DEA, which inherits the nice properties of the former, i.e. provides unique and unbiased efficiency scores. Also, the "weak-link" approach can be readily applied to various types of two-stage network structures. In this

approach, we introduce a novel definition about the overall efficiency of the DMU, inspired by the "weak link" notion in supply chains and the maximum-flow/minimum-cut problem in networks. We incorporate this notion into the assessment by assuming that given the stage efficiencies, the system efficiency can be viewed as the maximum flow through the network and can be estimated as the min-cut of the network, i.e. the system efficiency derives as the lowest of the stage efficiencies. We mathematically represent this concept by employing a two-phase max-min optimization method in a multi-objective programming framework, which seeks to maximize the minimum weighted achievement from zero-level efficiency, i.e. maximizing the lowest of the stage efficiencies (weak link). The proposed two-phase procedure estimates the stage efficiencies and the overall efficiency simultaneously by providing a unique Pareto optimal solution. The search direction towards the Pareto front is driven by the assumption that the stage efficiencies are proportional to their independent counterparts. External priorities can be also introduced to our methodology so as to obtain alternative Pareto optimal solutions. We conduct a systematic investigation of the sensitivity of the weak link so as to identify the source of inefficiency in the two-stage processes.

Finally, we revisit the work of Aviles-Sacoto et al (2015) who studied a peculiar situation of a two-stage process where some of the intermediate measures are inputs to the second stage and at the same time external outputs from that stage. We show that their modelling approach departs from the described setting and adapts a different situation, where the specific intermediate measure is viewed either as input to or as output from the second stage of the process. We alternatively propose a different modelling approach for the performance assessment of the two-stage process under examination, which rectifies the methodological problems that we observe.

Keywords: *Data Envelopment Analysis (DEA), Network DEA, composition approach, weak-link approach, multi-objective programming.*

Acknowledgements

This thesis could not be accomplished without the persistence, the determination and the passion for the subject of my professors and my colleagues. I feel extremely fortunate to have worked with them. I am greatly indebted to Professor D. Despotis for being more than an Advisor. He supervised me and guided me all these years and eventually he contributed greatly to my overall advancement and evolution as a student and as a person. I wish to express my sincerest appreciation to the Associate Professor D. Apostolou and the Associate Professor K. Metaxiotis for their support, collaboration and fertile conversations that broadened my horizons. I am also grateful to Dr. Ioannis Smyrlis for the collaboration, the steady encouragement and the constructive guidance.

As a PhD student I had the opportunity to be for a month as a guest with the Operations and Information Management Group at Aston University, UK. I would like to thank professor E. Thanassoulis for the invitation, as well as the members of the group for the hospitality and the exchange of ideas. While too numerous to mention, I would like to thank my teachers of all levels of education, each one contributed to my education and personal evolution. Moreover, I thank all current and former members of the Decision Support Systems laboratory of University of Piraeus and especially A. Mikeli, A. Zigomitros, Dr. K. Liagkouras and Dr. D. Sotiropoulos for the perfect atmosphere in the working group, the fruitful comments and the helpful advices. I have to make special mention and to thank my dear friend D. Sotiros for all his multilateral support; we still have a lot to go.

I would like to honor and dedicate this thesis to my family and my beloved aunts. Finally, I wish to express my deep gratitude and dedicate also this thesis to the wonderful people Matina, Diamantis, Ioanna and Giannis that stand by me.

Koronakos Grigorios,

Piraeus, March 2017

Contents

List of Figures

List of Tables

Chapter 1

Introduction

Performance measurement deals with ongoing monitoring and evaluation of the operations of the organizations so as to be able to improve their productivity and performance. Thus, the performance measurement is a subject of major importance. Improving organization's performance requires accurate understanding as well as systematic assessment of its internal structure, which is often a tough task because of organization's complexity.

Two main approaches, the parametric and the non-parametric, are suggested in the literature for the performance measurement of production units. In the parametric approach, a production function is explicitly assumed so as to describe the relationships among the inputs and the outputs that participate in the production process. However, the production function can be hardly formulated or is completely unknown. On the contrary, the non-parametric approach does not require any a priori specification of the underlying functional form that relates the inputs with the outputs. Data Envelopment Analysis (DEA) is a powerful non-parametric technique that is widely used for evaluating the performance of a set of comparable entities, called decision making units (DMUs), which use multiple inputs to produce multiple outputs. Charnes et al (1978) introduced DEA, based on Farrell's (1957) work of estimating technical efficiency with respect to a production frontier. DEA circumvents the problem of specifying an explicit form of the production function by constructing an empirical best practice production frontier by enveloping the observed data of the DMUs. The linear programming is the underlying mathematical method that enables DEA to determine the efficient production frontier and calculate the efficiency score of each DMU. The efficiencies provided by DEA are relative rather than absolute, because each unit is evaluated relative to the production frontier, i.e. the best practice units. DEA is capable of uncovering the sources of inefficiency and providing prescriptions for improving the inefficient units. DEA takes into account the returns to scale and the orientation of the analysis in calculating efficiency. The CCR (Charnes et al, 1978) and the BCC (Banker et al, 1984) models have established the foundation for further research in this field. A rapid and continuous growth has been reported since then, both in theoretical and application level. The theoretical advances of DEA as well as procedural issues are sufficiently described in Dyson et al (2001), Thanassoulis et al (2008) and Cook and Seiford (2009). A remarkable body of literature has been developed with a wide range of applications to measure the efficiency in various sectors such as business and finance, public services, education, health care, transportation, agriculture, supply chains etc. (Gattoufi et al, 2004).

The DMUs may have a complex structure that includes several interdependent operations (also known as stages, sub-processes, sub-units, divisions etc.) with a series, parallel or series-parallel arrangement. The conventional DEA models, however, regard the DMU as a *black box*, treating them as single stage production processes that transform some external inputs to final outputs. In such a setting, the internal structure and the interactions among the comprised operations of the DMUs are not taken into consideration. However, a significant number of studies has focused on assessing efficiency in multi-stage production processes, where outputs from some stages, characterized as intermediate products (measures), are used either as inputs to the other stages or as external outputs of the production process.

Seiford and Zhu (1999) assessed the efficiency of commercial banks in US by considering the bank operations as a two-stage process. They assessed, however, the stage efficiencies and the overall system efficiency independently with distinct standard DEA models. The network DEA extends and complements the conventional DEA by considering not only the internal structure of the DMUs but also the interactions among the sub-processes. When theses interactions are not taken into account, the results may be distorted and misleading. Hence, in contrast to traditional DEA models, the network DEA models provide more accurate results and further insights concerning the sources of inefficiency. Fare and Grosskopf (1996, 2000) were among the first to deal with efficiency assessments in network DEA. Thorough classifications of network DEA models and methods developed for various network configurations are given in Castelli et al (2010) and Kao (2014b). Moreover, a collection of network DEA models is given in Cook and Zhu (2014). In this thesis, we focus on two-stage series processes that are extensively studied in the literature. We show that basic network DEA methods proposed in the literature suffer from shortcomings that should be rectified before moving to more complex structures with many stages. Prominent approaches developed to deal with two-stage series processes are the multiplicative decomposition approach (Kao and Hwang, 2008), the additive decomposition approach (Chen et al, 2009b) and the slacks-based network DEA model (Tone and Tsutsui, 2009).

These are the first approaches to assess the stage efficiencies and the overall system efficiency jointly in one program.

Kao and Hwang (2008) introduced an innovative approach by taking into account a series relationship of the two stages and developed a model to estimate the overall efficiency of the production process as the product of the efficiencies of the two individual stages. Their multiplicative decomposition approach is based on the reasonable assumption that the values of the intermediate measures (virtual intermediate measures) are the same, no matter if they are considered as outputs of the first stage or inputs to the second stage. As they noted, the decomposition of the overall efficiency to the stage efficiencies is not unique. In order to check the uniqueness, they proposed a post-optimality procedure, to obtain the largest first (or second) stage efficiency score while keeping the overall efficiency unchanged. Liang et al (2008) under the same framework view the efficiency assessments in two-stage process in terms of a game approach.

Maintaining the series relationship between the two stages, Chen et al (2009b) introduced the additive efficiency decomposition in two-stage processes. They derive the overall efficiency of the production process as a weighted average of the efficiencies of the individual stages. Their modeling approach facilitates the linearization of a non-linear mathematical program by assuming that the weights of the two stages derive endogenously by the optimization process. However, we prove that this assumption leads to biased efficiency assessments.

An issue investigated further in the literature is the derivation of the efficient frontier in two-stage DEA. Chen et al (2010a) pointed out that adjusting the inputs and the outputs by the efficiency scores is not sufficient to yield a frontier projection, when the additive decomposition model is assumed. They developed instead, a model for deriving the efficient frontier within the Kao and Hwang (2008) multiplicative framework. The inability of the two-stage DEA models to locate correctly the efficient frontier, as it is the case with standard DEA, is further examined in Chen et al (2013). In this paper, it was demonstrated that under general network structures, the multiplier and the envelopment network DEA models are two different approaches, thus, alternative methods to overcome this deficiency were reviewed.

In this thesis, we describe the advantages of the network DEA methods over the classical DEA ones. We present in detail the most important of them and we provide a survey of the network DEA studies across the literature. In addition, we carry out a critical review of the fundamental approaches in two-stage network DEA, namely the additive and the multiplicative efficiency decomposition approaches and we discuss their inherent limitations and shortcomings. The decomposition approaches provide non-unique efficiency scores, while in the additive approach the assessment is biased. Based on a reverse perspective on how to obtain and aggregate the stage efficiencies, that of the composition as opposed to the decomposition, we introduce the *composition approach* to two-stage network DEA that effectively overcomes the deficiencies of the aforementioned decomposition methods. In other words, it provides unique and unbiased efficiency scores for the individual stages, which are then composed to obtain the overall efficiency, by selecting the aggregation method a posteriori. Also, we develop an envelopment model to derive the efficient projections and render efficient the inefficient units.

Based on the composition approach we build the "weak-link" approach, which can be applied to two-stage network structures of varying complexity. Inspired by the "weak link" notion in supply chains and the maximum-flow/minimum-cut problem in networks we introduce a novel definition of the system efficiency in two-stage network DEA. We adapt this notion to the performance assessment of the two-stage processes by employing a two-phase max-min optimization model in a multi-objective programming framework. We drive the quest towards the Pareto front by rationally assuming that the stage efficiencies are proportional to their independent counterparts. The proposed two-phase multi-objective procedure provides a unique Pareto optimal solution, i.e. unique stage efficiency scores, and the overall efficiency is derived as the lowest of the stage efficiencies. The properties of unique and unbiased efficiency scores enable us to identify sufficiently the source of inefficiency and demonstrate that the *"weak-link"* approach excels the decomposition approaches.

Finally, in this thesis, we revisit the paper of Aviles-Sacoto et al (2015) who assessed 37 undergraduate business programs in U.S. as two-stage processes within the peculiar situation of some of the intermediate measures, namely the internships, playing both an input and output role in regard to the second stage. We reveal that the proposed modelling approach deviates from the described scenario and depicts a different situation where the specific intermediate measure is viewed either as input to or as output from the second stage of the process. We develop instead an alternative modelling approach, within the context of network DEA, so as to amend this issue.

1.1 Motivation and objectives of the thesis

Network DEA broadens the application field of standard DEA so as to allow for efficiency assessments when the DMUs have complex internal structure that consists of several sub-processes. There is an increasing literature body on the field of network DEA and a variety of methods, however, there is only one critical review (Chen et al, 2013) and four surveys (Castelli et al, 2010, Agrell and Hatami-Marbini, 2013, Halkos et al, 2014 and Kao, 2014) with divergent views on the classification of the studies concerning the modelling approach and the network structure of the process. The most significant methods on the field of network DEA, namely the multiplicative and the additive decomposition, have inherent defects because they yield non-unique efficiency scores as reported in the literature. Moreover, we show that the latter provides biased efficiency scores. In addition, the slacks-based measure (SBM) approach (Tone and Tsutsui, 2009), which has already received much attention from the research community, cannot be formulated and applied to production processes with specific structures.

The above motivates us to define the objectives of this thesis as follows:

- To unveil relations and differences among the existing network DEA methods and present the origin and evolution of the most important ones.
- To offer a thorough categorization and critical survey of the state-of-the art network DEA methods.
- To uncover the deficiencies of the existing network DEA methods.
- To provide a deep examination of these defects so as to unveil their effects and give comprehensive interpretations.
- To develop alternative network DEA methods that amend the reported defects.

1.2 Contribution of the thesis

The contribution of the thesis to the network DEA literature is outlined below:

- Provides a thorough survey of the network DEA literature and classifies the network DEA studies according to the modelling approach they follow.
- Reveals the shortcomings of the network DEA methods concerning the returns to scale, the inconsistency between the multiplier and the envelopment models, the non-unique efficiency scores, the biased assessments and the inability to be universally applied.
- Establishes the properties that the network DEA methods should meet.
- Introduces the *composition approach* to two-stage network DEA, as opposed to the efficiency decomposition approach, by formulating the efficiency assessment of multi-stage processes as multi-objective mathematical programming problem. The new approach provides unique and unbiased efficiency scores.
- Provides methods to derive the efficient frontier in two-stage DEA.
- Introduces a novel definition of the system efficiency in two-stage processes, inspired by the "*weak link*" notion in supply chains and the maximum-flow/minimum-cut problem in networks.
- Develops the *"weak-link" approach* to network DEA for the performance assessment of two-stage processes of varying complexity, which identifies effectively the source of inefficiency in a multi-stage process.
- Reveals that the modelling approach of Aviles-Sacoto et al (2015) is misleading and proposes enhancements by developing alternative modelling formulations for the performance assessment of the specific two-stage process under examination.

1.3 Organization of the thesis

The thesis is organized as follows:

Chapter 2: This chapter discusses the history and the origins of Data Envelopment Analysis (DEA). It also provides the basic DEA concepts and definitions and describes the properties that render DEA a powerful technique for performance measurement. We present the classical DEA models, namely the CCR, the BCC, the additive and the slacks based measure (SBM), and we discuss the relations and differences among them.

Chapter 3: Demonstrates the advantages of network DEA over the standard DEA for the assessment of multi-stage processes. It also describes the evolution of network DEA, it provides a detailed survey of the network DEA literature and categorizes the existing studies based on the assessment approach that they follow. We discern two assessment paradigms, the independent assessment paradigm and the joint assessment paradigm. In the independent paradigm the performances of the DMUs and the sub-processes are assessed independently whereas in the joint paradigm the internal structure of the DMUs and the interdependencies among the sub-processes are taken into consideration. There are three general approaches that follow the joint assessment paradigm: the efficiency decomposition approach, the slacks-based measure approach and the system-centric approach. The most influential approaches are the efficiency decomposition and the slacks-based measure. Thus, we discuss in detail the most important network DEA methods of those categories and their extensions. In particular, we present the additive and the multiplicative decomposition methods as well as the network slacks based measure (NSBM). Also, we report the limitations of the aforementioned methods, concerning the uniqueness of the stage efficiency scores, the returns to scale assumed, the inconsistency between the multiplier and the envelopment models as well as the insufficient information that provide for the calculation of efficient projections.

Chapter 4: In this chapter, we revisit the additive and the multiplicative efficiency decomposition methods to discuss their shortcomings. Then, based on a reverse perspective on how to obtain and aggregate the stage efficiencies, that of the composition as opposed to the decomposition, we develop the *composition approach* to two-stage network DEA that overcomes the deficiencies of the decomposition methods. Selecting an output orientation for the first stage and an input orientation for the second stage, we show that it is possible to obtain unbiased efficiency scores for the two stages in a bi-objective optimization framework. We propose two alternative models by employing different scalarizing functions in a multi-objective linear programming (MOLP) model. Firstly, we aggregate additively the two objectives in a single objective that locates an extreme (vertex) Pareto-optimal solution. Then, we develop a min-max model that provides unique efficiency scores by locating a point on the Pareto front, not necessarily extreme. In the latter case, the stage efficiencies obtained are more balanced. The individual efficiency scores are then used to calculate the overall efficiency of the production process, by selecting the aggregation (composition) method a posteriori. As the conflicting role of the intermediate measures gives a peculiar character to two-stage processes that obscures the standard DEA premises, we develop an envelopment model to locate the efficient frontier, whose derivation from our primal multiplier model is justified.

Chapter 5: In this chapter, we introduce a novel definition of the system efficiency in two-stage network DEA, inspired by the “weak link” notion in supply chains and the maximum-flow/minimum-cut problem in networks. We adapt this notion to fit the multi-stage processes dealt with in network DEA, by assuming that given the stage efficiencies, the system efficiency can be viewed as the maximum flow through the network and can be estimated as the min-cut of the network. Thus, our primary goal is to estimate the stage efficiency scores in a manner that the minimal stage efficiency (the weak link) and, thus, the overall system efficiency gets the maximum possible value. The mathematical representation of this concept is expressed by the weighted max-min formulation which seeks to maximize the minimum weighted achievement from zero-level efficiency. For this purpose, we develop a two-phase max-min optimization technique in a multi-objective programming setting to estimate the stage efficiencies and the overall efficiency simultaneously. The "weak-link" approach is an advancement of the *composition approach*, therefore it gains the nice properties of providing unique and unbiased efficiency scores unlike the decomposition methods. The "weak-link" approach can be applied to various types of two-stage network structures. The proposed two-phase procedure secures the Pareto optimality condition and provides a unique point on the Pareto front, i.e. unique stage efficiency scores, by maximizing the lowest of the stage efficiencies (weak link). We derive the stage efficiency scores based on the assumption that they are proportional to their independent counterparts, i.e. the independent stage efficiencies define the search direction towards the Pareto front. However, beyond this rational assumption, external priorities can be explicitly introduced to our models to obtain alternative Pareto optimal solutions, i.e. different stage and overall efficiencies. Also, we provide a systematic investigation of the sensitivity of the weak link in order to identify adequately the source of inefficiency. Finally, a detailed comparison with the multiplicative approach highlights the advantages of our method.

Chapter 6: Revisits the work of Aviles-Sacoto et al (2015) who studied the efficiency assessment of 37 undergraduate business programs in U.S. as two-stage processes, where some of the intermediate measures are inputs to the second stage and at the same time external outputs from that stage. Under this peculiar situation they developed a network DEA assessment framework based on the additive decomposition method. However, as we demonstrate, the original modelling approach followed in Aviles-Sacoto (2015) arbitrarily, yet unnecessarily, departs from the described setting and adapts a different situation, where the specific intermediate measure is viewed either as input to or as output from the second stage of the process. Thus, we propose an alternative modelling approach for the performance assessment of the undergraduate business programs, in the context of network DEA.

Chapter 7: Concluding remarks are drawn based on the research findings and directions for future research on the field of network DEA are provided.

1.4 Publications based on the thesis

International journals

- Despotis DK, Koronakos G, Sotiros D (2016). Composition versus decomposition in two-stage network DEA: a reverse approach. Journal of Productivity Analysis, 45(1), pp. 71-87.

- Despotis DK, Koronakos G, Sotiros D (2016). The "weak-link" approach to network DEA for two-stage processes. European Journal of Operational Research, 254(2), pp. 481-492.

In proceedings of international conferences

- Despotis DK, Koronakos G (2014). Efficiency assessment in two-stage processes: A novel network DEA approach. Procedia Computer Science, 31, pp. 299-307.

- Koronakos G, Despotis DK. A novel DEA approach to assess individual and overall efficiencies in two-stage processes. 11th International Conference on Data Envelopment Analysis, Samsun, Turkey, June 27-30, 2013, pp. 7-12.

- Despotis DK, Koronakos G, Sotiros D. The "weak link" approach to network DEA: The case of two-stage processes. DEA2016-14th International conference on Data Envelopment Analysis, Wuhan, China, May 23-26, 2016, pp. 145-151.

In proceedings of national conferences

- Koronakos G, Sotiros D, Despotis DK. Περιβάλλουσα Ανάλυση Δεδομένων σε παραγωγικές διαδικασίες δύο σταδίων. 4ο Φοιτητικό συνέδριο της ΕΕΕΕ, Δεκέμβριος 17-18, 2015, Αθήνα, Ελλάδα, pp. 76-80.

Presentations (book of abstracts) in international conferences

- Koronakos G, Despotis DK, Sotiros D. Additive decomposition in two-stage DEA: An alternative approach. 25th European Conference on Operational Research, Vilnius, Lithuania, July 8-11, 2012.

- Despotis DK, Koronakos G. Two-Stage Data Envelopment Analysis: Foundation and recent developments. International conference on Optimization, Computing and Business Analytics (ICOCBA2012), Kolkata, India, December 20-22, 2012.

- Despotis DK, Koronakos G. Models for unbiased efficiency assessments in an additive two-stage DEA framework. 26th European Conference on Operational Research, Rome, Italy, July 1-4, 2013.

- Koronakos G, Despotis DK. A study on two-stage Network DEA Models with some improvements. 20th Conference of the International Federation of Operational Research Societies, Barcelona, Spain, July 13-18, 2014.

Chapter 2

Data Envelopment Analysis

Data Envelopment Analysis (DEA) is a non-parametric and "*data driven*" technique that is based on linear programming. It was developed for measuring the relative efficiencies of a set of comparable entities, called Decision Making Units (DMUs), which convert multiple inputs into multiple outputs. The definition of a DMU is generic since the application field of DEA includes various forms of DMUs, such as hospitals, universities, banks, supply chains, countries and so forth. As noted in Cooper (2014), the term "*DMU*" is coined so as to emphasize the purpose of DEA "*to be useful to managers and policy makers not only in measuring the efficiency of different organizations but also in guiding them how that could be accomplished*". The fundamental concept of DEA is to identify the best practice DMUs so as to form an efficient frontier and act as benchmark for the non-frontier units. Using this information, the non-frontier (inefficient) units can be compared with the benchmarks and their level of efficiency can be measured.

DEA was developed as an alternative to the econometric (parametric) approach for the efficiency measurement of production units. In the econometric approach an explicit production function is assumed (e.g. Cobb-Douglas) and the parameters of this function are estimated so as to fit the observations. On the other hand, in DEA no assumptions are made for the underlying functional form that transforms inputs to outputs. DEA builds an empirical best practice production frontier based on the observed data and provides performance measures by comparing the observations by the best practice units.

As noticed in Cooper et al (2007), Cooper and Lovell (2011) and Cooper (2014), the basic concepts of DEA were initiated by Farell (1957), who established a piecewise linear envelopment of the data as the most pessimistic specification of the production frontier by solving systems of linear equations. Farrell reported the inadequacies of the existing efficiency and productivity measures in the presence of multiple inputs, i.e. the average productivity for an individual input (ignoring all other inputs) and the efficiency index in which a weighted average of inputs is compared with output. Being inspired by the activity analysis of Koopmans (1951) and Debreu (1951), he provided new efficiency measures which are based on radial equiproportionate reductions or expansions of the inefficient observations to the production frontier. That is to say, Farell (1957) introduced a measure of technical efficiency based on the relative notion of comparing the inputs and outputs of similar DMUs, i.e. he proposed a measure of relative technical efficiency. However, Farell (1957) failed to generalize his formulations to the multiple inputs-outputs case and confined to situations with many inputs but a single output.

Charnes et al (1978) built upon Farrell’s work and the works of Charnes and Cooper (1961), (1962), on linear and fractional programming respectively, and made the connection between Farrell’s technical efficiency measure and the classic output to input ratio measure of efficiency. Actually, Charnes et al (1978) enabled the efficiency assessment to deal with multiple inputs and outputs by generalizing the single output-input ratio measure of efficiency in terms of a fractional linear program. In order to be aligned with the classic definition of efficiency, i.e. the ratio of output to input, they transformed the multiple inputs and outputs into single “virtual” input and “virtual” output. The virtual input and output are formed as weighted sums, where the weights are not given *a priori* but they derive from the optimization process. Also, they presented the conversion of the fractional linear program to an equivalent linear program, which as Cooper et al (2007) noticed is “dual to the problem formulated by Farrell”.

It is worth to note that Farrell and Fieldhouse (1962) based on Hoffman’s comments about Farrell’s framework (cf. the discussion section for his work included in Farrell’s paper), provided sufficiently the linear programming formulation for the single-output case. However, they did not provide rigorous mathematical details and interpretations. Bringing them all together, A. Charnes, W.W. Cooper and E. Rhodes established the DEA for the performance measurement of DMUs and extended its power and computational convenience based on mathematical programming. They also provided a strict mathematical framework that eases the analysts and the decision makers to follow and comprehend.

The origins and the historical evolution of DEA are succinctly outlined by Seiford (1996), Forsund and Sarafoglou (2002), Cooper and Lovell (2011) and Cooper (2014).

2.1 Basic concepts and definitions

The DMUs are homogeneous production units that transform multiple inputs into multiple outputs, which are non-negative scalars. In real-world problems the identification of the inputs and the outputs for the performance assessment of DMUs may be a rough task. It is however of great importance, because the chosen inputs and outputs should effectively portray the DMUs' operations. The inputs and the outputs of DMUs could be of different type and of unit of measurement. A schematic representation of a DMU is given in Figure 2.1 below.

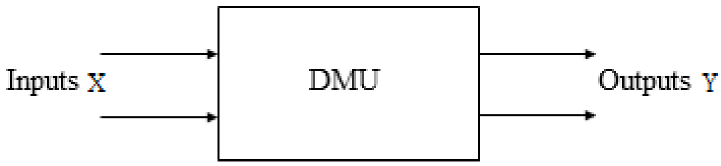


Fig. 2.1: A typical representation of a DMU

2.1.1 Production Possibility Set (PPS)

As noticed, the primary purpose of DEA is to identify the best production practice among the DMUs so as to derive an empirical production frontier. The best production practice is derived from any DMU of the production possibility set (PPS) that produces the highest possible levels of output given its level of inputs. The PPS, also known as technology, is related to the production process operated by the DMUs. It is a convex set that contains all the feasible levels of outputs that can be produced from the available levels of inputs, no matter if they are not observed in practice. These feasible input-output correspondences are obtained using interpolations between the observed input-output bundles of the DMUs. Assuming n DMUs, each using m inputs $X_j = (x_{ij}, i = 1, \dots, m)$, $\mathrm{j} = 1, \dots, \mathrm{n}$ to produce s outputs $Y_j = (y_{rj}, r = 1, \dots, s)$, we denote the PPS as T:

$$T = \{(X, Y) \in \Re_{+}^{m+s} |\; X \; can \; produce \; Y\}$$

In DEA, a certain assumption concerning the returns-to-scale (RTS) is required for the construction of the PPS. The RTS describe the environment in which the DMUs operate. In particular, the RTS indicate for each DMU the relation between a proportional change in inputs and the resulting change in outputs. Under the constant returns to scale (CRS) assumption a proportional increase in input levels results in an equiproportionate increase in the output levels. If this assumption does not hold, then the DMUs operate under variable returns to scale (VRS). Specifically, when the output levels increase by a greater proportion than the proportional increase in the input levels, then increasing returns to scale (IRS) exist. On the other hand, when the output levels increase by a lower proportion than the proportional increase in the input levels, then decreasing returns to scale (DRS) exist. Conclusions about the RTS should be drawn from the specific characteristics of the DMUs and the environment in which they operate.

Beyond the assumption of the RTS, the PPS (T) construction is based on the principle of minimal extrapolation, i.e. the PPS is the smallest convex set enveloping all the observations (DMUs). The PPS is build on the following axioms:

- *Inclusion of observations*:

  Each observed DMU $\begin{bmatrix} X_j \\ Y_j \end{bmatrix} \in T$.

- *Monotonicity*:

  If $\begin{bmatrix} X_j \\ Y_j \end{bmatrix} \in T$ and $X_0^+ \geq X_j$ then $\begin{bmatrix} X_0^+ \\ Y_j \end{bmatrix} \in T$.

  If $\begin{bmatrix} X_j \\ Y_j \end{bmatrix} \in T$ and $Y_0^- \leq Y_j$ then $\begin{bmatrix} X_j \\ Y_0^- \end{bmatrix} \in T$.

- *No output can be produced without some input:*

  If $Y_j \geq 0$ and $Y_j \neq 0$ then $\begin{bmatrix} 0 \\ Y_j \end{bmatrix} \notin T$.

- *Ray unboundedness (CRS)*:

  If $\begin{bmatrix} X_j \\ Y_j \end{bmatrix} \in T \Longrightarrow k \begin{bmatrix} X_j \\ Y_j \end{bmatrix} \in T, \ \forall \ k \in \Re_0^+; \ j = 1, \dots, n$.

- *Convexity*:

  Any convex combination of DMUs that belong to *T*, belong to *T*, i.e.

  $\sum_{j=1}^{n} \lambda_j \begin{bmatrix} X_j \\ Y_j \end{bmatrix} \in T, \sum_{j=1}^{n} \lambda_j = 1, \lambda_j \geq 0, j = 1, \dots, n$.

The mathematical representation of the PPS for the n DMUs under constant returns to scale (CRS) assumption is as follows:

$$T^{CRS} = \left\{ (X, Y) \mid \sum_{j=1}^{n} \lambda_j X_j \leq X, \sum_{j=1}^{n} \lambda_j Y_j \geq Y, \lambda_j \geq 0, j = 1, \dots, n \right\} \quad (2.1)$$

The corresponding PPS under variable returns to scale (VRS) is given below:

$$T^{VRS} = \left\{ (X, Y) \mid \sum_{j=1}^{n} \lambda_j X_j \leq X, \sum_{j=1}^{n} \lambda_j Y_j \geq Y, \sum_{j=1}^{n} \lambda_j = 1, \lambda_j \geq 0, j = 1, \dots, n \right\} \quad (2.2)$$

The difference between T^{CRS} and T^{VRS} is the convexity constraint $\sum_{j=1}^{n} \lambda_j = 1$. The incorporation of this constraint allows for constant, increasing and decreasing returns to scale. Also, it leads to smaller PPS with tighter envelopment of the observations (DMUs). We provide in Figure 2.2 a schematic representation of T^{CRS} and T^{VRS} for a case of seven DMUs, labelled A to G, that use one input to produce one output, as presented in Table 2.1.

Table 2.1: Single input-output case

DMU	A	B	C	D	E	F	G
Input (X)	3	4	10	7	5	9	8
Output (Y)	1	5	8	2	3	5	4

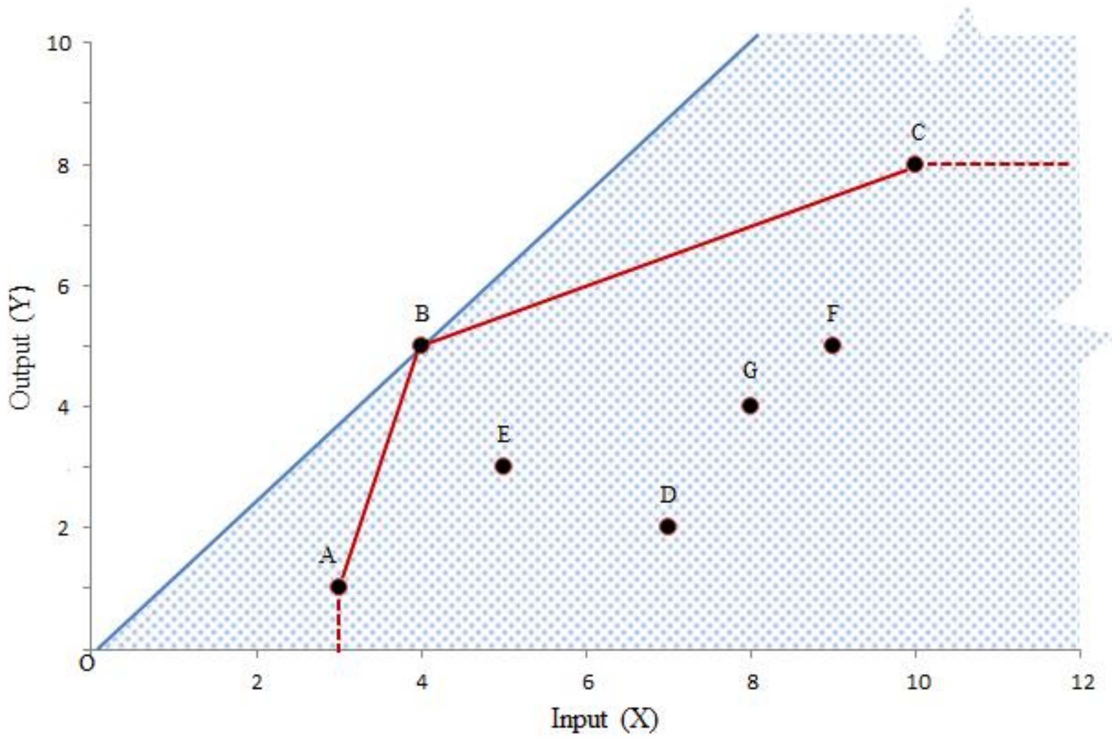


Fig. 2.2: Production possibility sets and efficient frontiers under CRS and VRS assumption

Under CRS assumption, the PPS that envelops all the data points (DMUs) is defined by the horizontal axis and the ray from the origin through DMU B. This ray constitutes the CRS efficient frontier, i.e. the efficient subset of the PPS. Under VRS assumption, the efficient frontier is determined by the DMUs A, B and C. The piecewise PPS envelops all the DMUs within the region bounded by the efficient frontier ABC, the horizontal line passing through C and the vertical line passing through A. The returns to scale are increasing along AB, constant to point B and decreasing along BC. As can be deduced, the assumption of the returns to scale affects the shape of the PPS and therefore the performance of the DMUs.

2.1.2 Efficiency measurement

The technical efficiency is a radial distance measure that derives for each evaluated DMU by means of a maximal feasible radial contraction of its input levels given its levels of outputs (input orientation). With an output orientation the distance is measured by means of a maximal feasible radial expansion of its output levels given its levels of inputs. Koopmans (1951) extended the optimality concept of Pareto optimality and provided a definition of technical efficiency, which is adapted in DEA:

Definition 2.1: A DMU is technically efficient if, and only if, an increase in any output or a decrease in any input is possible only by decreasing at least some other output or by increasing at least some other input.

2.1.2.1 Illustrative example: the CRS case

We revisit now the typical example of the seven DMUs (A to G) with the single input-output, as given in the previous section.

Table 2.2: Single input-output case

DMU	A	B	C	D	E	F	G
Input (X)	3	4	10	7	5	9	8
Output (Y)	1	5	8	2	3	5	4
Output/Input	0.333	1.25	0.8	0.286	0.6	0.556	0.5
CRS efficiency	0.267	1	0.64	0.229	0.48	0.444	0.4

The output/input ratio of each DMU is the slope of the ray passing through the origin of the axes and the DMU. The ray with the highest slope that passes from the origin through point B is the CRS efficient frontier. Therefore DMU B is efficient and defines the CRS efficient frontier. The other DMUs, namely A, C, D, E, F and G are inefficient as they lie below the efficient frontier. The output per input ratio for each DMU is presented in the third row of Table 2.2. The CRS technical efficiency of the DMUs is a relative measure that can be calculated by taking the output/input ratio with that of output/input ratio of the efficient DMU (B):

$$0 \leq \frac{output\ per\ input\ of\ DMU\ i}{output\ per\ input\ of\ DMU\ B} \leq 1, i = A, B, C, D, E, F, G$$

The last row of Table 2.2 presents the technical efficiency for each DMU. When input orientation is selected, the technical efficiency of the DMU E is derived by the ratio $\frac{\mathrm{E_1E_2}}{\mathrm{E_1E}} = 0.48$. When output orientation is selected, the technical efficiency is derived by the inverse of the ratio $\frac{\mathrm{E_3E_4}}{\mathrm{E_3E}} = 2.083$, i.e. the technical efficiency is $\frac{1}{2.083} = 0.48$. Hence, under the CRS assumption the input and output oriented measures provide the same efficiency score.

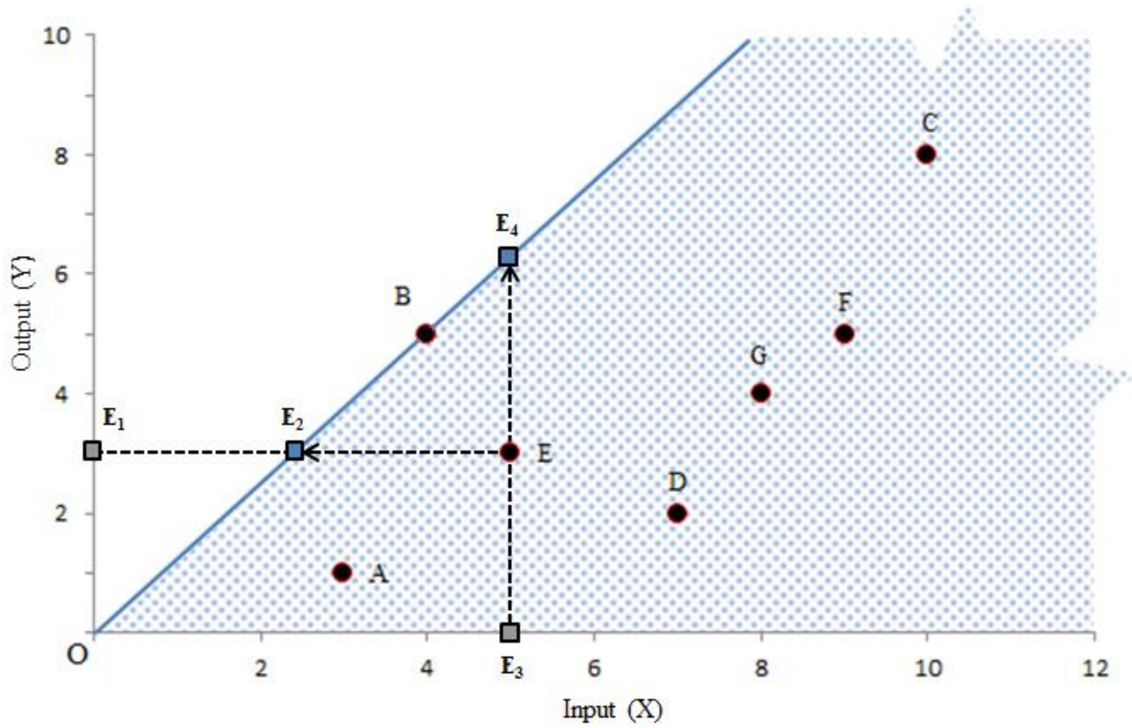


Fig. 2.3: CRS technical efficiency and projections of DMU E

The inefficient DMU E can be rendered efficient by a horizontal projection to the efficient frontier at point E_2 (2.4, 3), when input orientation is selected. This is accomplished by reducing its input level to the extent that the technical efficiency indicates, i.e. 0.48•5 = 2.4, while its output level is maintained. Alternatively, when output orientation is selected, the DMU E can be projected vertically to point E_4 (5, 6.25) of the efficient frontier,. This is achieved by increasing its output levels using the inverse of technical efficiency, i.e. 2.0833•3 = 6.25.

2.1.2.2 Illustrative example: the VRS case

When variable returns to scale are assumed the technical efficiency can be analyzed to two components, namely the VRS technical efficiency and the *"scale"* efficiency (SE). Banker et al (1984) showed that the CRS technical efficiency can be derived as the product of the VRS technical efficiency and the scale efficiency. The technical efficiency (TE) obtained under CRS and VRS assumptions, is called *global* and *pure* technical efficiency accordingly (cf. Cooper et al, 2007). The scale inefficiency is attributed to either decreasing or increasing returns to scale. A DMU is scale inefficient when it operates away from its *most productive scale size* (MPSS), see Banker (1984) and Banker et al (1984). The MPSS is a point on the CRS efficient frontier that maximizes the average productivity for its input-output bundle.

In Figure 2.4 below we illustrate the example of single input-output case under VRS assumption. The VRS efficient frontier is constituted by the DMUs A,B and C. As we can see, the CRS efficient DMU B is also VRS and scale efficient, i.e. it operates at the MPSS. On the other hand, the VRS efficient DMU A is scale inefficient because it operates on the IRS part of the VRS frontier. Analogously, the VRS efficient DMU C is scale inefficient because it operates on the DRS part of the VRS frontier. The inefficient DMU E can be rendered VRS efficient by projecting it to the points E_5 or E_6 onto the VRS efficient frontier, depending the orientation selected. Assuming an input orientation, the VRS technical efficiency and the scale efficiency can be determined by the ratios $\frac{E_1E_5}{E_1E}$ and $\frac{E_1E_2}{E_1E_5}$ respectively. The CRS technical efficiency for the DMU E is decomposed as:

$$CRS\,TE = VRS\,TE \times SE = \frac{\mathrm{E_1E_2}}{\mathrm{E_1E}} = \frac{E_1E_5}{E_1E} \times \frac{E_1E_2}{E_1E_5}$$

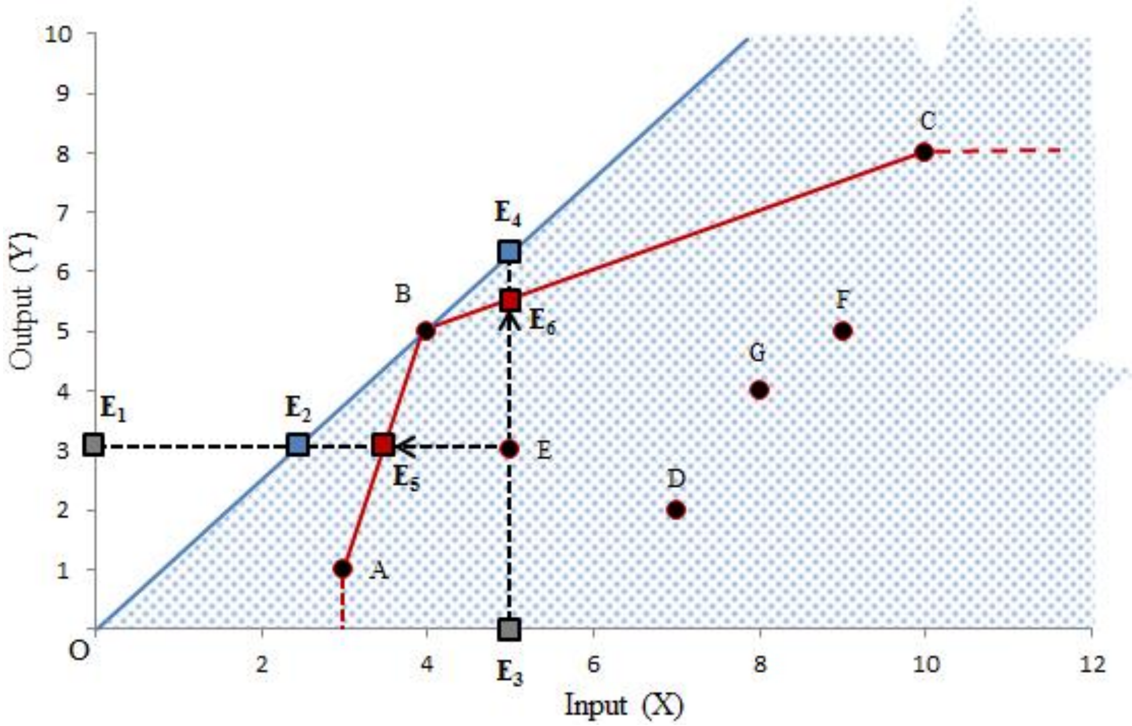


Fig. 2.4: Technical and scale efficiency for DMU E

Summing up, the technical inefficiency expresses for an inefficient DMU the failure to use the lowest possible level of input given its output or to produce the highest possible level of output volume given its input. The scale inefficiency is attributed to the size of operations of a DMU and represents deviation from the MPSS (CRS efficient frontier). We shall see in the next sections, how the aforementioned concepts can be accommodated in an operationally implementable form.

2.2 Basic DEA models

In this section we outline the four basic DEA models that originate from different assumptions on the way the inefficient units are projected on the efficient frontier (radial or non-radial). The radial efficiency measure assumes a proportional contraction (expansion) of inputs (outputs) towards the frontier. The radial measure is used in the two milestone DEA models, namely the CCR (Charnes et al, 1978) and the BCC (Banker et al, 1984) models. These models have become standards in the literature of performance measurement under the assumption of constant and variable returns to scale respectively. The non-radial efficiency measure differs from the radial efficiency measure in that the projections on the efficient frontier are made by non-proportional reductions (increases) of the input (output) levels.

2.2.1 The CCR model

The CCR model was introduced by Charnes et al (1978) to measure the relative efficiency of decision making units under the constant returns to scale assumption. The efficiency of a DMU is defined as the ratio of the weighted sum of outputs (total virtual output) to the weighted sum of inputs (total virtual input), with the weights being obtained in favor of each evaluated unit by the optimization process.

2.2.1.1 Multiplier form

Assume n DMUs, each using m inputs to produce s outputs. We denote by y_{rj} the level of the output $r(r=1,\ldots,s)$ produced by unit j ($j=1,\ldots,n$) and x_{ij} the level of the input i (i=1,…, m) used by unit j. The vector of inputs for DMU j is $X_j = (x_{1j},\ldots,x_{mj})^T$ and the vector of outputs is $Y_j = (y_{1j},\ldots,y_{rj})^T$. The basic CCR model that provides the CCR efficiency for the DMU$_{jo}$ is given below:

$$
\begin{aligned}
& max\ e_{j_0} = \frac{\omega Y_{j_0}}{\eta X_{j_0}} \\
& s.t. \\
& \frac{\omega Y_j}{\eta X_j} \leq 1, \ \ j = 1,\ldots,n \\
& \eta \geq 0, \omega \geq 0
\end{aligned}
\tag{2.3}
$$

The model (2.3) is formulated and solved for each DMU in order to obtain its efficiency score. The variables $\eta=(\eta_1,\ldots,\eta_m)$ and $\omega=(\omega_1,\ldots,\omega_s)$ are the weights associated with the inputs and the outputs respectively. These weights are calculated in a manner that they provide the highest possible efficiency score for the evaluated DMU$_{jo}$. The constraints of model (2.3) limit the efficiency scores of the DMUs in the interval (0, 1].

The CCR model (2.3) can be transformed to a linear program by applying the Charnes and Cooper (1962) transformation (C-C transformation hereafter). The transformation is carried out by considering a scalar $t \in \Re^+$ such as $t\eta X_{j_0} = 1$ and multiplying all terms of model (2.3) with $t>0$ so that $v = t\eta$, $u = t\omega$. The linear equivalent of model (2.3) is expressed as:

$$\begin{aligned}
&max\, e_{j_0} = uY_{j_0} \\
&s.t. \\
&vX_{j_0} = 1 \\
&uY_j - vX_j \leq 0,\ \ j = 1, \dots, n \\
&v \geq 0, u \geq 0
\end{aligned} \qquad (2.4)$$

The decision variables (v,u) of model (2.4) often are constrained to positive values by replacing the non-negativity constraints $v,u\geq 0$ with $v,u\geq\varepsilon$, where ε is a non-Archimedean infinitesimal. This is done in order to avoid giving a zero weight to some of the inputs and the outputs. Once an optimal solution v^*,u^* of model (2.4) is derived, the input oriented CCR-efficiency $e_{j_0}^*$ for DMU$_{jo}$ under evaluation is obtained directly from the objective function.

Definition 2.2 (CCR-Efficiency): The DMU$_{jo}$ is *CCR-efficient* if and only if $e_{j_0}^* = 1$ and there exists at least one optimal solution (v^*, u^*), with $u^*>0$ and $v^*>0$. Otherwise, the DMU$_{jo}$ is *CCR-inefficient.*

At the optimal solution of model (2.4) at least one constraint of the inequality constraints is binding (i.e. it holds as equality). The binding constraints that hold as equality ($u^*Y_j - v^*X_j = 0$) indicate the reference units for the evaluated unit.

2.2.1.2 Envelopment form

The envelopment form of the CCR model derives as the dual of the multiplier form (2.4) as follows:

$$\begin{aligned}
&min\, \theta \\
&s.t. \\
&\theta X_{j_0} - X\lambda \geq 0 \\
&Y\lambda - Y_{j_0} \geq 0 \\
&\lambda \geq 0
\end{aligned} \qquad (2.5)$$

The *envelopment* form (2.5) represents explicitly the way that the evaluated units are projected on the efficient frontier. The correspondence of the constraints and the variables of models (2.4) and (2.5) is presented in Table 2.3.

Table 2.3: Correspondence between models (2.4) and (2.5)

Constraints of multiplier model (2.4)	Variables of envelopment model (2.5)	Variables of multiplier model (2.4)	Constraints of envelopment model (2.5)
$vX_{j_0} = 1$	θ	$v \geq 0$	$\theta X_{j_0} - X\lambda \geq 0$
$uY_j - vX_j \leq 0, j = 1, \dots, n$	$\lambda \geq 0$	$u \geq 0$	$Y\lambda - Y_{j_0} \geq 0$

The constraints of the envelopment model (2.5) make evident its relation with the CRS production possibility set T^{CRS}(2.1). These constraints require the proposed activity (θX_{jo}, Y_{jo}) to belong to the T^{CRS}. The objective of the input oriented model (2.5) is to seek for the minimum θ that reduces radially the input levels (X_{jo}) to θX_{jo} onto the boundary of T^{CRS}, maintaining the output levels (Y_{jo}). In order to account for *input excesses* and *output shortfalls* the non-negative slack variables $s^- = (s_1^-, \dots, s_m^-)^T$ and $s^+ = (s_1^-, \dots, s_s^+)^T$ are introduced in model (2.5) to get its standard form as follows:

$$\begin{aligned} & min\, \theta \\ & s.t. \\ & \theta X_{j_0} - X\lambda - s^- = 0 \\ & Y_{j_0} - Y\lambda + s^+ = 0 \\ & \lambda \geq 0, s^- \geq 0, s^+ \geq 0 \end{aligned} \tag{2.6}$$

The model (2.6) is solved by the following two-phase LP procedure is solved:

Phase I

The model (2.6) is solved in order to derive the efficiency score θ^* for the evaluated DMU_{jo}.

Phase II

Using the optimal value θ^* from Phase I, the following model is solved to find a solution that maximizes the sum of input excesses and output shortfalls, where e=(1,…,1) is a vector of ones.

$$
\begin{aligned}
& max\, es^{-} + es^{+} \\
& s.t. \\
& X\lambda + s^{-} = \theta^{*} X_{j_0} \\
& Y\lambda - s^{+} = Y_{j_0} \\
& \lambda \geq 0, s^{-} \geq 0, s^{+} \geq 0
\end{aligned}
\tag{2.7}
$$

Definition 2.3: Given an optimal solution $(\theta^{*}, \lambda^{*}, s^{-*}, s^{+*})$ from Phases I and II, the evaluated DMU$_{jo}$ is CCR-inefficient if and only if θ^{*}=1 and all slacks are zero ($s^{-*} = 0, s^{+*}$=0). Otherwise, the DMU$_{jo}$ is CCR-inefficient.

The Definition 2.2 characterizes the efficient units in terms of the multiplier form whereas the Definition 2.3 characterizes the efficient units in terms of the envelopment form. These two definitions are equivalent according to complementary slackness theorem. Accordingly, the Definition 2.3 is equivalent to Pareto-Koopmans efficiency (Definition 2.1).

Given the optimal solution $(\theta^{*}, \lambda^{*}, s^{-*}, s^{+*})$ derived by Phases I and II (models (2.6) and (2.7)) the projections of the evaluated unit j_o $(\hat{X}_{j_0}, \hat{Y}_{j_0})$ on the efficient frontier are obtained as follows:

$$
\begin{aligned}
& \hat{X}_{j_0} = \theta^{*} X_{j_0} - s^{-*} = X_j \lambda^{*} \\
& \hat{Y}_{j_0} = Y_{j_0} + s^{+*} = Y_j \lambda^{*}
\end{aligned}
\tag{2.8}
$$

2.2.1.3 Input-Output oriented models

The technical efficiency of a DMU can be measured by adopting either an input or an output orientation. If the conservation of inputs is of greater importance an input orientation is selected. Alternatively, if the expansion of outputs is considered more important according to the analyst's perspective an output orientation is selected. The two variants of the CCR model in multiplier and envelopment forms are presented in Table 2.4 below.

Table 2.4: Input and output oriented CCR models

	Multiplier form		Envelopment form	
Input oriented	$max\, e_{j_0} = uY_{j_0}$ $s.t.$ $vX_{j_0} = 1$ $uY_j - vX_j \leq 0,\ \ j = 1, \dots, n$ $v \geq 0, u \geq 0$	(2.4)	$min\, \theta$ $s.t.$ $\theta X_{j_0} - X\lambda - s^- = 0$ $Y_{j_0} - Y\lambda + s^+ = 0$ $\lambda \geq 0, s^- \geq 0, s^+ \geq 0$	(2.6)
Output oriented	$min \frac{1}{e_{j_0}} = vX_{j_0}$ $s.t.$ $uY_{j_0} = 1$ $uY_j - vX_j \leq 0,\ \ j = 1, \dots, n$ $v \geq 0, u \geq 0$	(2.9)	$max\, \varphi$ $s.t.$ $X_{j_0} - X\mu - \tau^- = 0$ $\varphi Y_{j_0} - Y\mu + \tau^+ = 0$ $\mu \geq 0, \tau^- \geq 0, \tau^+ \geq 0$	(2.10)

The optimal solutions of the input and output oriented models are related as follows:

- $\theta^* = 1/\varphi^*$
- $\lambda_j^* = \mu_j^*/\varphi^*, j = 1, \dots, n$
- $s^{-*} = \tau^{-*}/\varphi^*, i = 1, \dots, m$
- $s^{+*} = \tau^{+*}/\varphi^*, r = 1, \dots, s$

(2.11)

2.2.1.4 Illustrative example

We illustrate the above models and concepts with an example of seven DMUs (A to G) that use a single input to produce two outputs as shown in Table 2.5 below:

Table 2.5: Example of seven DMUs with single input and two outputs

DMU	A	B	C	D	E	F	G
Input (X^1)	2	2	10	4	4	1	1
Output (Y^1)	4	6	30	4	12	1	4
Output (Y^2)	14	6	10	20	24	7	2
Output (Y^1) / Input (X^1)	2	3	3	1	3	1	4
Output (Y^1) / Input (X^1)	7	3	1	5	6	7	2

To make the representation of the units in two dimensional space possible the outputs per unit of input are calculated. Table 2.6 exhibits the results obtained from the multiplier model (2.9) or its dual, the envelopment model (2.10), i.e. the inverse of the efficiency scores (φ^*), the optimal weights (v_1^*, u_1^*, u_2^*) in terms of the multiplier model, the reference set of the inefficient units with the intensities and the input excesses (τ_1^{-*}) and the output shortfalls τ_1^{+*}, τ_2^{+*}) .

Table 2.6: CCR output oriented results

DMU	φ^*	v_1^*	u_1^*	u_2^*	Reference set	Intensity variables	τ_1^{-*}	τ_1^{+*}	τ_2^{+*}
A	1	1	0.111	0.111	-	$\mu_A^* = 1$	0	0	0
B	1.2	1.2	0.267	0.067	E, G	$\mu_E^* = 0.4$, $\mu_G^* = 0.6$	0	0	0
C	1.333	1.333	0.333	0	G	$\mu_G^* = 1$	0	0	0.667
D	1.4	1.4	0	0.2	A	$\mu_A^* = 1$	0	0.6	0
E	1	1	0.111	0.111	-	$\mu_E^* = 1$	0	0	0
F	1	1	0	0.143	A	$\mu_A^* = 1$	0	1	0
G	1	1	0.222	0.056	-	$\mu_G^* = 1$	0	0	0

The DMUs A, E and G are efficient, while the other DMUs are inefficient. Figure 2.5 below depicts the production possibility set (shadowed region), the efficient frontier defined by the efficient units A, E, and G and the projections of the inefficient DMUs on the frontier.

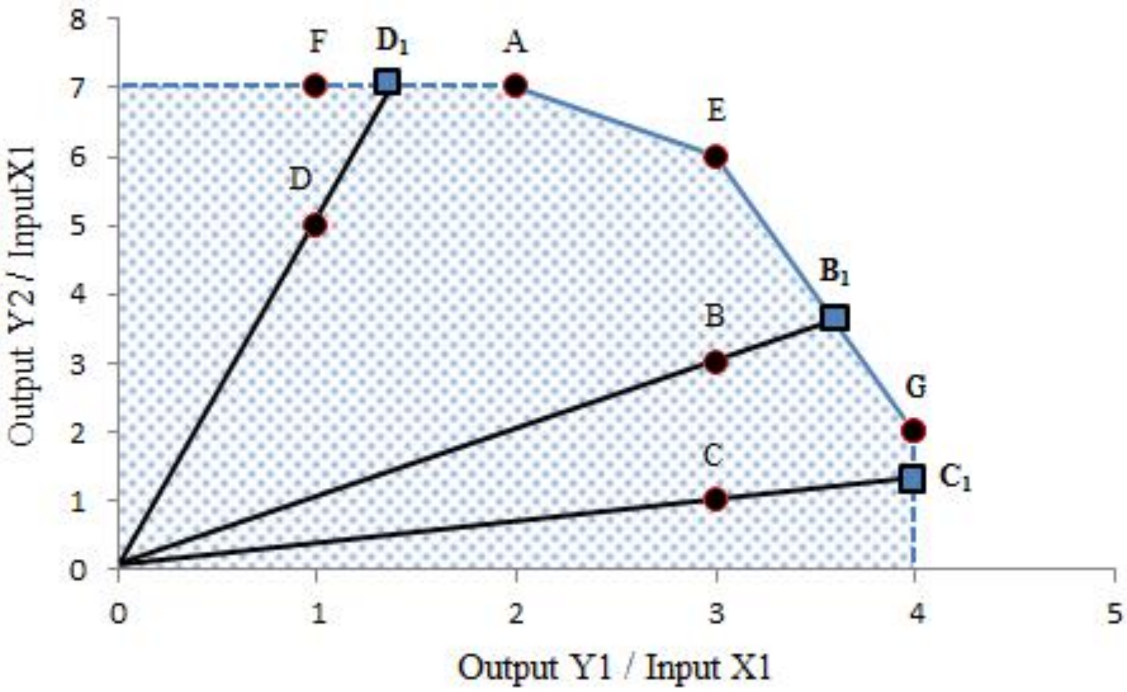


Fig. 2.5: Production possibility set, efficient frontier and projections

The output oriented projections of the inefficient units are obtained from the following relations:

$$\hat{X}_{j_0} = X_{j_0} - \tau^{-*} = X_j \mu_j^*$$
$$\hat{Y}_{j_0} = \varphi^* Y_{j_0} + \tau^{+*} = Y_j \mu_j^* \qquad (2.12)$$

The DMU F is inefficient because although φ^*=1 there is an output short fall ($\tau_1^{+*} = 1$) with respect with output Y^1. Hence, DMU F can be rendered efficient by increasing the output Y_F^1 indicated by the amount $\tau_1^{+*} = 1$. Graphically, this is accomplished by moving from point F to point A. Using the formulations (2.12) we derive the improvements ($\hat{X}_F^1, \hat{Y}_F^1, \hat{Y}_F^2$) for the DMU F as follows:

$$\hat{X}_F^1 = X_F^1 - \tau_1^{-*} = 1 - 0 = 1 \ \ or \ \ \hat{X}_F^1 = \mu_A^* X_A^1 = 1 \cdot 1 = 1$$

$$\hat{Y}_F^1 = \varphi^* Y_F^1 + \tau_1^{+*} = 1 \cdot 1 + 1 = 2 \ \ or \ \ \hat{Y}_F^1 = \mu_A^* Y_A^1 = 1 \cdot 2 = 2$$

$$\hat{Y}_F^2 = \varphi^* Y_F^2 + \tau_2^{+*} = 1 \cdot 7 + 0 = 7 \ \ or \ \ \hat{Y}_F^2 = \mu_A^* Y_A^2 = 1 \cdot 7 = 7$$

The DMU D is inefficient as φ^*=1.4>1. Its projection D1 on the frontier can be derived by increasing proportionally the two outputs by φ^*=1.4, however this radial improvement is not sufficient to restore the efficiency of unit D as there is an extra improvement (non-radial) of output Y_D^1 by $\tau_1^{+*} = 0.6$. Therefore, DMU D has to move first to point D_1 and then on point A of the efficient frontier. The projections ($\hat{X}_D^1, \hat{Y}_D^1, \hat{Y}_D^2$) for the DMU D are calculated as follows:

$$\hat{X}_D^1 = X_D^1 - \tau_1^{-*} = 1 - 0 = 1 \ \ or \ \ \hat{X}_D^1 = \mu_A^* X_A^1 = 1 \cdot 1 = 1$$

$$\hat{Y}_D^1 = \varphi^* Y_D^1 + \tau_1^{+*} = 1.4 \cdot 1 + 0.6 = 2 \ \ or \ \hat{Y}_D^1 = \mu_A^* Y_A^1 = 1 \cdot 2 = 2$$

$$\hat{Y}_D^2 = \varphi^* Y_D^2 + \tau_2^{+*} = 1.4 \cdot 5 + 0 = 7 \ \ or \ \ \hat{Y}_D^2 = \mu_A^* Y_A^2 = 1 \cdot 7 = 7$$

The projections of DMUs B and C are obtained accordingly:

$$\hat{X}_B^1 = X_B^1 - \tau_1^{-*} = 1 - 0 = 1 \;\; or \;\; \hat{X}_B^1 = \mu_E^* X_E^1 + \mu_G^* X_G^1 = 0.4 \cdot 1 + 0.6 \cdot 1 = 1$$

$$\hat{Y}_B^1 = \varphi^* Y_B^1 + \tau_1^{+*} = 1.2 \cdot 3 + 0 = 3.6 \;\; or \;\; \hat{Y}_B^1 = \mu_E^* Y_E^1 + \mu_G^* Y_G^1 = 0.4 \cdot 3 + 0.6 \cdot 4 = 3.6$$

$$\hat{Y}_B^2 = \varphi^* Y_B^2 + \tau_2^{+*} = 1.2 \cdot 3 + 0 = 3.6 \;\; or \;\; \hat{Y}_B^2 = \mu_E^* Y_E^2 + \mu_G^* Y_G^2 = 0.4 \cdot 6 + 0.6 \cdot 2 = 3.6$$

$$\hat{X}_C^1 = X_C^1 - \tau_1^{-*} = 1 - 0 = 1 \;\; or \;\; \hat{X}_C^1 = \mu_G^* X_G^1 = 1 \cdot 1 = 1$$

$$\hat{Y}_C^1 = \varphi^* Y_C^1 + \tau_1^{+*} = 1.333 \cdot 3 + 0 = 4 \;\; or \;\; \hat{Y}_C^1 = \mu_G^* Y_G^1 = 1 \cdot 4 = 4$$

$$\hat{Y}_C^2 = \varphi^* Y_C^2 + \tau_2^{+*} = 1.333 \cdot 1 + 0.667 = 2 \;\; or \;\; \hat{Y}_C^2 = \mu_G^* Y_G^2 = 1 \cdot 2 = 2$$

In Table 2.7 below, we provide the projections for the inefficient DMUs.

Table 2.7: Projections

DMU	B	C	D	F
Input $\hat{X}^1$	1	1	1	1
Output $\hat{Y}^1$	3.6	4	2	2
Output $\hat{Y}^2$	3.6	2	7	7

2.2.2 The BCC model

The CCR model was extended to the BCC model by Banker et al (1984) so as to accommodate variable returns to scale (VRS). The incorporation of variable returns to scale in DEA allows for decomposing the overall efficiency to technical and the scale efficiency in contrast to the CCR model which aggregates them into a single measure.

A structural difference of the CCR and the BCC models is the additional free of sign variable u_o in the multiplier form of the latter, which is the dual variable associated with the additional convexity constraint (eλ=1) in the envelopment form. Table 2.8 exhibits the multiplier and the envelopment forms of the input oriented and the output oriented BCC models.

Table 2.8: Input and output oriented BCC models

	Multiplier form		Envelopment form	
Input oriented	$max\, uY_{j_0} - u_0$ $s.t.$ $vX_{j_0} = 1$ $uY_j - u_0 - vX_j \leq 0, \ \ j = 1, \dots, n$ $v \geq 0, u \geq 0$	(2.13)	$min\, \theta$ $s.t.$ $\theta X_{j_0} - X\lambda - s^- = 0$ $Y_{j_0} - Y\lambda + s^+ = 0$ $e\lambda = 1$ $\lambda \geq 0, s^- \geq 0, s^+ \geq 0$	(2.14)
Output oriented	$min\, vX_{j_0} - u_0$ $s.t.$ $uY_{j_0} = 1$ $uY_j - vX_j + u_0 \leq 0, \ \ j = 1, \dots, n$ $v \geq 0, u \geq 0$	(2.15)	$max\, \varphi$ $s.t.$ $X_{j_0} - X\mu - \tau^- = 0$ $\varphi Y_{j_0} - Y\mu + \tau^+ = 0$ $e\lambda = 1$ $\lambda \geq 0, \tau^- \geq 0, \tau^+ \geq 0$	(2.16)

We give in Table 2.9 the primal and dual correspondences of the constraints and the variables of the BCC input oriented models.

Table 2.9: Correspondence between models (2.13) and (2.14)

Constraints of multiplier model (2.4)	Variables of envelopment model (2.5)	Variables of multiplier model (2.4)	Constraints of envelopment model (2.5)
$vX_{j_0} = 1$	θ	$v \geq 0$	$\theta X_{j_0} - X\lambda \geq 0$
$uY_j - vX_j - eu_o \leq 0, j = 1, \dots, n$	$\lambda \geq 0$	$u \geq 0$	$Y\lambda \geq Y_{j_0}$
		u_o	$e\lambda=1$

The definition of efficiency is analogous to the Definitions 2.2 and 2.3. The efficiency scores obtained from the BCC model are not less than the corresponding ones obtained from the CCR model. The boundary of the VRS production possibility set is closer to the observations (DMUs) since, the convexity constraint imposed in the BCC envelopment model spans a subset of the CRS production possibility set (see Fig. 2.2).

The prevailing returns to scale can be identified by the optimal solution of both multiplier and envelopment BCC models. Banker et al (1984) determined the returns to scale (RTS) using the optimal value of the free variable in the multiplier models. Banker (1984) and Banker and Thrall (1992) identified the RTS by the intensity variables (lambdas) of the envelopment models.

Theorem: Given the point (x_0, y_0) that lies on the efficient frontier, the returns to scale at this point are identified by the following conditions:

1. Increasing returns to scale (IRS) prevail at (x_0, y_0) if and only if $u_o^* < 0$ for all optimal solutions.
2. Decreasing returns to scale (DRS) prevail at (x_0, y_0) if and only if $u_o^* > 0$ for all optimal solutions.
3. Constant returns to scale (CRS) prevail at (x_0, y_0) if and only if $u_o^* = 0$ in any optimal solutions.

We point out that the input and output oriented BCC models may determine different returns to scale, i.e. the results may depend on the orientation adopted (cf. Banker et al, 2004). For the same DMU, the BCC input oriented may identify IRS while the BCC output oriented may identify DRS. This is attributed to the nature of the VRS technology (see Figures 2.2 and 2.4) and the projections of the evaluated unit onto the VRS frontier. In the example portrayed in Figure 2.4, concerning the inefficient DMU E, IRS prevail at the frontier point E_5, which is obtained by the input oriented projection, while DRS prevail at the frontier point E_6, which is derived by its output oriented projection.

2.2.3 The additive model

A natural evolution of the CCR and the BCC models is the *additive* model proposed by Charnes et al (1985). The CCR and BCC models are based on input or output orientation and they provide efficiency scores via a radial measure. On the contrary, in the additive model both orientations are combined, thus it is non-oriented. The additive model deals directly with the input excesses and output shortfalls and, in contrast to the CCR and the BCC models, it does not provide a direct measure of efficiency, but it distinguishes only among efficient and inefficient units. Table 2.10 below exhibits the multiplier and the envelopment forms of the additive model under both CRS and VRS assumptions.

Table 2.10: Additive DEA models

	Multiplier form		Envelopment form	
CRS assumption	$\min vX_{j_0} - uY_{j_0}$ $s.t.$ $uY_j - vX_j \leq 0, \;\; j = 1, \dots, n$ $v \geq 1, u \geq 1$	(2.17)	$\max es^- + es^+$ $s.t.$ $X\lambda + s^- = X_{j_0}$ $Y\lambda - s^+ = Y_{j_0}$ $\lambda \geq 0, s^- \geq 0, s^+ \geq 0$	(2.18)
VRS assumption	$\min vX_{j_0} - uY_{j_0} + u_o$ $s.t.$ $uY_j - u_o - vX_j \leq 0, \;\; j = 1, \dots, n$ $v \geq 1, u \geq 1$	(2.19)	$\max es^- + es^+$ $s.t.$ $X\lambda + s^- = X_{j_0}$ $Y\lambda - s^+ = Y_{j_0}$ $e\lambda = 1$ $\lambda \geq 0, s^- \geq 0, s^+ \geq 0$	(2.20)

A DMU is efficient according to the additive model if and only if at the optimal solution all slacks are zero ($s^{-*} = 0, s^{+*} = 0$), i.e. there are not any input excesses or output shortfalls. Under the CRS assumption, a DMU is deemed efficient by the additive model if and only if is deemed efficient by the CCR model. Also, under VRS assumption, a DMU is deemed efficient by the additive model if and only if is deemed efficient by the BCC model.

2.2.4 The Slacks-Based Measure (SBM)

A limitation of the additive model is that only discriminates the efficient from the inefficient DMUs without providing an efficiency measure for the evaluated units. Tone (2001) introduced the slacks based measure (SBM) which is a successor of the additive model that is capable of providing the efficiency scores of the evaluated units. Tone (2001) built upon the additive model and introduced a non-radial measure of efficiency that takes into account the input excesses and the output shortfalls. The fractional form of the SBM model is as follows:

$$
\begin{aligned}
& min \frac{1-\frac{1}{m}\sum_{i=1}^{m} s_i^- / x_{ij_o}}{1+\frac{1}{s}\sum_{r=1}^{s} s_r^+ / y_{rj_o}} \\
& s.t. \\
& X\lambda + s^- = X_{j_0} \\
& Y\lambda - s^+ = Y_{j_0} \\
& \lambda \geq 0, s^- \geq 0, s^+ \geq 0
\end{aligned}
\tag{2.21}
$$

The model (2.21) can be converted to the following non-linear program by multiplying both terms of the objective function by a positive scalar variable t such that $t + \left(\frac{1}{s}\sum_{r=1}^{s} ts_r^+ / y_{rj_o}\right) = 1$.

$$
\begin{aligned}
& min\, t - \frac{1}{m}\sum_{i=1}^{m} ts_i^- / x_{ij_o} \\
& s.t. \\
& t + \left(\frac{1}{s}\sum_{r=1}^{s} ts_r^+ / y_{rj_o}\right) = 1 \\
& X\lambda + s^- = X_{j_0} \\
& Y\lambda - s^+ = Y_{j_0} \\
& \lambda \geq 0, s^- \geq 0, s^+ \geq 0, t > 0
\end{aligned}
\tag{2.22}
$$

Then model (2.22) can be transformed to a linear equivalent by applying the following variable transformation $\mu = t\lambda$, $\sigma^- = s^-$ and $\sigma^+ = ts^+$.

$$
\begin{aligned}
& \min \rho = t - \frac{1}{m}\sum_{i=1}^{m} \sigma_i^- / x_{ij_o} \\
& s.t. \\
& t + \left(\frac{1}{s}\sum_{r=1}^{s} \sigma_r^+ / y_{rj_o}\right) = 1 \\
& X\mu + \sigma^- = tX_{j_0} \\
& Y\mu - \sigma^+ = tY_{j_0} \\
& \mu \geq 0, \sigma^- \geq 0, \sigma^+ \geq 0, t > 0
\end{aligned}
\tag{2.23}
$$

Note that t>0 by virtue of the first constraint of model (2.23). A DMU is efficient if the optimal value of the objective function is one ($\rho^*=1$), which indicates that there are no input excesses or output shortfalls ($\sigma^{-*}=0$ and $\sigma^{+*}=0$).

Conclusion

Data Envelopment Analysis enabled the efficiency assessment of units that use multiple inputs to produce multiple outputs, without the need of a priori knowledge of the production function. Also DEA requires very few assumptions, therefore it opened up the possibilities for use in a wide range of applications. Great attention has been paid to DEA from the research community as well as from operations analysts, because of its practical usefulness on providing performance measures and handling effectively the sources of inefficiency.

However, the further development of DEA is needed so as to address the new real world problems and the emerging issues caused by the growing complexity of the production conditions. For instance the DMUs may consist of sub-processes with a complex internal structure. Despite the exquisite properties of the DEA models presented in this chapter, these models do not take into account the possible internal structures of the DMUs and they carry out the assessments of the DMUs as single stage processes. Thus, they cannot provide meaningful results when applied to multi-stage processes as we will see in the following chapters.

Chapter 3

Review of Network DEA methods

Network DEA is an extension of conventional DEA developed to take account of the internal structure of DMUs. The DMUs often consist of sub-units that interact and perform various operations. The traditional DEA models treat the DMUs as single stage production processes that transform some external inputs to final outputs. In such a setting, the internal structure of the DMUs is not taken into consideration and the linking activities are neglected. Cook and Zhu (2014) stressed out that in conventional DEA the DMU is treated as a *black box* and its internal structure and operations are ignored. In addition, Kao (2017) pointed out that "*when interactions among divisions are not taken into account, the results will be distorted and misleading*". Kao and Hwang (2008) showed that the standard DEA models may deem a DMU overall system efficient even though all their sub-units are inefficient. Conclusively, the standard DEA models fail to adequately capture and mathematically represent the aforementioned characteristics of the DMUs. Also, they fall short to shed light on the sources of inefficiency as well as to provide succinct guidance for the improvement of the inefficient DMUs and sub-units. On the other hand, in network DEA, the DMU is considered as a network of interconnected sub-units, with the connections indicating the flow of intermediate products (commonly called intermediate measures or links). In the literature, these sub-units are also known as stages, divisions, sub-DMUs, sub-systems, sub-processes, processes, procedures, components and functions. Albeit in this thesis we may use these terms interchangeably, we mainly adopt the term "stage" when we refer to the sub-units of the DMUs. The advantage of the network DEA models is their ability to reflect accurately the DMUs' internal operations as well as to incorporate their relationships and interdependences. Therefore, they yield more representative and precise results than the conventional DEA models and provide more information regarding the sources that cause inefficiency. Cook and Seiford (2009) included the network DEA models to the methodological developments of DEA and mentioned that these models allow the detailed examination of the inner workings of a production process, which leads to a greater understanding of that process. Indeed, having a full picture of the internal structure of DMUs and examining their sub-units in a coordinated manner will provide further insights for the performance assessment and will assist better the decision making.

Fare and Primont (1984) and Charnes et al (1986) were the first studies, to the best of our knowledge, in the field later named network DEA by Fare and Grosskopf (2000). Fare and Primont (1984) distinguished the internal structure of multi-plant firms, i.e. firms that own many plants. They defined the firm's technology by constructing first the technology of each plant. Their approach was applied to a selected sample of electric generating firms which consists of nineteen plants in Illinois. Later, Kao (1998) applied their approach for the efficiency assessment of eight Taiwanese forest districts with 34 working circles. The performance of each working circle was measured based on the technology constructed from all of them. Charnes et al (1986) assumed that the US army recruitment is comprised of two processes, namely the awareness creation and the contract establishment. Charnes et al (1994) characterized the two-stage process assumed in Charnes et al (1986) as a *stratified* production process where a *hybrid* modelling approach was used. The work of Fare and Grosskopf (2000) is considered pioneering in the field of network DEA. Although the terms "*black box*" and *network technology* had been earlier reported in the studies of Fare (1991), Fare and Whittaker (1995), Fare and Grosskopf (1996) and Fare et al (1997), it was Fare and Grosskopf (2000) who first coined the term *network DEA* and provided a consolidated framework of the aforementioned studies for multi-stage processes with various structures.

Network DEA has already attracted the interest of researchers and a significant body of research is devoted to both theory and applications. Kao (2014b) noticed that the number of publications before 2000 was two or three per year, thereafter though it has rapidly grown. Liu et al (2013), in their citation-based literature survey for DEA for the period 1978–2010, considered the field of network DEA as a subarea which is relatively active in recent years. However, from 2010 onwards there has been a blast of publications on network DEA. Some of these studies explore the properties of the existing methods while others apply them to real world problems. The application field of network DEA as we will also see below is very wide, e.g. supply chains, banking, education, sports just to mention some. The network DEA methods can now be straightforwardly and effectively employed for the performance evaluation of a supply chain and its members which is undoubtedly a rough task. Agrell and Hatami-Marbini (2013) provided a thorough review for network DEA methods including studies devoted to supply chain performance analysis. They also remarked that the supply chains are complex multi-stage systems with interrelations, which use multiple inputs to produce multiple outputs. Hence, the network DEA methods can be adequately employed for their performance assessment. Many prominent approaches are developed to deal with the variety of the structures, the interdependencies and the conflicting interests of the sub-units. Most of the network DEA studies are dedicated to the performance assessment of DMUs with a specific internal structure. A DMU may consist of sub-units arranged in series, in parallel or a mixture of these. Cook et al (2010a) and Chen et al (2013) provide insights and directions for further research for the two-stage network structures arranged in series. Castelli et al (2010) and Halkos et al (2014) provided comprehensive categorized overviews of models and methods developed for different multi-stage production architectures. Kao (2014b) provides an excellent review and classification of network DEA methods according to the model they use and the network structure that they examine. Moreover, a collection of network DEA methods is given in Cook and Zhu (2014).

The aim of this chapter is to describe the underlying notions of network DEA, to present the state of the art in the field and to review the most significant network DEA methods. Also, our goal is to provide a comprehensive insight and categorization of the network DEA literature in a unified manner. In particular, we present the possible network structures that a DMU may be characterized of, we demonstrate the advantages of the network DEA over the standard DEA, we provide a critical review of the most influential approaches and we discuss their extensions, inherent limitations and shortcomings. In addition, we track the majority of multi-stage DEA applications and we classify them according to the method they utilize. Hence, this chapter presents a complete survey of the network DEA literature.

3.1 Network structures and assessment paradigms

The DMUs may have various types of internal structures. However, we discern that their production processes may be arranged either in series, in parallel or in series-parallel. The series and the parallel production processes are two distinctive network architectures studied extensively in the literature. In this section, we provide some illustrative examples of network structures which are used as the basis for the development of the most significant network DEA methods. The four types of series two-stage network processes depicted in Figure 3.1 are the basis for the development of network DEA theory and applications.

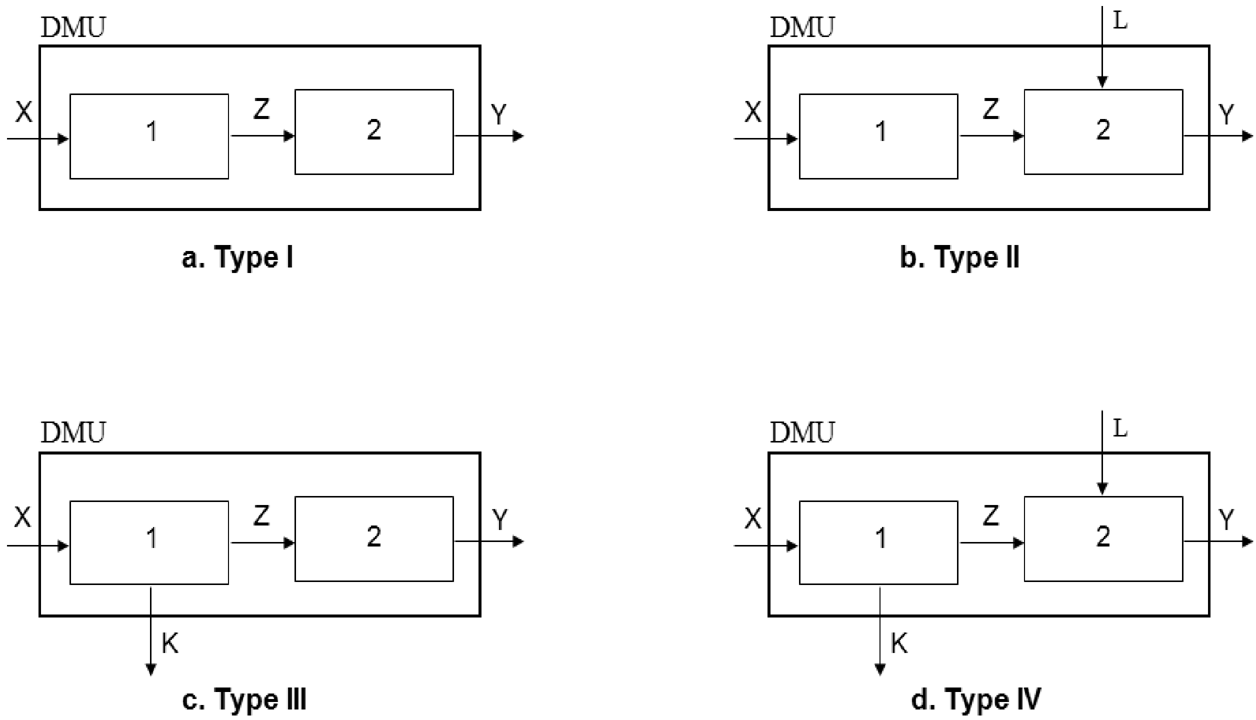


Fig. 3.1: The four types of series two-stage processes

In the Type I two-stage process (Fig. 3.1a) the first stage uses external inputs (*X*) to produce the intermediate measures (*Z*), which are subsequently used as inputs to the second stage which produces the final outputs (*Y*). In Type I structure, nothing but the external inputs to the first stage enters the system and nothing but the outputs of the second stage leaves the system. This is the elementary network structure that has drawn the attention of most of the research work. Wang et al (1997) and Seiford and Zhu (1999) are the first who studied processes of Type I structure.

In the production process of Type II (Fig. 3.1b) each DMU uses the external inputs (*X*) in the first stage to yield the intermediate measures (*Z*), which then are used along with the additional external inputs (*L*) to the second stage to yield the final outputs (*Y*), as depicted in Figure 3.1b. That is, the second stage uses except from the intermediate measures additional external inputs (*L*) for exclusive use. Liang et al (2006), under game theoretic concepts, studied a supply chain with two stages, the seller and the buyer, where the buyer (second stage) uses extra inputs. Notice that the Type II structure may be varied by assuming that the external inputs (*X*) can be freely shared between the stages in conjunction with or without the additional inputs (*L*). Such a variation is considered in Chen et al (2006), where the impact of the Information Technology (IT) on firm performance is examined.

In the production process of Type III (Fig. 3.1c), the first stage produces some final outputs (*K*) beyond the intermediate measures (*Z*).

In the production process of Type IV (Fig. 3.1d) external inputs and final outputs appear in both stages. The first stage uses the inputs (*X*) to generate the final outputs (*K*) and the intermediate measures (*Z*). The second stage uses the intermediate measures (*Z*) and the additional external inputs (*L*) for the production of the final outputs (*Y*). This type of network structure was first studied in Charnes et al (1986), Fare (1991), Fare and Whittaker (1995) and Fare and Grosskopf (1996).

The four types of network structures portrayed in Figure 3.1 can be generalized to series structures with more than two stages. Figure 3.2 exhibits a series network structure where each DMU is considered as a multi-stage process with *v* stages. Actually, the general series multi-stage process depicted in Figure 3.2 is the multiple of Type IV.

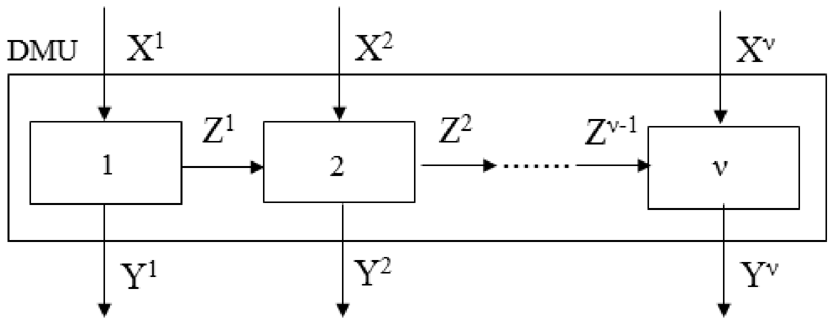


Fig. 3.2: A general multi-stage series process

Another basic network structure of the DMUs is a production process whose sub-processes are configured in parallel. Figure 3.3 below depicts the internal structure of a DMU with *v* parallel processes without interdependencies. Each sub-process transforms the external inputs (*X*) to final outputs (*Y*). A characteristic example of units that can be considered as parallel production processes are the academic departments, where teaching and research are two separate functions with specific resources and outputs. Analogously, a university can be viewed as a DMU and its departments can be regarded as the individual parallel sub-units. A modification of the parallel structure of Figure 3.3 may involve shared flows among the stages, i.e. the sub-processes, instead of consuming dedicated inputs, they share common resources (external inputs). Such a case was examined by Fare et al (1997) who assessed the performance of 57 grain farms with one shared input, namely the land. The land is allocated among four different agricultural operations, specifically the crops of corn, soybeans, wheat and the double crop soybeans. Based on Fare et al (1997), Vaz et al (2010) studied 78 Portuguese retail stores, each one comprised of five sections, namely groceries, perishables, light bazaar, heavy bazaar and textiles. These sections operate in parallel and share the floor area. In the educational sector, Beasley (1995) and Mar Molinero (1996) developed nonlinear models to measure teaching and research performance as parallel academic operations, in the presence of two shared inputs, namely general and equipment expenditures.

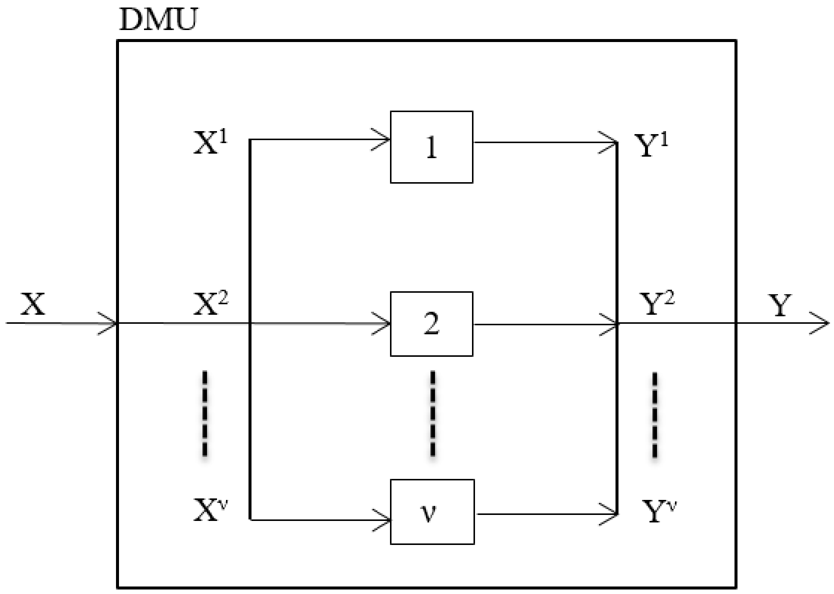


Fig. 3.3: A production process with parallel sub-processes

The basic series and parallel configurations are not always sufficient to describe real world situations. Therefore, more complex network structures, mixtures of aforementioned ones can be used to represent in detail the relationships among the sub-processes. Figure 3.4 below portrays for example such a network structure composed of a combination of series and parallel structures.

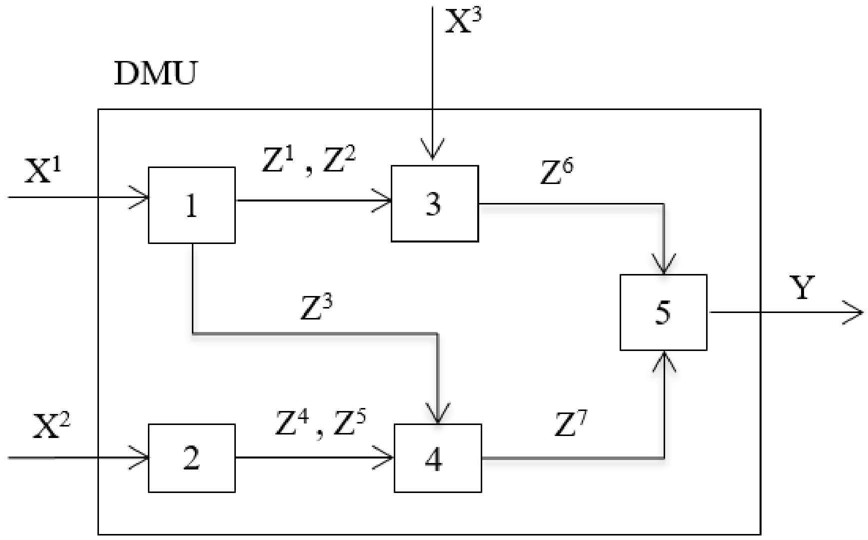


Fig. 3.4: Mixed network structure with series and parallel sub-processes

Such a structure was studied by Lewis and Sexton (2004) who assessed the performance of 30 teams (DMUs) of the Major League Baseball in the US. They assumed that the operation of each team is represented by a network of five distinct sub-processes. The first and the second correspond to the team's front office operations, while the other three correspond to the team's on-field operations.

The performance assessment of DMUs within the network DEA framework is carried out by a variety of approaches, which can be categorized into two assessment paradigms, namely, the *independent assessment paradigm* and the *joint assessment paradigm*. In the *independent assessment paradigm* the standard DEA models are used to assess the performance of the DMUs and the sub-processes independently. In the *joint assessment paradigm* the internal structure of the DMUs and the interdependencies among the sub-processes are taken into account, also the efficiency assessment of the sub-processes and the whole system is made simultaneously. There are three general approaches that follow the *joint assessment paradigm*: the *efficiency decomposition approach*, the *slacks-based measure approach* and the *system-centric approach*. The categorization is based on the perspective of each approach about the relationships between the system (DMU) and the stage efficiencies as well as on the kind of information provided for the performance of the individual stages and the system. The interaction between the sub-processes is taken into account by these approaches, however their difference lies on the way that the overall and the stage efficiencies are derived. In particular, the efficiency decomposition approach measures the system efficiency first and then the stage efficiencies are calculated ex post. In the slacks-based measure approach the overall efficiency is obtained from the aggregation of the stage efficiencies with various ways as we will see further on. As system-centric we characterize the network DEA methods that take into account the internal structure of the DMUs and the interdependencies among the stages, but they provide only an overall performance measure without generating the stage efficiencies. In the system-centric methods there is no functional form that connects the overall and the stage efficiencies.

We give below the notation that will be employed in the current chapter:

$j \in J = \{1, \dots, n\}$: The index set of the *n* DMUs.

$j_0 \in J$: Denotes the evaluated DMU.

$\delta \in \Delta = \{1, \dots, \nu\}$: The index set of the ν processes that each DMU_j is composed.

$X_j = \left(x_{ij}, i = 1, \dots, m\right)$: The vector of external inputs used by DMU_j.

$Z_j = \left(z_{pj}, p = 1, \dots, q\right)$: The vector of intermediate measures for DMU_j.

$Y_j = \left(y_{sj}, r = 1, \dots, s\right)$: The vector of final outputs produced by DMU_j.

$L_j = \left(l_{dj}, d = 1, \dots, a\right)$: The vector of extra inputs in structures of Type I and IV.

$K_j = \left(k_{cj}, c = 1, \dots, b\right)$: The vector of extra outputs in structures of Types III and IV.

$\eta = (\eta_1, \dots, \eta_m)$: The vector of weights for the external inputs in the fractional model.

$v = (v_1, \dots, v_m)$: The vector of weights for the external inputs in the linear model.

$\varphi = (\varphi_1, \dots, \varphi_q)$: The vector of weights for the intermediate measures in the fractional model.

$w = (w_1, \dots, w_q)$: The vector of weights for the intermediate measures in the linear model.

$\omega = (\omega_1, \dots, \omega_s)$: The vector of weights for the outputs in the fractional model.

$u = (u_1, \dots, u_s)$: The vector of weights for the outputs in the linear model.

$g = (g_1, \dots, g_a)$: The vector of weights of extra inputs in the fractional model.

$\gamma = (\gamma_1, \dots, \gamma_a)$: The vector of weights for the extra inputs in the linear model.

λ: The intensity vector for the first stage.

μ: The intensity vector for the second stage.

e_j^o: The overall efficiency of DMU_j.

e_j^1: The efficiency of the first stage for DMU$_j$.

e_j^2: The efficiency of the second stage for DMU$_j$.

e_j^ν: The efficiency of the ν stage for DMU$_j$.

E_j^1: The independent efficiency score of the first stage for DMU$_j$.

E_j^2: The independent efficiency score of the second stage for DMU$_j$.

3.2 Independent assessments

The *independent approach* is an elementary method for the assessment of DMUs that consist of sub-processes. Although the internal structure of the DMUs is recognized, the stage efficiencies and the overall system efficiency are calculated independently. The standard DEA model is used separately in each stage without considering possible conflicts and connections among them. In this approach the stages are treated as operating independently of one another and are assessed as independent DMUs, hence the impact of each stage to the overall efficiency cannot be measured.

Consider the basic input oriented CRS-DEA models that estimate independently the stage-1, the stage-2 and the overall efficiency for the evaluated unit j_0 with the simple Type-I (see Fig. 3.1a):

<u>*Stage 1:*</u>

$$E_{j_0}^1 = max\frac{\varphi Z_{j_0}}{\eta X_{j_0}} \quad s.t. \quad \frac{\varphi Z_j}{\eta X_j} \leq 1, \;\; j = 1, \ldots, n \quad \eta \geq 0, \varphi \geq 0 \qquad (3.1)$$

<u>*Stage 2:*</u>

$$E_{j_0}^2 = max\frac{\omega Y_{j_0}}{\hat{\varphi} Z_{j_0}} \quad s.t. \quad \frac{\omega Y_j}{\hat{\varphi} Z_j} \leq 1, \;\; j = 1, \ldots, n \quad \hat{\varphi} \geq 0, \omega \geq 0 \qquad (3.2)$$

<u>*Overall:*</u>

$$E_{j_0}^o = max\frac{\omega Y_{j_0}}{\eta X_{j_0}} \quad s.t. \quad \frac{\omega Y_j}{\eta X_j} \leq 1, \;\; j = 1, \ldots, n \quad \eta \geq 0, \omega \geq 0 \qquad (3.3)$$

Notice that the output oriented variants of the above models can be also used, as well as different assumptions concerning the returns to scale (VRS etc.). The CRS input oriented models (3.1) and (3.2) yield the independent stage efficiencies while model (3.3) provides the overall efficiency of the DMU j_0. In model (3.3) only the external inputs (X) and the final outputs (Y) are used for the assessment of the evaluated unit j_0, whereas in models (3.1) and (3.2) for stage-1 and stage-2 respectively, only their individual inputs and outputs are taken into account, i.e. (X)–(Z) for stage-1 and (Z)–(Y) for stage-2. As a result, the fact that the outputs of the first stage are the inputs to the second stage is ignored. Moreover, the overall efficiency is not connected to the individual efficiencies since they are evaluated independently. In effect, the efficiency scores derived by the independent approach are misleading. This has been reported by Kao and Hwang (2008, 2010), as a DMU may be overall efficient while the individual stages are not. Such irregular results are attributed to the fact that no coordination between the stages is assumed. Finally, the stage and the overall efficiency scores obtained by the independent approach serve as upper bounds of the stage and system efficiencies respectively. Because of its simplicity the independent approach can be applied to any network structure since the relationships among the stages are not taken into account.

Significant studies that employed the independent approach and have attracted the scientific interest are, among others, that of Charnes et al (1986), Chilingerian and Sherman (1990), Wang et al (1997), Seiford and Zhu (1999), Zhu (2000), Sexton and Lewis (2003) and Lewis and Sexton (2004). In Charnes et al (1986), the army recruitment was viewed as a two-stage process, namely the awareness creation and the contract establishment. Chilingerian and Sherman (1990) modeled the medical service as a two-stage process, where the first stage is under the control of the management and the second stage is controlled by the physician. In stage-1 the management handles the assets of the hospitals and provides with clinical outputs which are used as inputs to the stage-2. In the second stage, the physicians decide how to utilize these inputs so as to provide medical care to the patients. To be more specific, the inputs of stage-1 are nurses, management and support staff, medical supplies, various expenditures, capital and fixed costs. The intermediate measures generated by stage-1 and conveyed to stage-2 include hours of nursing care, counseling services and therapy, volume of diagnostic tests, drugs dispensed and other quantitative indicators about the medical treatment issued. The final outputs of stage-2 are research grants and quantitative indicators for the patients and the trained staff. Wang et al (1997) studied the impact of IT on the performance of 36 banks. They assumed a simple two-stage process of Type I (Fig. 3.1a) where the first stage represents the *funds collection* and the second the *investment*. Seiford and Zhu (1999) studied the performance of the top 55 commercial banks in USA by considering both the operational and the market performance. They modelled the bank operations as a simple two-stage process of Type I, with the stage-1 representing profitability and the stage-2 marketability. Within a similar framework, Zhu (2000) evaluated the performance of profitability and marketability of the Fortune 500 companies. Sexton and Lewis (2003) evaluated the performance of 30 teams of the Major League Baseball in the USA, by modelling the whole team's operations as a two-stage process of Type I, with stage-1 representing front-office operation and stage-2 representing on-field operation. Lewis and Sexton (2004) extended their previous study by modelling the team's operations with the network structure depicted in Figure 3.4. In particular, the first two stages correspond to the team's front office operations, which use funds (player salaries) to acquire talent, whereas the rest three stages correspond to the team's on-field operations, which utilize talent to win games.

3.3 Joint assessments

The independent approach neglects the conflicts or connections between the stages. Contrarily, according to the joint assessment paradigm the overall and the stage efficiencies are simultaneously estimated from one program. The efficiency decomposition approach and the slacks-based measure approach are two characteristic families that follow the joint assessment paradigm.

3.3.1 Efficiency decomposition approach

A major characteristic of the decomposition approach is that, apart from the definition of the efficiency of the individual stages (stage efficiencies), it premises the definition of the overall efficiency of the DMU together with a model to decompose the overall efficiency to the stage efficiencies. Then, the efficiency scores of the stages derive as offspring of the overall efficiency of the unit. The two basic decomposition methods dominating the literature on two-stage DEA, i.e. the multiplicative method of Kao and Hwang (2008) and the additive method of Chen et al (2009b) assume the same definitions of stage efficiencies but they differ substantially in the definition of the overall system efficiency as well as in the way they conceptualize the decomposition of the overall efficiency to the efficiencies of the individual stages. In multiplicative efficiency decomposition, the overall efficiency is defined as a product of the stage efficiencies, whereas in the additive efficiency decomposition, the overall efficiency is defined as a weighted average of the stage efficiencies.

3.3.1.1 Multiplicative efficiency decomposition

The multiplicative efficiency decomposition method is introduced by Kao and Hwang (2008) and Liang et al (2008) for the simple two-stage network structure of Type I (Fig. 3.1a). Specifically, Liang et al (2008) studied the efficiency decomposition of the two-stage process using game theoretic concepts. Under the multiplicative decomposition method the efficiency of the entire process is decomposed into the product of the efficiencies of the two individual stages. The overall efficiency and the stage efficiencies of the DMU*j*, under the CRS assumption, are defined as follows:

$$e_j^o = \frac{\omega Y_j}{\eta X_j}, e_j^1 = \frac{\varphi Z_j}{\eta X_j}, e_j^2 = \frac{\omega Y_j}{\hat{\varphi} Z_j} \qquad (3.4)$$

In order to link the efficiency assessments of the two stages, it is universally accepted that the values of the intermediate measures (virtual intermediate measures) should be same for both stages, i.e. the weights associated with the intermediate measures should be the same ($\hat{\varphi} = \varphi$), no matter if these measures are considered as outputs of the first stage or inputs to the second stage. The decomposition model assumed is as follows:

$$e_j^o = \frac{\omega Y_j}{\eta X_j} = \frac{\varphi Z_j}{\eta X_j} \cdot \frac{\omega Y_j}{\hat{\varphi} Z_j} = e_j^1 \cdot e_j^2 \qquad (3.5)$$

i.e. the overall efficiency is defined as the *square geometric average* of the stage efficiencies.

Given the above definitions, the CRS input oriented model below assesses the overall efficiency and the stage efficiencies of the evaluated unit j_0:

$$\begin{aligned} & e_{j_0}^o = max \frac{\omega Y_{j_0}}{\eta X_{j_0}} \\ & s.t. \\ & \frac{\varphi Z_j}{\eta X_j} \leq 1, \; j = 1, \dots, n \\ & \frac{\omega Y_j}{\varphi Z_j} \leq 1, \; j = 1, \dots, n \\ & \eta \geq 0, \varphi \geq 0, \omega \geq 0 \end{aligned} \qquad (3.6)$$

Notice that the constraints $\omega Y_j/\eta X_j \leq 1,\ j = 1,\dots,n$ are redundant and thus omitted. Model (3.6) is a fractional linear program that can be modeled and solved as a linear program (3.7) by applying the C-C transformation. The correspondence of variables is: $v = t\eta$, $u = t\omega$, $w = t\varphi$ where t is a scalar variable such that: $t\eta X_{j_0} = 1$.

$$
\begin{aligned}
& e_{j_0}^{o} = max\, uY_{j_0} \\
& s.t. \\
& vX_{j_0} = 1 \\
& wZ_j - vX_j \leq 0,\ \ j = 1,\dots,n \\
& uY_j - wZ_j \leq 0,\ \ j = 1,\dots,n \\
& v \geq 0, w \geq 0, u \geq 0
\end{aligned}
\tag{3.7}
$$

Once an optimal solution (v^*, w^*, u^*) of model (3.7) is obtained, the overall efficiency and the stage efficiencies are calculated as follows:

$$
e_{j_0}^{o} = u^* Y_{j_0},\ e_{j_0}^{1} = w^* Z_{j_0}, e_{j_0}^{2} = \frac{e_{j_0}^{0}}{e_{j_0}^{1}} \tag{3.8}
$$

In parallel, Liang et al (2008) developed the multiplicative decomposition in the light of game theoretic concepts. They characterized the multiplicative decomposition method described above as a cooperative or a centralized game, i.e. they refer to model (3.7) as centralized. In addition, they presented the case of non-cooperative game between the stages, where preemptive priority is given to one stage like the leader-follower situations in decentralized control systems. Liang et al (2008) and Cook et al (2010a) investigated the relations among the cooperative, the non-cooperative and the conventional DEA approaches.

Notice that the overall efficiency is obtained as the optimal value of the objective function of model (3.7), the stage-1 efficiency is given by the total virtual intermediate measure, whereas the stage-2 efficiency derives as offspring of the overall and stage-1 efficiencies. A major shortcoming of the multiplicative method is that the decomposition of the overall efficiency to the stage efficiencies is not unique. Indeed, as the term wZ_{j_0} does not appear in neither the objective function or in the normalization constraint, its value may vary and still maintain the optimal value of the objective function (i.e. the overall efficiency) and the inequality constraints of model (3.7). Also, as noted by Kao and Hwang (2008) the above deficiency renders the comparison of stage efficiencies among all DMUs lack a common basis. That is why Kao and Hwang (2008) and Liang et al (2008) propose solving a pair of linear programs, in a post-optimality phase, to obtain the largest scores for $e_{j_0}^{1}$ and $e_{j_0}^{2}$ while maintaining the overall efficiency score obtained by model (3.7). In particular, they developed a procedure for testing the uniqueness of efficiency decomposition by maximizing the efficiency of one stage under the constraint that the optimal overall efficiency obtained by (3.7) is maintained. Then the efficiency of the other stage is calculated from (3.5). The highest efficiency for the first stage is obtained by the following model.

$$
\begin{aligned}
& e_{j_0}^{1U} = max\, wZ_{j_0} \\
& s.t. \\
& vX_{j_0} = 1 \\
& uY_{j_0} = e_{j_0}^{o} \\
& wZ_j - vX_j \leq 0, \;\; j = 1, \dots, n \\
& uY_j - wZ_j \leq 0, \;\; j = 1, \dots, n \\
& v \geq 0, w \geq 0, u \geq 0
\end{aligned}
\tag{3.9}
$$

Once an optimal solution (v^*, w^*, u^*) of model (3.9) is obtained then $e_{j_0}^{1U} = w^* Z_{j_0}$ and the efficiency of the second stage is derived by $e_{j_0}^{2L} = e_{j_0}^{0} / e_{j_0}^{1U}$. Alternatively, if priority is given to the second stage the corresponding model to find its highest efficiency level is the following:

$$
\begin{aligned}
& e_{j_0}^{2U} = max\, uY_{j_0} \\
& s.t. \\
& wZ_{j_0} = 1 \\
& uY_{j_0} - e_{j_0}^{o} vX_{j_0} = 0 \\
& wZ_j - vX_j \leq 0, \;\; j = 1, \dots, n \\
& uY_j - wZ_j \leq 0, \;\; j = 1, \dots, n \\
& v \geq 0, w \geq 0, u \geq 0
\end{aligned}
\tag{3.10}
$$

Given an optimal solution (v^*, w^*, u^*) of model (3.10), the highest efficiency score of stage-2 is $e_{j_0}^{2U} = u^* Y_{j_0}$ and the resulting efficiency of the first stage is $e_{j_0}^{1L} = e_{j_0}^{0} / e_{j_0}^{2U}$. If $e_{j_0}^{1U} \neq e_{j_0}^{1L}$ or $e_{j_0}^{2U} \neq e_{j_0}^{2L}$ then the efficiency decomposition is not unique, in other words there are alternative optimal solutions that yield the same level of overall efficiency, i.e. $e_{j_0}^{0} = e_{j_0}^{1} \cdot e_{j_0}^{2} = e_{j_0}^{1U} \cdot e_{j_0}^{2L} = e_{j_0}^{1L} \cdot e_{j_0}^{2U}$.

The purpose of models (3.9) and (3.10), beyond checking the uniqueness of the efficiency decomposition, is to provide also alternative solutions in case of non-uniqueness. The argument is that one might wish giving priority to the first or the second stage in the efficiency assessments. Although there is a rationale in this argument, the non-uniqueness of the decomposition is still a problem, especially in the case that no priority is conceived by the management. Notice that the above procedure can be also applied when output orientation is selected using the output oriented models accordingly.

As mentioned above, Liang et al (2008) and Cook et al (2010a) viewed the efficiency assessment of the two-stage process as a non-cooperative game under the leader-follower assumption of a decentralized control system; this paradigm is also referred to as the Stackelberg game. A DMU may be seen as a supply chain with two parts, consisting for example of a manufacturer and a retailer. In such a setting, the manufacturer acts as a leader whereas the retailer is treated as a follower. Assuming that the first stage is the leader then its performance is computed first by applying the conventional DEA model. The leader (first stage) seeks to maximize its performance without considering the follower (second stage). The performance of the follower (second stage) is calculated subject to the requirement that the leader's efficiency is fixed at its optimal value. The following pair of LP models provides the leader-follower solution given that the stage-1 is the leader.

$$
\begin{aligned}
& e_{j_0}^{1\,Leader} = max\, wZ_{j_0} \\
& s.t. \\
& vX_{j_0} = 1 \\
& wZ_j - vX_j \leq 0, \;\; j = 1, \dots, n \\
& v \geq 0, w \geq 0
\end{aligned}
\qquad (3.11)
$$

Once the leader's efficiency ($e_{j_0}^{1\,Leader}$) is obtained the efficiency of the follower (second stage) is obtained by the following model:

$$
\begin{aligned}
& e_{j_0}^{2\,\text{Follower}} = max\ \frac{uY_{j_0}}{e_{j_0}^{1\,Leader}} \\
& s.t. \\
& vX_{j_0} = 1 \\
& wZ_{j_0} = e_{j_0}^{1\,Leader} \\
& wZ_j - vX_j \leq 0,\ \ j = 1, \dots, n \\
& uY_j - wZ_j \leq 0,\ \ j = 1, \dots, n \\
& v \geq 0, w \geq 0, u \geq 0
\end{aligned}
\tag{3.12}
$$

Alternatively, if the second stage is assumed to be the leader then its efficiency score is optimized first. The leader-follower modelling approach yields the maximum achievable efficiency score for each stage when it acts as a leader, i.e. it generates the independent efficiency scores ($e_{j_0}^{1\,Leader} = E_{j_0}^{1}, e_{j_0}^{2\,Leader} = E_{j_0}^{2}$). Liang et al (2008) remarked that contrary to the cooperative model (3.7), the non-cooperative one yields always unique efficiency decomposition.

The envelopment form of the multiplicative decomposition model is studied by Chen et al (2010a) and Chen et al (2013). The dual to the CRS input oriented model (3.7) is formulated as follows:

$$
\begin{aligned}
& \min \theta \\
& s.t. \\
& X\lambda \leq \theta X_{j_0} \\
& Y\mu \geq Y_{j_0} \\
& Z\lambda\text{-}Z\mu \geq 0 \\
& \lambda \geq 0, \mu \geq 0
\end{aligned}
\tag{3.13}
$$

Contrary to the standard DEA context where the multiplier and envelopment DEA models are dual models and equivalent, as also remarked by Chen et al (2010a) and Chen et al (2013), such is not necessarily true for the two forms of network DEA models. As they further noted, the duals to the multiplier-based network DEA models may not provide the frontier projections without exerting appropriate modifications to them. The above are also observed for model (3.13), where the usual procedure of adjusting the inputs and outputs by the efficiency scores is not adequate to provide a frontier projection. Also, model (3.13) does not provide the stage efficiency scores. These irregularities may be attributed to the conflicting nature of the intermediate measures and to the fact that may none DMU be overall efficient, i.e. may none DMU be efficient in both stages. Thus, new techniques are needed for the determination of the efficient frontier of a two-stage process. Chen et al (2010a) developed the alternative envelopment model (3.14), in order to overcome the reported inadequacies and generate the efficient frontier. They replaced the observed levels of intermediate measures by variables and separated the constraints associated with the intermediate measures.

$$\begin{aligned} & \min \tilde{\theta} \\ & s.t. \\ & X\lambda \leq \tilde{\theta} X_{j_0} \\ & Y\mu \geq Y_{j_0} \\ & Z\lambda \geq \tilde{Z}_{j_0} \\ & Z\mu \leq \tilde{Z}_{j_0} \\ & \lambda \geq 0, \mu \geq 0, \tilde{Z}_{j_0} \geq 0 \end{aligned} \quad (3.14)$$

Chen et al (2010a) showed that model (3.14) and model (3.13) yield the same overall efficiency score i.e. $\theta = \tilde{\theta}$, and model (3.14) provides additionally sufficient information on how to project inefficient DMUs onto the efficient frontier. The projection $(\hat{X}_{j_0}, \hat{Z}_{j_0}, \hat{Y}_{j_0})$ for DMU$_{jo}$ is derived by the optimal solution of model (3.14) as $\tilde{\theta}^* X_{j_0}, \tilde{Z}^*_{j_0}, Y_{j_0}$. The dual to model (3.14) is as follows:

$$\begin{aligned} & max\ uY_{j_0} \\ & s.t. \\ & vX_{j_0} = 1 \\ & w^1 Z_j - vX_j \leq 0,\ \ j = 1, \dots, n \\ & uY_j - w^2 Z_j \leq 0,\ \ j = 1, \dots, n \\ & w^2 - w^1 \leq 0 \\ & v \geq 0, w^1 \geq 0, w^2 \geq 0, u \geq 0 \end{aligned} \quad (3.15)$$

In model (3.14) the constraints $\tilde{Z}_{j_0} \geq 0$ are redundant thus can be omitted. This affects model (3.15) by converting the constraint $w^2 - w^1 \leq 0$ to equality ($w^2 - w^1 = 0$), i.e. the weights concerning the intermediate measures are the same and model (3.15) is identical to model (3.7). The findings discussed above are characterized as pitfalls of network DEA models by Chen et al (2013). They proposed that under network DEA the envelopment models should be used for deriving the frontier projection for inefficient DMUs and the multiplier ones for the determination of the efficiency scores.

Chen et al (2009a) and Cook et al (2010a) examined the relations and equivalences of the multiplicative decomposition approach with other existing network DEA methods. In particular, they established the equivalence between the studies of Fare and Grosskopf (1996), Chen and Zhu (2004), Kao and Hwang (2008) and Liang et al (2008). In particular, Chen et al (2009a) showed that the model of Chen and Zhu (2004) under CRS assumption is equivalent to the output oriented model of Kao and Hwang (2008) and the centralized output oriented model of Liang et al (2008). Also, Cook et al (2010a) illustrated that model (3.13), the dual of model (3.7), is equivalent to the models proposed by Fare and Grosskopf (1996). All these models, under CRS assumption, provide the same overall efficiency score for the two-stage process of Figure 3.1a.

A major limitation of the multiplicative decomposition method is its inability to be straightforwardly applied under the VRS assumption. This is because the extra free-in-sign variables introduced in the VRS model will render the resulting model highly non-linear. Later, Kao and Hwang (2011) proposed an approach to decompose technical and scale efficiencies of the two-stage process. They derived the scale efficiencies for the two stages assuming an input oriented VRS model for the first stage and an output oriented VRS model for the second stage. Thus, the system efficiency is decomposed into the product of the technical and scale efficiencies of the stages.

Extensions of the multiplicative efficiency decomposition

The multiplicative decomposition method can be readily applied to series multi-stage processes of Type I but not to general network structures because the assumption that the overall efficiency is the product of the stage efficiencies renders the resulting models highly nonlinear. In Kao (2009a), (2009b), (2012), (2014a) and Kao and Hwang (2010) it is shown that the overall efficiency of a DMU with the parallel structure of Figure 3.3 is the weighted average of the stage efficiencies, where the weights are derived from the proportions of inputs utilized by each stage. In the above studies the multiplicative decomposition method of Kao and Hwang (2008) is modified so as to be applied to any type of series and series-parallel multi-stage processes. Their modelling approach is based on the common assumption that the weights associated with the intermediate measures are the same. Also, they deal with general series and series-parallel multi-stage processes by transforming the multi-stage process under evaluation. In particular, dummy sub-processes are introduced in the original multi-stage process, which operate in a parallel configuration with the actual sub-processes. By applying this transformation the overall efficiency of the system is derived as the product of the efficiencies of the sub-systems, where the efficiency of each modified process (sub-system) is a weighted average of the efficiencies of the processes (real and dummies).

Below we give an example of the aforementioned technique applied to the Electricity Service System (Fig. 3.5) originally discussed in Tone and Tsutsui (2009). The first process (Generation division) generates electric power (Z^1), which then is used to the second process (Transmission division) in order to be sold to large customers as output (Y^2) or to be sent as intermediate measure (Z^2) to the third process (Distribution division) so as to provide electricity to small customers (Y^3).

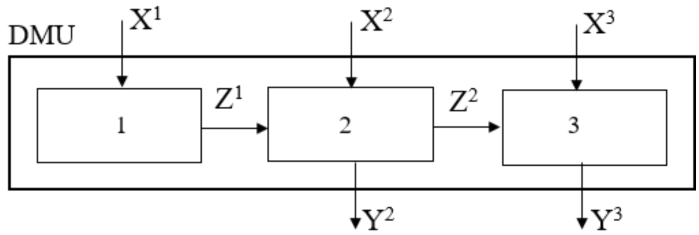


Fig. 3.5: Electric power generation, transmission and distribution (Tone and Tsutsui, 2009).

The overall efficiency of the DMU*j* and the stage efficiencies are defined as follows:

$$e_j^o = \frac{u_2 Y_j^2 + u_3 Y_j^3}{v_1 X_j^1 + v_2 X_j^2 + v_3 X_j^3}$$

$$e_j^1 = \frac{w_1 Z_j^1}{v_1 X_j^1}, e_j^2 = \frac{w_2 Z_j^2 + u_2 Y_j^2}{w_1 Z_j^1 + v_2 X_j^2}, e_j^3 = \frac{u_3 Y_j^3}{w_2 Z_j^2 + v_3 X_j^3} \quad (3.16)$$

The series network structure described above can be transformed via the approach introduced in Kao (2009a) to the network structure of Figure 3.6 below. The squares and circles represent the actual and dummy processes respectively.

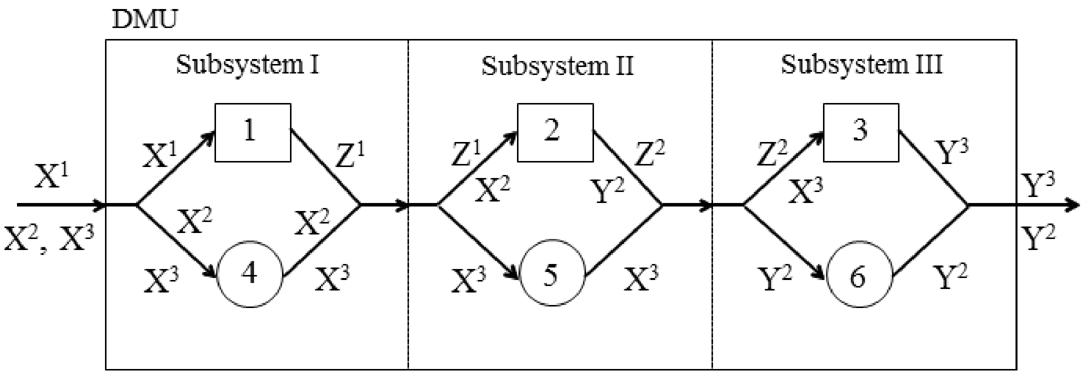


Fig. 3.6: The transformed network process of Fig. 3.5

The modified network structure contains three subsystems arranged in series, where each of them consists of an actual process and a dummy process operating in parallel. The dummy processes are introduced so as to convey the inputs and the outputs dedicated to specific processes throughout the system. The efficiencies of the three sub-systems are computed by the ratio of their aggregate output to aggregate input.

$$e_j^I = \frac{w_1 Z_j^1 + v_2 X_j^2 + v_3 X_j^3}{v_1 X_j^1 + v_2 X_j^2 + v_3 X_j^3}, e_j^{II} = \frac{w_2 Z_j^2 + u_2 Y_j^2 + v_3 X_j^3}{w_1 Z_j^1 + v_2 X_j^2 + v_3 X_j^3}, e_j^{III} = \frac{u_2 Y_j^2 + u_3 Y_j^3}{w_2 Z_j^2 + v_3 X_j^3 + u_2 Y_j^2}$$

As mentioned above, the overall efficiency of a system, whose stages are in parallel, is the weighted average of the stage efficiencies. Thus, in the above transformed network (Fig. 3.6) the efficiency of each sub-system is obtained as the weighted average of the efficiencies of the actual and dummy process. The weights are derived endogenously from the optimization process as the proportions of inputs consumed by each process. Notice that the dummy processes have the same inputs and outputs, therefore their efficiency score is one.

- $e_j^I = t_j^1 e_j^1 + t_j^4 e_j^4 = t_j^1 e_j^1 + (1 - t_j^1)$ where $t_j^1 = (v_1 X_j^1)/(v_1 X_j^1 + v_2 X_j^2 + v_3 X_j^3)$ and $t_j^1 + t_j^4 = 1$

- $e_j^{II} = t_j^2 e_j^2 + t_j^5 e_j^5 = t_j^2 e_j^2 + (1 - t_j^2)$ where $t_j^2 = (w_1 Z_j^1 + v_2 X_j^2)/(w_1 Z_j^1 + v_2 X_j^2 + v_3 X_j^3)$ and $t_j^2 + t_j^5 = 1$

- $e_j^{III} = t_j^3 e_j^3 + t_j^6 e_j^6 = t_j^3 e_j^3 + (1 - t_j^3)$ where $t_j^3 = (w_2 Z_j^2 + v_3 X_j^3)/(w_2 Z_j^2 + v_3 X_j^3 + u_2 Y_j^2)$ and $t_j^3 + t_j^6 = 1$

From the above mathematical relationships it follows that the overall (system) efficiency of the DMU*j* can be calculated as the product of the three sub-systems efficiencies.

$$e_j^o = e_j^I \cdot e_j^{II} \cdot e_j^{III} = \left[t_j^1 e_j^1 + \left(1 - t_j^1\right)\right] \cdot \left[t_j^2 e_j^2 + \left(1 - t_j^2\right)\right] \cdot \left[t_j^3 e_j^3 + \left(1 - t_j^3\right)\right]$$

The resulting model for the performance assessment for DMU j_0 with the above network structure is given as:

$$\begin{aligned}
& e_{j_0}^o = max\, u_2 Y_{j_0}^2 + u_3 Y_{j_0}^3 \\
& s.t. \\
& v_1 X_{j_0}^1 + v_2 X_{j_0}^2 + v_3 X_{j_0}^3 = 1 \\
\textbf{System} \quad & u_2 Y_j^2 + u_3 Y_j^3 - v_1 X_j^1 - v_2 X_j^2 - v_3 X_j^3 \leq 0,\ \ j = 1, \dots, n \\
\textbf{1}^{\textbf{st}} \textbf{ process} \quad & w_1 Z_j^1 - v_1 X_j^1 \leq 0,\ \ j = 1, \dots, n \\
\textbf{2}^{\textbf{nd}} \textbf{ process} \quad & u_2 Y_j^2 + w_2 Z_j^2 - w_1 Z_j^1 - v_2 X_j^2 \leq 0,\ \ j = 1, \dots, n \\
\textbf{3}^{\textbf{rd}} \textbf{ process} \quad & u_3 Y_j^3 - w_2 Z_j^2 - v_3 X_j^3 \leq 0,\ \ j = 1, \dots, n \\
& v_1, v_2, v_3, w_1, w_2, u_1, u_2 \geq 0
\end{aligned} \qquad (3.17)$$

Once an optimal solution of model (3.17) is obtained, the overall and the actual stage efficiencies are calculated from the relationships (3.16). However, the decomposition of the overall efficiency to the stage efficiencies might be not unique (Fukuyama and Mirdehghan, 2012). To summarize, the shortcoming of non-unique efficiency scores may occur in the assessment of any type of network structure when the multiplicative decomposition method is applied.

Alternative multiplicative efficiency decomposition methods

As shown above, a modified version of the multiplicative decomposition approach is proposed by Kao (2009a), (2014a) so as to be applicable to any type of series and series-parallel multi-stage processes. Beyond that, alternative methods have been developed whose common characteristic is that the overall efficiency is defined as the product of the stage efficiencies. However, generalizing this assumption to multi-stage networks different to Type I leads to high non-linear models which are difficult to solve. A common solution practice is to use parametric techniques.

Zha and Liang (2010) studied a modified two-stage process of Type II (Fig. 3.1b) where the external inputs are freely allocated between the stages. In the study of Zha and Liang (2010), the overall efficiency of the system is derived as the product of the efficiencies of the two individual stages, as in Kao and Hwang (2008). Zha and Liang (2010) incorporated game-theory framework and a heuristic procedure so as to overcome the linearization issues raised by the adoption of the multiplicative format of the overall efficiency. In particular, using the concept of Stackelberg (non-cooperative) game, they first computed the lower and upper bounds of the stage efficiencies. Then, they incorporated this information into a non-linear cooperative model and by treating the efficiency of one stage as a parameter they succeeded to transform it to a parametric linear program. Their method is illustrated by using the dataset of 30 top U.S. commercial banks which originally studied by Seiford and Zhu (1999).

Li et al (2012) studied also a two-stage production process of Type II (Fig. 3.1b) in the view of cooperative (centralized control) and non-cooperative games (decentralized control). They developed a parametric approach in order to obtain the stage efficiency scores and then the overall efficiency is computed by the product of the stage efficiencies. Their approach is demonstrated by evaluating the research and development of 30 Chinese regions. The stage efficiencies as well as the overall efficiency are defined as follows:

$$e_j^1 = \frac{\varphi Z_j}{\eta X_j}, e_j^2 = \frac{\omega Y_j}{\varphi Z_j + g L_j}, e_j^o = e_j^1 \cdot e_j^2 \tag{3.18}$$

The extra inputs (L) which are utilized by the second stage render non-linear the function of the overall efficiency. As a result, the objective function of the evaluation model proposed by Li et al (2012) is non-linear:

$$\begin{aligned}
& e_{j_0}^o = max \frac{\varphi Z_{j_0}}{\eta X_{j_0}} \cdot \frac{\omega Y_{j_0}}{\varphi Z_{j_0} + g L_{j_0}} \\
& s.t. \\
& \frac{\varphi Z_j}{\eta X_j} \leq 0, \ \ j = 1, \dots, n \\
& \frac{\omega Y_j}{\varphi Z_j + g L_j} \leq 0, \ \ j = 1, \dots, n \\
& \eta \geq 0, \varphi \geq 0, \omega \geq 0, g \geq 0
\end{aligned} \tag{3.19}$$

In order to deal with the non-linearity issues the authors resorted to a heuristic search procedure in order to estimate a global optimal solution. Firstly, they calculated the largest efficiency score for one stage and then they included it as a parameter to the evaluation model (3.19) so as to handle it as a parametric linear program. Assuming that in model (3.19) the efficiency score of stage-1 is chosen to be treated as a parameter, then its upper bound is obtained from the following linear model:

$$
\begin{aligned}
& E_{j_0}^1 = max\ wZ_{j_0} \\
& s.t. \\
& vX_{j_0} = 1 \\
& wZ_j - vX_j \leq 0,\ \ j = 1, \dots, n \\
& uY_j - wZ_j + \gamma L_j \leq 0,\ \ j = 1, \dots, n \\
& v \geq 0, w \geq 0, u \geq 0, \gamma \geq 0
\end{aligned}
\qquad (3.20)
$$

The largest first stage efficiency score ($E_{j_0}^1$ – independent efficiency score) is derived from the optimal solution of model (3.20). Hence, the efficiency of the first stage ($e_{j_0}^1$) in model (3.19) can be treated as a parameter in the interval $[0, E_{j_0}^1]$. Thus model (3.19), after applying the C-C transformation, can be rewritten as follows:

$$
\begin{aligned}
& e_{j_0}^o = max\ e_{j_o}^1 \cdot uY_{j_0} \\
& s.t. \\
& wZ_{j_0} + \gamma L_{j_0} = 1 \\
& wZ_j - vX_j \leq 0,\ \ j = 1, \dots, n \\
& uY_j - wZ_j + \gamma L_j \leq 0,\ \ j = 1, \dots, n \\
& wZ_{j_0} - e_{j_o}^1 vX_{j_0} = 0 \\
& v \geq 0, w \geq 0, u \geq 0, \gamma \geq 0 \\
& e_{j_o}^1 \in [0, E_{j_0}^1]
\end{aligned}
\qquad (3.21)
$$

The parameter $e_{j_0}^1$ is progressively increased by a small step until the upper bound $E_{j_0}^1$ of the interval is reached. For any given value of parameter $e_{j_0}^1$ the overall and the second stage efficiency are calculated unless the program (3.21) is infeasible. Once the heuristic procedure is finished then reasonably they select the stage efficiencies that yield the maximum achievable level of overall efficiency.

3.3.1.2 Additive efficiency decomposition

The additive efficiency decomposition method is introduced by Chen et al (2009b) for the assessment of the two-stage process of Type I (Fig. 3.1a) and then is extended by Cook et al (2010b) for the evaluation of multi-stage processes of varying structures. Both studies have already received great attention from the research community. In contrast to the multiplicative efficiency decomposition method, the overall efficiency is obtained as a weighted average of the stage efficiencies, where the weights represent the portion of all inputs utilized by each stage. Notably, this aggregation method is used previously in some network DEA studies without, however, being part of a well-established efficiency decomposition framework. For instance, it is first appeared in Beasley (1995), who evaluated the efficiency of teaching and research of the UK chemistry and physics departments and viewed them as two processes that operate in parallel and share some resources. The aforementioned aggregation method was also adopted by Cook and Hababou (2001), Cook and Green (2004) and Jahanshahloo et al (2004), who similarly examined parallel production processes with shared inputs. Amirteimoori and Kordrostami (2005) and Amirteimoori and Shafiei (2006) aimed to measure the performance of series processes using the aforementioned aggregation method about the overall and the stage efficiencies, however they treated the stages in a non-coordinated manner. In particular, the weights associated with the intermediate measures were different for each stage.

In the context of additive efficiency decomposition method, the overall efficiency and the stage efficiencies, under CRS assumption, of the DMU j are defined as follows:

$$e_j^o = \frac{\omega Y_j + \varphi Z_j}{\eta X_j + \varphi Z_j}, e_j^1 = \frac{\varphi Z_j}{\eta X_j}, e_j^2 = \frac{\omega Y_j}{\varphi Z_j} \tag{3.22}$$

The definition of the stage efficiencies are the same as in the multiplicative method, but the additive method differentiates in the definition of the overall efficiency. In (3.22) the intermediate measures appear in both terms of the fraction that defines the overall efficiency, meaning that they are considered as inputs and as outputs simultaneously. The decomposition model used is as follows:

$$e_j^o = \frac{\omega Y_j + \varphi Z_j}{\eta X_j + \varphi Z_j} = t_j^1 \frac{\varphi Z_j}{\eta X_j} + t_j^2 \frac{\omega Y_j}{\varphi Z_j},\ t_j^1 + t_j^2 = 1 \tag{3.23}$$

i.e. the overall efficiency is expressed as a *weighted arithmetic average* of the stage efficiencies. The functional forms of the weights derive by solving the system (3.23) for t_j^1 and t_j^2, as follows:

$$t_j^1 = \frac{\eta X_j}{\eta X_j + \varphi Z_j}, t_j^2 = \frac{\varphi Z_j}{\eta X_j + \varphi Z_j} \tag{3.24}$$

Chen et al (2009b) noticed that the weights t_j^1 and t_j^2 "*are intended to represent the relative importance or contribution of the performance of stages 1 and 2, respectively, to the overall performance of the DMU*" and argued that the "size" of a stage reflects its importance. They also noted that the size can be computed by the portion of total resources devoted to each stage. It is worth to note that as the weights are functions of the virtual intermediate measures, they depend on the unit being evaluated and, obviously, they generally differentiate from one unit to another. Given the above definitions, the input oriented CRS model below assesses the overall efficiency of the evaluated unit j_0:

$$\begin{aligned} & e_{jo}^o = max \frac{\omega Y_{j_o} + \varphi Z_{j_o}}{\eta X_{j_o} + \varphi Z_{j_o}} \\ & s.t. \\ & \frac{\varphi Z_j}{\eta X_j} \leq 1, \ j = 1, \dots, n \\ & \frac{\omega Y_j}{\varphi Z_j} \leq 1, \ j = 1, \dots, n \\ & \eta \geq 0, \varphi \geq 0, \omega \geq 0 \end{aligned} \tag{3.25}$$

Applying the C-C transformation to the linear fractional model (3.25), the following linear program is modeled and solved:

$$\begin{aligned} & e_{jo}^o = max\, uY_{j_o} + wZ_{j_o} \\ & s.t. \\ & vX_{j_o} + wZ_{j_o} = 1 \\ & wZ_j - vX_j \leq 0, \ j = 1, \dots, n \\ & uY_j - wZ_j \leq 0, \ j = 1, \dots, n \\ & v \geq 0, w \geq 0, u \geq 0 \end{aligned} \tag{3.26}$$

Once an optimal solution (v^*,w^*,u^*) of model (3.26) is obtained, the overall efficiency and the stage efficiencies are calculated as follows:

$$\begin{aligned} e^o_{jo} &= u^* Y_{j_o} + w^* Z_{j_o} \\ t^1_{jo} &= v^* X_{j_o}\ ,\ t^2_{jo} = w^* Z_{j_o} \\ e^1_{jo} &= \frac{w^* Z_{j_o}}{v^* X_{j_o}} \\ e^2_{jo} &= \frac{e^0_{jo} - t^1_{jo}\, e^1_{jo}}{t^2_{jo}} = \frac{u^* Y_{j_o}}{w^* Z_{j_o}} \end{aligned} \tag{3.27}$$

The overall efficiency e^o_{jo} is obtained as the optimal value of the objective function, the weight t^1_{jo} is obtained as the optimal virtual input, the weight t^1_{jo} is obtained as the optimal virtual intermediate measure and the efficiency of the first stage e^1_{jo} is given by the ratio of the two weights whereas the efficiency of the second stage e^2_{jo} is obtained as offspring of $e^o_{jo}, e^1_{jo}, t^1_{jo}, t^2_{jo}$.

In case an output orientation is selected, the "size" of each stage (*weight*) is measured by the portion of the total outputs produced from each stage:

$$t^1_j = \frac{wZ_j}{uY_j + wZ_j}, t^2_j = \frac{uY_j}{uY_j + wZ_j} \tag{3.28}$$

As above, the overall efficiency is expressed as a *weighted arithmetic average* of the stage efficiencies:

$$e^o_j = \frac{vX_j + wZ_j}{uY_j + wZ_j} = t^1_j \frac{vX_j}{wZ_j} + t^2_j \frac{wZ_j}{uY_j}\ , t^1_j + t^2_j = 1 \tag{3.29}$$

The output oriented model under CRS assumption is formulated as:

$$\begin{aligned} & e^o_{jo} = min\, vX_{j_o} + wZ_{j_o} \\ & s.t. \\ & uY_{j_o} + wZ_{j_o} = 1 \\ & wZ_j - vX_j \le 0, \ \ j = 1, \dots, n \\ & uY_j - wZ_j \le 0, \ \ j = 1, \dots, n \\ & v \ge 0, w \ge 0, u \ge 0 \end{aligned} \tag{3.30}$$

Similarly to the case of the multiplicative efficiency decomposition, the additive decomposition of the overall efficiency to the stage efficiencies is non-unique. Chen et al (2009b) developed a procedure, similar to that of Kao and Hwang (2008) and Liang et al (2008), so as to derive extreme efficiency decompositions. Let t_j^{1*}, t_j^{2*} and e_{jo}^{o*} the optimal weights and the overall efficiency obtained from model (3.26). Then, if pre-emptive priority is given to the first stage, the efficiency of that stage is calculated first in a manner that the overall efficiency is maintained via the following model.

$$
\begin{aligned}
& e_{j_0}^{1U} = max\, wZ_{j_o} \\
& s.t. \\
& vX_{j_o} = 1 \\
& \left(1 - e_{jo}^{o*}\right) wZ_{j_o} + uY_{j_o} = e_{jo}^{o*} \\
& wZ_j - vX_j \leq 0, \;\; j = 1, \dots, n \\
& uY_j - wZ_j \leq 0, \;\; j = 1, \dots, n \\
& v \geq 0, w \geq 0, u \geq 0
\end{aligned}
\qquad (3.31)
$$

The corresponding efficiency for the second stage is given by $e_{jo}^{2L} = \left(e_{jo}^{o*} - t_{jo}^{1*} e_{jo}^{1U}\right) / t_{jo}^{2*}$. If pre-emptive priority is given to stage-2, then the second stage efficiency is first estimated with the constraint that the overall efficiency is preserved, by the following model:

$$
\begin{aligned}
& e_{j_0}^{2U} = max\, uY_{j_o} \\
& s.t. \\
& wZ_{j_o} = 1 \\
& wZ_{j_o} + uY_{j_o} - e_{jo}^{o*} vX_{j_o} = e_{jo}^{o*} \\
& wZ_j - vX_j \leq 0, \;\; j = 1, \dots, n \\
& uY_j - wZ_j \leq 0, \;\; j = 1, \dots, n \\
& v \geq 0, w \geq 0, u \geq 0
\end{aligned}
\qquad (3.32)
$$

The corresponding efficiency of the first stage is calculated by $e_{jo}^{1L} = \left(e_{jo}^{o*} - t_{jo}^{2*} e_{jo}^{2U}\right) / t_{jo}^{1*}$.

The modelling approach adopted by Chen et al (2009b) for the additive efficiency decomposition, enables the straightforward assessment of the two-stage process of Type I

(Fig. 3.1a) under variable returns to scale. The overall efficiency of the DMU j as well as the stage efficiencies under VRS assumption are defined as follows:

$$e_j^{o\_vrs} = \frac{wZ_j+\xi_1+uY_j+\xi_2}{wZ_j+vX_j}, e_j^{1\_vrs} = \frac{wZ_j+\xi_1}{vX_j}, e_j^{2\_vrs} = \frac{uY_j+\xi_2}{wZ_j} \tag{3.33}$$

When input orientation is chosen, the weights that reflect the "size" of each stage are defined as follows:

$$t_j^1 = \frac{vX_j}{vX_j + wZ_j}, t_j^2 = \frac{wZ_j}{vX_j + wZ_j} \tag{3.34}$$

The overall efficiency is defined as a weighted arithmetic average of the stage efficiencies:

$$e_j^{o\_vrs} = \frac{wZ_j+\xi_1+uY_j+\xi_2}{wZ_j+vX_j} = t_j^1 \frac{wZ_j+\xi_1}{vX_j}, + t_j^2 \frac{uY_j+\xi_2}{wZ_j}, \; t_j^1 + t_j^2 = 1 \tag{3.35}$$

The resulting VRS input oriented model that provides the overall efficiency of DMU j_0 is:

$$\begin{aligned}
& e_{jo}^{o_vrs} = max\, wZ_{j_o} + \xi_1 + uY_{j_o} + \xi_2 \\
& s.t. \\
& vX_{j_o} + wZ_{j_o} = 1 \\
& wZ_j - vX_j + \xi_1 \le 0, \;\; j = 1, \dots, n \\
& uY_j - wZ_j + \xi_2 \le 0, \;\; j = 1, \dots, n \\
& v \ge 0, w \ge 0, u \ge 0
\end{aligned} \tag{3.36}$$

Once an optimal solution of model (3.36) is derived, the stage efficiencies can be computed from (3.33). Although the additive efficiency decomposition can be straightforwardly applied to variable returns to scale, the standard property that the VRS efficiency scores are not less than the CRS efficiency scores does not hold.

Extensions of the Additive Efficiency Decomposition

Cook et al (2010b) extended the additive decomposition method of Chen et al (2009b) to series, parallel and series-parallel multi-stage processes. For instance, we provide below the additive modelling approach for a netwok stucture which consists of four sub-processes arranged in a series-parallel (mixed) configuration.

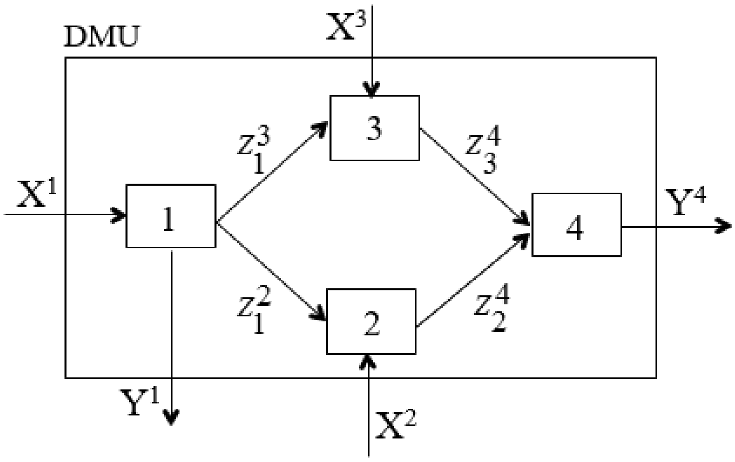


Fig. 3.7: Multi-stage process (DMU) with series-parallel (mixed) internal structure

Following Chen et al (2009b), the overall efficiency is defined as a weighted arithmetic average of the stage efficiencies:

$$e_j^o = t_j^1 e_j^1 + t_j^2 e_j^2 + t_j^3 e_j^3 + t_j^4 e_j^4,\ t_j^1 + t_j^2 + t_j^3 + t_j^4 = 1$$

$$e_j^1 = \frac{u_1 Y_j^1 + w_1^2 Z_{1_j}^2 + w_1^3 Z_{1_j}^3}{v_1 X_j^1}, e_j^2 = \frac{w_2^4 Z_{2_j}^4}{v_2 X_j^2 + w_1^2 Z_{1_j}^2}, \quad (3.37)$$

$$e_j^3 = \frac{w_3^4 Z_{3_j}^4}{v_3 X_j^3 + w_1^3 Z_{1_j}^3}, e_j^4 = \frac{u_4 Y_j^4}{w_2^4 Z_{2_j}^4 + w_3^4 Z_{3_j}^4}$$

The weight associated with each stage is obtained from the proportion of inputs that used by this stage as follows:

$$t_j^1 = \frac{v_1 X_j^1}{TVI}, t_j^2 = \frac{v_2 X_j^2 + w_1^2 Z_{1_j}^2}{TVI}, t_j^3 = \frac{v_3 X_j^3 + w_1^3 Z_{1_j}^3}{TVI}, t_j^4 = \frac{w_2^4 Z_{2_j}^4 + w_3^4 Z_{3_j}^4}{TVI} \quad (3.38)$$

where $TVI = v_1X_j^1 + v_2X_j^2 + v_3X_j^3 + w_1^2Z_{1_j}^2 + w_1^3Z_{1_j}^3 + w_2^4Z_{2_j}^4 + w_3^4Z_{3_j}^4$. The efficiency assessment of the multi-stage process depicted in Figure 3.7 is carried out by the following model:

$$
\begin{aligned}
& e_{j_0}^o = max\; u_1Y_{j_0}^1 + u_2Y_{j_0}^2 + w_1^2Z_{1_{j_0}}^2 + w_1^3Z_{1_{j_0}}^3 + w_2^4Z_{2_{j_0}}^4 + w_3^4Z_{3_{j_0}}^4 \\
& s.t. \\
& v_1X_{j_0}^1 + v_2X_{j_0}^2 + v_3X_{j_0}^3 + w_1^2Z_{1_{j_0}}^2 + w_1^3Z_{1_{j_0}}^3 + w_2^4Z_{2_{j_0}}^4 + w_3^4Z_{3_{j_0}}^4 = 1 \\
& u_1Y_j^1 + w_1^2Z_{1_j}^2 + w_1^3Z_{1_j}^3 - v_1X_j^1 \le 0,\;\; j = 1, \dots, n \\
& w_2^4Z_{2_j}^4 - v_2X_j^2 - w_1^2Z_{1_j}^2 \le 0,\;\; j = 1, \dots, n \\
& w_3^4Z_{3_j}^4 - v_3X_j^3 - w_1^3Z_{1_j}^3 \le 0,\;\; j = 1, \dots, n \\
& u_4Y_j^4 - w_2^4Z_{2_j}^4 - w_3^4Z_{3_j}^4 \le 0,\;\; j = 1, \dots, n \\
& v_1, v_2, v_3, w_1^2, w_1^3, w_2^4, w_3^4, u_1, u_4 \ge 0
\end{aligned}
\tag{3.39}
$$

Once an optimal solution of model (3.39) is obtained the overall and the stage efficiencies for DMU$_{jo}$ as well as the weights are calculated from the relationships (3.37) and (3.38) respectively. Notably, model (3.39) inherits the defects of additive decomposition method in the sense that the overall efficiency decomposition to stage efficiencies is not unique. However, model (3.39) can be adapted to meet the VRS assumption.

Chen et al (2010b) extended the work of Chen et al (2009b) for the efficiency assessment of two-stage production processes with shared resources. Particularly, they assumed a two-stage process as in Figure 3.8 where the second stage uses, beyond the intermediate measures (*Z*), a portion of the external inputs (*X*). They applied their models to the assessment of the benefits of information technology in banking industry, originally studied by Wang et al (1997).

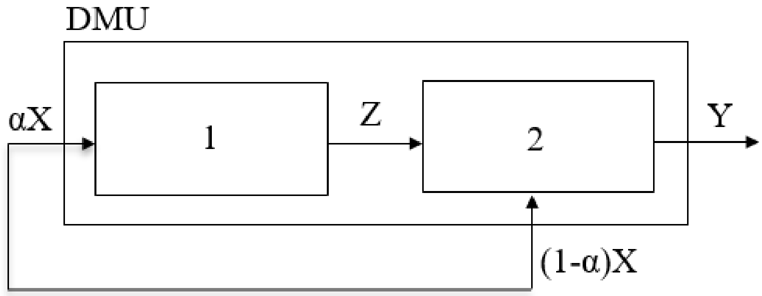


Fig. 3.8: Two-stage production process with shared resources

Alternative additive efficiency decomposition methods

As alternative additive decomposition methods are considered those in which, the overall efficiency is defined as a weighted arithmetic average of the stage efficiencies and the weights are predetermined and given as parameters instead of being endogenously estimated by the optimization process. However, notice that using this aggregation method for the efficiency assessment of network structures of any form, leads to non-linear models. Parametric techniques are commonly used to handle the non-linearity issues.

Liang et al (2006) proposed that the operations of a seller-buyer supply chain can be modelled, under both cooperative and non-cooperative concepts, as a two-stage process of Type II (Fig. 3.1b). They unified the performance assessment models of the two stages, based on the common assumption that the weights of the intermediate measures are the same in both stages. The overall efficiency is defined as the simple arithmetic average of the stage efficiencies. When a leader-follower (non-cooperative game) situation is assumed, then similar models to (3.11) and (3.12) are employed, which are adapted to Type II structure. Under the cooperative concept, the efficiencies of the seller and the buyer are jointly maximized, thus the following non-linear model is derived:

$$e_{j_0}^{o} = max \frac{1}{2}\left[\frac{\varphi Z_{j_0}}{\eta X_{j_0}} + \frac{\omega Y_{j_0}}{\varphi Z_{j_0} + gL_{j_0}}\right]$$
$$s.t.$$
$$\frac{\varphi Z_j}{\eta X_j} \leq 0, \;\; j = 1, \ldots, n \qquad (3.40)$$
$$\frac{\omega Y_j}{\varphi Z_j + gL_j} \leq 0, \;\; j = 1, \ldots, n$$
$$\eta \geq 0, \varphi \geq 0, \omega \geq 0, g \geq 0$$

Liang et al (2006) assumed the following transformation in order to transform the above model to a parametric linear program.

$$\tau^1 = 1/\, \eta X_{j_0}, \tau^2 = 1/\big(\, \varphi Z_{j_0} + gL_{j_0}\big)$$
$$v = \eta\tau^1, w^1 = \varphi\tau^1$$
$$u = \omega\tau^2, w^2 = \varphi\tau^2, \gamma = g\tau^2$$

From the above transformation, a linear relationship is implied between w^1 and w^2, i.e. $w^2 = \zeta w^1$, with $\zeta \geq 0$ and $\zeta = \big(1 - \gamma L_{j_0}\big)/w^1 Z_{j_0} < \big(1/w^1 Z_{j_0}\big)$ since $\gamma L_{j_0} + \zeta w^1 Z_{j_0} = 1$ and $w^1 Z_{j_0} \leq 1$, $\gamma L_{j_0} > 0$.

$$e_{j_0}^{o} = max \frac{1}{2}\big(w^1 Z_{j_0} + uY_{j_0}\big)$$
$$s.t.$$
$$vX_{j_0} = 1$$
$$\gamma L_{j_0} + \zeta w^1 Z_{j_0} = 1 \qquad (3.41)$$
$$w^1 Z_j - vX_j \leq 0, \;\; j = 1, \ldots, n$$
$$uY_j - \zeta w^1 Z_j - \gamma L_j \leq 0, \;\; j = 1, \ldots, n$$
$$v \geq 0, w^1 \geq 0, u \geq 0, \gamma \geq 0, \zeta \geq 0$$

Notice that the first stage efficiency ($w^1 Z_{j_0}$) in model (3.41) will not be less than the lowest efficiency score $\big(e_{j_0}^{1\,Follower}\big)$ obtained when the stage-1 is treated as a follower. Therefore, the variable ζ in model (3.41) can be treated as a parameter in the interval [0, $1/e_{j_0}^{1\,Follower}$). The model (3.41) is solved for different values of the parameter ζ and the pair of efficiency scores that provides the maximum overall efficiency is selected.

Chen et al (2006) examined the IT impact on banking industry previously studied by Wang et al (1997) and Chen and Zhu (2004), they remarked though that these studies do not fully characterize the IT impact on firm performance. Therefore, they proposed that the external IT-related inputs of stage-1 should be shared with stage-2 (Fig. 3.8). Similar to Liang et al (2006), they treat the resulting non-linear assessment model using the aforementioned procedure and they calculate the overall efficiency from the simple arithmetic average of the stage efficiencies.

Liang et al (2011) studied a serial two-stage production process with feedback, as depicted in Figure 3.9. In this system, some outputs from the second process are fed back as inputs to the first process, i.e. they have a double role serving both as inputs and outputs.

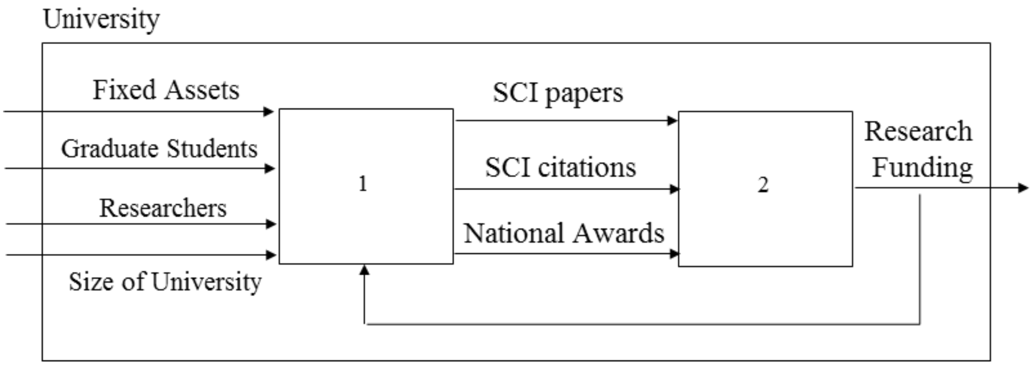


Fig. 3.9: Two-stage process with feedback

Similarly to the aforementioned studies the overall efficiency is derived as the arithmetic average of the stage efficiencies. Again, the resulting non-linear model is transformed to a parametric LP and a global optimal solution is obtained as in Liang et al (2006). Liang et al (2011) illustrated their approach by measuring the performance of 50 Chinese universities. In particular, they assumed as inputs to the first stage the *fixed assets*, the *researchers*, the *graduate students* and the *size* of each university, while they assumed as outputs the numbers of SCI papers, SCI citations and national awards. These outputs serve as inputs to the second stage, i.e. they are the intermediate measures of the system, in order to attract research funds from the granting agency. The research funding, which is the only output of the second stage, is fed back to the first stage i.e. it serves also as input.

3.3.2 Slacks-Based Measure approach

Tone and Tsutsui (2009) introduced the network slacks-based measure (NSBM) based on the SBM and the weighted SBM proposed by Tone (2001) and Tsutsui and Goto (2009) respectively. Their approach assesses simultaneously the overall and the stage efficiencies of the evaluated units. They built their method in the envelopment form based on a generalized production possibility set that describes the relationships of the multi-stage processes. In particular, they assumed that a DMU consists of ν sub-processes ($\delta=1,\ldots,\nu$), where each sub-process consumes external inputs X^{δ} to produce some outputs Y^{δ} (the superscript δ denotes the sub-process). The sub-processes are connected and interact via the intermediate measures $Z^{(\delta,\psi)}$, where the superscripts δ and ψ ($\delta\neq\psi$) represent the source sub-process and the recipient sub-process respectively. The generalized production possibility set $\{X^{\delta}, Z^{(\delta,\psi)}, Y^{\delta}\}$ under VRS assumption is defined as:

$$
\begin{aligned}
& X^{\delta}h^{\delta} \leq X^{\delta}, \quad \delta = 1,\ldots,\nu \\
& Z^{(\delta,\psi)}h^{\delta} = Z^{(\delta,\psi)}, \quad \forall\,(\delta,\psi)(as\ outputs\ from\ \delta) \\
& Z^{(\delta,\psi)}h^{\psi} = Z^{(\delta,\psi)}, \quad \forall\,(\delta,\psi)(as\ inputs\ to\ \psi) \\
& Y^{\delta}h^{\delta} \geq Y^{\delta}, \quad \delta = 1,\ldots,\nu \\
& h^{\delta} \geq 0\,, \qquad \forall\,(\delta) \\
& eh^{\delta} = 1\,, \qquad \forall\,(\delta)
\end{aligned}
\tag{3.42}
$$

Notice that the intensity vector h^{δ} is specific to each sub-process δ ($\delta=1,\ldots,\nu$). The above VRS production possibility set can be also used under CRS assumption by removing the last set of convexity constraints ($h^{\delta}=1$). Tone and Tsutsui (2009) proposed two options for representing the constraints corresponding to the intermediate measures:

a) The “free” link case, where the linking flows are freely determined, i.e. $Z^{(\delta,\psi)}h^{\delta} = Z^{(\delta,\psi)}h^{\psi}, \forall(\delta,\psi)$. In this case, the intermediate measures that link the stages are tested in the light of the other DMUs. Hence, the intermediate measures may increase or decrease in order to preserve the continuity of being simultaneously outputs of one stage and inputs to some other.

b) The “fixed” link case, where the intermediate measures are kept unchanged on their initial levels, i.e. $Z^{(\delta,\psi)}h^{\delta} = Z_{j_o}^{(\delta,\psi)}$, $Z^{(\delta,\psi)}h^{\psi} = Z_{j_o}^{(\delta,\psi)}$, $\forall(\delta,\psi)$. Notice that the subscript j_o denotes the DMU under evaluation. Tone and Tsutsui (2009) remarked that if all the intermediate measures are fixed to their original levels then the analysis to follow will treat the stages separately similar to the independent assessment.

By incorporating the input and output slacks and one of the above options for the constraints of the intermediate measures, the DMUj_o under evaluation is expressed as follows:

$$\begin{aligned} & X^{\delta} h^{\delta} + s^{\delta -} = X^{\delta}_{j_o}, \quad \delta = 1, \dots, \nu \\ & Y^{\delta} h^{\delta} - s^{\delta +} = Y^{\delta}_{j_o}, \quad \delta = 1, \dots, \nu \\ & e h^{\delta} = 1, \quad \delta = 1, \dots, \nu \\ & h^{\delta} \geq 0, s^{\delta -} \geq 0, s^{\delta +} \geq 0, \quad \delta = 1, \dots, \nu \end{aligned} \tag{3.43a}$$

$$Z^{(\delta,\psi)} h^{\delta} - Z^{(\delta,\psi)} h^{\psi} = 0, \ \forall (\delta, \psi) \qquad (\textit{free link}) \tag{3.43b}$$

or

$$\begin{aligned} & Z^{(\delta,\psi)} h^{\delta} = Z^{(\delta,\psi)}_{j_o}, \ \forall (\delta, \psi) \\ & Z^{(\delta,\psi)} h^{\psi} = Z^{(\delta,\psi)}_{j_o}, \ \forall (\delta, \psi) \end{aligned} \qquad (\textit{fixed link}) \tag{3.43c}$$

Tone and Tsutsui (2009), similar to the conventional SBM, proposed three different efficiency measures based on the orientation, they formed the input, the output and the non-oriented situation. When input orientation is selected then the overall efficiency of the DMUj_o is derived as a weighted arithmetic mean of the slacks-based measures of the individual stages, i.e. $e^{o}_{j_0} = \sum_{\delta=1}^{\nu} w^{\delta} \cdot e^{\delta}_{j_0}$, with $\sum_{\delta=1}^{\nu} w^{\delta} = 1$ and $w^{\delta} \geq 0$. The weights w^{δ} are predefined by the analyst and represent the importance of each stage. The input oriented NSBM model for the efficiency assessment of the DMUj_o is as follows:

$$\begin{aligned} & e^{o}_{j_0} = min \sum_{\delta=1}^{\nu} w^{\delta} \left[1 - \frac{1}{m_{\delta}} \left(\sum_{i=1}^{m_{\delta}} \frac{s_i^{\delta -}}{x^{\delta}_{i_{jo}}} \right) \right] \\ & \textit{subject to } (3.43a), (3.43b) \textit{ or } (3.43c) \end{aligned} \tag{3.44}$$

In model (3.44), the number of inputs consumed by each stage δ is denoted by m_{δ}, also the *free* or the *fixed* link case can be used to represent the constraints corresponding to the intermediate measures. When an optimal solution of model (3.44) is obtained then the overall efficiency can be directly obtained from its objective function and the stage efficiencies are calculated using the optimal input slacks $s^{\delta -*}$ as follows:

$$e_{j_0}^{\delta} = 1 - \frac{1}{m_\delta}\left(\sum_{i=1}^{m_\delta} \frac{s_i^{\delta-*}}{x_{i_{jo}}^{\delta}}\right), \qquad \delta = 1, \dots, \nu \tag{3.45}$$

When output orientation is selected, then the following NSBM model is used for the performance assessment of the DMUj_o:

$$\frac{1}{\theta_{j_0}^{o}} = max \sum_{\delta=1}^{\nu} w^{\delta}\left[1 + \frac{1}{s_\delta}\left(\sum_{r=1}^{s_\delta} \frac{s_r^{\delta+}}{y_{r_{jo}}^{\delta}}\right)\right] \tag{3.46}$$

$subject\ to$ (3.43a), (3.43b) or (3.43c)

The output oriented overall efficiency for DMUj_o is derived from the optimal value of the objective function of model (3.46). The authors in order to confine the efficiency scores into the range [0, 1], they expressed the output oriented stage efficiency scores using the optimal output slacks $s^{\delta+*}$ as:

$$\theta_{j_0}^{\delta} = \frac{1}{1 + \frac{1}{s_\delta}\left(\sum_{r=1}^{s_\delta} \frac{s_r^{\delta+*}}{y_{r_{jo}}^{\delta}}\right)}, \ \delta = 1, \dots, \nu \tag{3.47}$$

As can be deduced the NSBM output oriented overall efficiency is the weighted harmonic mean of the stage efficiency scores:

$$\frac{1}{\theta_{j_0}^{o}} = \sum_{\delta=1}^{\nu} \frac{w^{\delta}}{\theta_{j_0}^{\delta}} \tag{3.48}$$

In case non-orientation is selected, i.e. when both input and output slacks are taken into consideration in the assessment, then the non-oriented NSBM model is expressed as:

$$\zeta_{j_0}^{o} = min \frac{\sum_{\delta=1}^{\nu} w^{\delta}\left[1 - \frac{1}{m_\delta}\left(\sum_{i=1}^{m_\delta} \frac{s_i^{\delta-}}{x_{i_{jo}}^{\delta}}\right)\right]}{\sum_{\delta=1}^{\nu} w^{\delta}\left[1 + \frac{1}{s_\delta}\left(\sum_{r=1}^{s_\delta} \frac{s_r^{\delta+}}{y_{r_{jo}}^{\delta}}\right)\right]} \tag{3.49}$$

$subject\ to$ (3.43a), (3.43b) or (3.43c)

Given the optimal solution of model (3.49), then the non-oriented overall efficiency is straightforwadly derived from the objective function while the non-oriented stage efficiency scores are calculated as follows:

$$\zeta_{j_0}^{\delta} = min \frac{1 - \frac{1}{m_\delta}\left(\sum_{i=1}^{m_\delta} \frac{s_i^{\delta-*}}{x_{i_{jo}}^{\delta}}\right)}{1 + \frac{1}{s_\delta}\left(\sum_{r=1}^{s_\delta} \frac{s_r^{\delta+*}}{y_{r_{jo}}^{\delta}}\right)}, \qquad \delta = 1, \ldots, \nu \tag{3.50}$$

Once an optimal solution $(h^{\delta*}, s^{\delta-*}, s^{\delta+*})$ of models (3.44), (3.46) or (3.49) is obtained, then the projections onto the efficient frontier can be calsulated as follows:

$$\begin{aligned} X_{j_o}^{\delta*} &= X_{j_o}^{\delta} - s^{\delta-*} = X^{\delta} h^{\delta*}, \qquad \delta = 1, \ldots, \nu \\ Y_{j_o}^{\delta*} &= Y_{j_o}^{\delta} + s^{\delta+*} = Y^{\delta} h^{\delta*}, \qquad \delta = 1, \ldots, \nu \end{aligned} \tag{3.51}$$

If the *fixed link* case is used in the assessment, then the intermediate measures will remain unchanged to their intial levels. Otherwise, if the *free link* case is selected, then the projections of the intermediate measures are computed as follows:

$$Z_{j_o}^{(\delta,\psi)*} = Z^{(\delta,\psi)} h^{\delta*}, \ \forall\, (\delta, \psi) \tag{3.52}$$

From the above we conclude that in the non-oriented case the relationship between the overall efficiency and the stage efficiencies cannot be defined explicitly. Tone and Tsutsui (2009) noticed that alternative forms of the overall efficiency could be used in the non-oriented case. For instance, Lu et al (2014) modified the non-oriented NSBM (3.49) by deriving the the non-oriented overall efficiency as the simple arithmetic mean of the non-oriented stage efficiencies.

Notice that the above models are given under VRS assumption, however the CRS models can be also formed regardless the orientation by removing the corresponding convexity constraints. The experimentation of Tone and Tsutsui (2009) revealed that under CRS assumption and employing the *free link* case the NSBM may deem inefficient all the DMUs under evalution in each individual stage. This finding contradicts with the characteristics of traditional DEA models where at least one DMU is deemed efficient so as to construct the efficient frontier. On the other hand, the authors proved that under VRS assumption there is always at least an efficient DMU in each sub-process. As they further pointed out, this also holds when the *fixed link* case is utilized under CRS assumption.

Fukuyama and Mirdehghan (2012) showed, by providing adequate examples, that the NSBM of Tone and Tsutsui (2009) fails to identify the efficiency status of DMUs because the slacks concerning the intermediate measures are not considered in the definitions of the efficiencies. They proposed a revised PPS and a two-phase approach which identifies sufficiently the efficiency status under the *fixed link* case only. Mirdehghan and Fukuyama (2016) developed another two-phase approach by incorporating the notions of mathematical dominance, which deals effectively with the *free link* case also.

Chen et al (2013) noticed that the network DEA methods that are developed on the basis of the production possibility set, such as Tone and Tsutsui's (2009) slacks-based method should be re-examined with respect to the definition of the stage efficiencies. Especially, they discovered that the NSBM of Tone and Tsutsui (2009) provides only the overall efficiency when it is applied for the performance assessment of the two-stage processes of Type I (Fig. 3.1a). Chen et al (2013) argued that since the intermediate measures are the only outputs from stage-1 and the only inputs to stage-2, then neither the input oriented NSBM for stage-2 nor the output oriented NSBM for stage-1 can be formed. This relates, as noted above, with the absence of the slacks associated with the intermediate measures in the definitions of the efficiencies. They regarded this finding as a pitfall and they concluded that the NSBM models can only yield the overall efficiency of the Type I two-stage process. In Table 3.1 below we demonstrate the applicability of the NSBM on various types of network structures.

Table 3.1: Applicability of NSBM

Network Structure	Input Oriented	Output Oriented	Non-Oriented
Series - Type I	-	-	-
Series - Type II	✓	-	-
Series - Type III	-	✓	-
Series - Type IV	✓	✓	✓
Generalized Series	✓	✓	✓
Parallel	✓	✓	✓
Series-Parallel (Mixed)	✓	✓	✓

3.3.3 System-centric approach

The network DEA methods that are characterized as system-centric, they do not provide the stage efficiencies but only the overall efficiency of the DMU under evaluation. Most of these methods are modelled in the envelopment form which is based on the unification of the production possibility sets of the individual stages. Kao (2014b) referred to such methods as "*system distance measure*" methods, where an input or output oriented distance measure model is employed to measure the overall efficiency of each DMU.

Notice that most system-centric methods originate from the pioneer work of Fare (1991). Fare (1991) studied DMUs with the structure of Type IV (Fig. 3.1d) and combined the production technology of the two stages to derive the entire-expanded technology of the DMU. Their proposed model however yields only the overall efficiency of the DMU. Fare and Whittaker (1995) employed the approach of Fare (1991) for the performance assessment of 137 dairy farms in USA. Fare and Grosskopf (1996) studied the same network structure and they followed the same practice to formulate the system technology. They built upon Fare (1991) to construct Malmquist productivity indices (cf. Caves et al, 1982; Fare and Grosskopf, 1992a) to draw efficiency comparisons between periods. Fare and Grosskopf (2000), as mentioned above, unified the methods introduced by Fare (1991), Fare and Whittaker (1995), Fare and Grosskopf (1996) and Fare et al (1997) to a generalized framework for modelling various types of network structures.

3.4 Classification of network DEA studies

In this section we provide a thorough classification of network DEA studies involving theoretical developments and applications. They are basically categorized according to the assessment paradigm they follow i.e. independent, decomposition, slacks-based measure and system-centric.

Table 3.2 below presents the studies that are based on independent assessments. For each study we provide the reference, the network structure of the DMUs, the number of stages and the returns to scale assumed to form the production possibility set (PPS). We also indicate the studies that provide theoretical developments or these that consists of applications and we give a short description of the application field.

Table 3.2: Studies based on independent assessments

Reference	Network Structure	No of Stages	PPS - Returns to Scale	Theoretic	Application	Application Field
Fare and Primont (1984)	Parallel	ν≥2	VRS	✓	✓	US coal-fired steam electric generating plants
Charnes et al (1986)	Series-Type IV	2	CRS	✓	✓	US Military
Chilingerian and Sherman (1990)	Series-Type I	2	CRS	✓	✓	Medical services
Fare et al (1992b)	Parallel	ν≥2	CRS	✓		-
Fare and Primont (1993)	Parallel	ν≥2	VRS	✓		-
Wang et al (1997)	Series-Type I	2	VRS	✓	✓	IT on banks
Kao (1998)	Parallel	ν≥2	VRS		✓	Taiwanese forests
Seiford and Zhu (1999)	Series-Type I	2	CRS/VRS	✓	✓	US commercial banks
Soteriou and Zenios (1999)	Mixed	3	VRS	✓	✓	Branches of a Cyprus bank
Zhu (2000)	Series-Type I	2	CRS/VRS		✓	Fortune 500 companies
Keh and Chu (2003)	Series-Type I	2	VRS		✓	Grocery stores
Sexton and Lewis (2003)	Series-Type I	2	VRS	✓	✓	Teams of USA Major League Baseball
Luo (2003)	Series-Type I	2	CRS / VRS		✓	US large banks
Lewis and Sexton (2004)	Mixed	5	VRS	✓	✓	Teams of USA Major League Baseball
Abad et al (2004)	Series-Type II	2	VRS		✓	Stocks in the Spanish manufacturing industry
Keh et al (2006)	Series-Type I	2	VRS		✓	Asia–Pacific hotels
Lu (2009)	Series-Type I	2	CRS / VRS		✓	Taiwanese IC-design firms
Lo and Lu (2009)	Series-Type I	2	VRS		✓	Taiwanese financial holding companies
Lo (2010)	Series-Type I	2	VRS		✓	US S&P 500 firms
Tsolas (2011)	Series-Type I	2	VRS		✓	Greek commercial banks
Tsolas (2013)	Series-Type I	2	VRS		✓	Greek construction firms
Adler et al (2013)	Mixed	3	VRS	✓	✓	European airports

In Table 3.3 below the fifth column (*Model*) indicates whether the study is seminal or modification and extension of an existing one.

Table 3.3: Seminal efficiency decomposition approaches, modifications and extensions

Reference	**Network Structure**	**No of Stages**	**PPS - Returns to Scale**	**Model**	**Application Field**
Beasley (1995)	Parallel	2	CRS	Seminal Study	UK Chemistry and Physics departments
Mar Molinero (1996)	Parallel	2	CRS	Modification of Beasley (1995)	UK Chemistry and Physics departments
Tsai and Mar Molinero (2002)	Parallel	5	VRS	Extension of Mar Molinero (1996)	National Health Service trusts in England
Cook et al (2000)	Parallel	2	CRS	Modification of Beasley (1995)	Branches of a Canadian bank
Cook and Hababou (2001)	Parallel	2	VRS	Extension of Cook et al (2000)	Branches of a Canadian bank
Cook and Green (2004)	Parallel	4	CRS	Extension of Cook et al (2000), Cook et al (2001)	Manufacturing plants in steel industry
Jahanshahloo et al (2004)	Parallel	3	CRS	Extension of Cook et al (2000)	Branches of an Iranian bank
Amirteimoori and Kordrostami (2005)	Generalized Series	ν≥2	CRS	Extension of Beasley (1995) and Cook et al (2000)	Illustrative data
Amirteimoori and Shafiei (2006)	Generalized Series / Series – Type IV	ν≥2	CRS	Extension of Beasley (1995) and Cook et al (2000)	Illustrative data
Chen et al (2006)	Series –Type II	2	CRS	Extension of Tsai and Mar Molinero (2002)	IT on banks
Liang et al (2006)	Series –Type II	2	CRS	Extension of Tsai and Mar Molinero (2002)	Illustrative data on Supply Chains
Kao and Hwang (2008)	Series –Type I	2	CRS	Seminal Study	Non-life Insurance Companies in Taiwan
Liang et al (2008)	Series –Type I	2	CRS	Seminal Study	IT on banks / US commercial banks
Chen et al (2009b)	Series –Type I	2	CRS/VRS	Extension of Beasley (1995), Amirteimoori and Kordrostami (2005) and Amirteimoori and Shafiei (2006)	Non-life Insurance Companies in Taiwan
Kao (2009a)	Generalized Series / Parallel / Mixed	ν≥2	CRS	Extension of Beasley (1995), Cook et al (2000), Amirteimoori and Kordrostami (2005), Amirteimoori and Shafiei (2006) and Kao and Hwang (2008)	Non-life Insurance Companies in Taiwan / Illustrative data
Kao (2009b)	Parallel	ν≥2	CRS	Extension of Beasley (1995) and Cook et al (2000)	Taiwanese forests
Chen et al (2010a)	Series –Type I	2	CRS	Extension of Kao and Hwang (2008)	Non-life Insurance Companies in Taiwan
Cook et al (2010b)	Generalized Series / Mixed	ν>2	CRS	Extension of Chen et al (2009b)	Electric power companies / Illustrative data

Reference	Network Structure	No of Stages	PPS - Returns to Scale	Model	Application Field
Chen et al (2010b)	Series –Type II	2	VRS	Extension of Chen et al (2009b)	IT on banks
Zha and Liang (2010)	Series –Type II	2	CRS	Extension of Kao and Hwang (2008)	US commercial banks
Kao and Hwang (2010)	Generalized Series / Parallel / Mixed	ν≥2	CRS	Unification of Kao (2009a) and Kao (2009b)	IT on banks / Illustrative data
Kao and Hwang (2011)	Series –Type I	2	VRS	Modification of Kao and Hwang (2008)	-
Liang et al (2011)	Series with feedback	2	CRS	Extension of Chen et al (2006) and Liang et al (2006)	Chinese universities
Li et al (2012)	Series –Type II	2	CRS	Extension of Kao and Hwang (2008)	Regional R&D in China
Kao (2014a)	Generalized Series / Parallel / Mixed	ν≥2	CRS	Extension of Kao (2009a), Kao (2009b) and Kao and Hwang (2010)	Electric power companies / Illustrative data
Li et al (2015)	Series –Type I	2	VRS	Extension of Kao and Hwang (2008)	Nations in 2012 London summer Olympic Games
An et al (2016)	Series –Type I	2	CRS	Extension of Kao and Hwang (2008)	Non-life Insurance Companies in Taiwan

The following directed graphs depict the starting points and the advancements of the *multiplicative* and the *additive* efficiency decomposition methods. Each node represents one or more studies that constitute a milestone on each efficiency decomposition approach. The edges indicate relationship between studies, i.e. the direction of each edge points from the study used as theoretical basis to the study that extend this basis. By employing this representation method we highlight the development of the efficiency decomposition approaches and the knowledge flow paths. Notice that the colors on each node indicate the type of network structures that examined in each study.

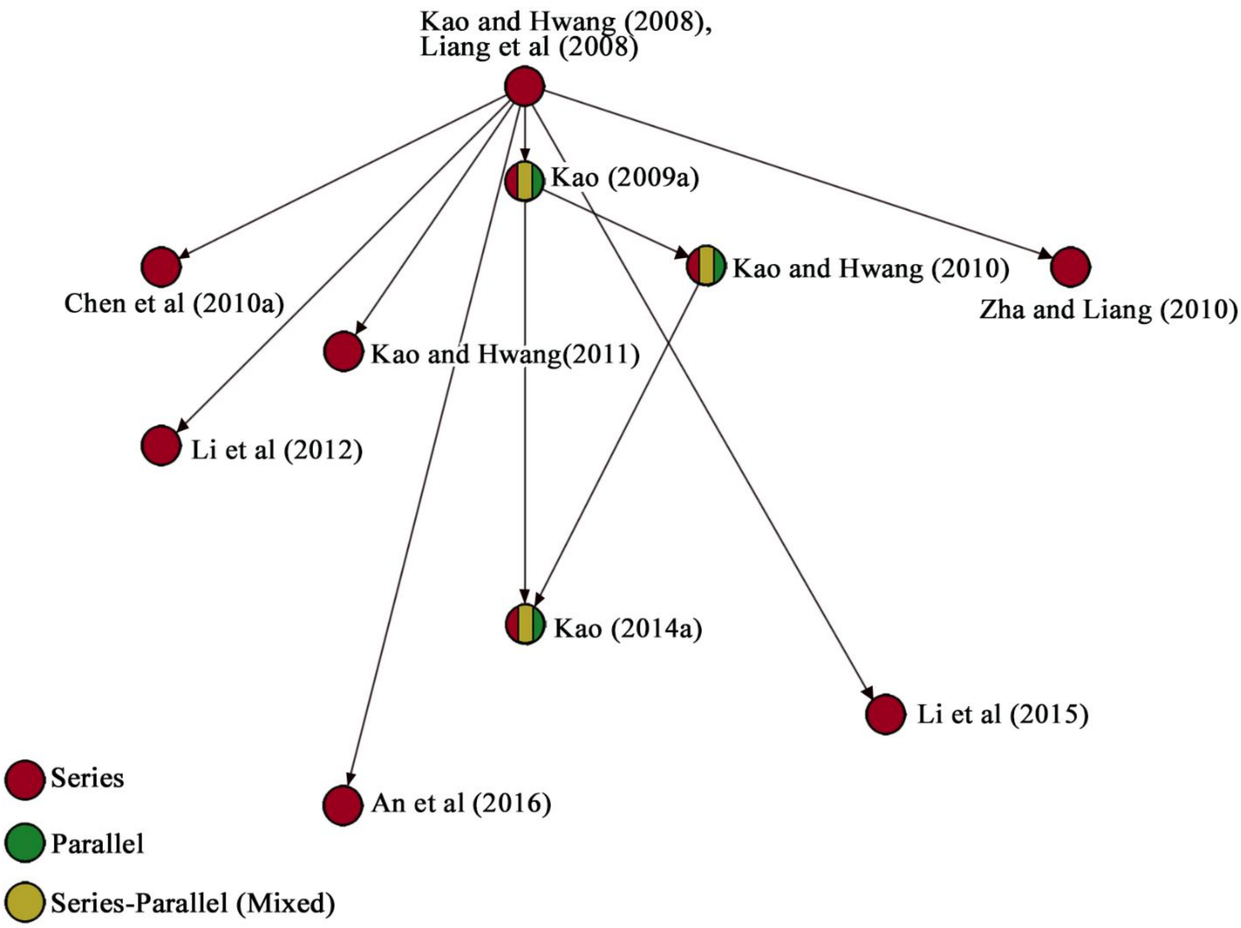


Fig. 3.10: Evolution of the multiplicative efficiency decomposition method

Below we provide the schematic representation of the evolution of the additive efficiency decomposition method.

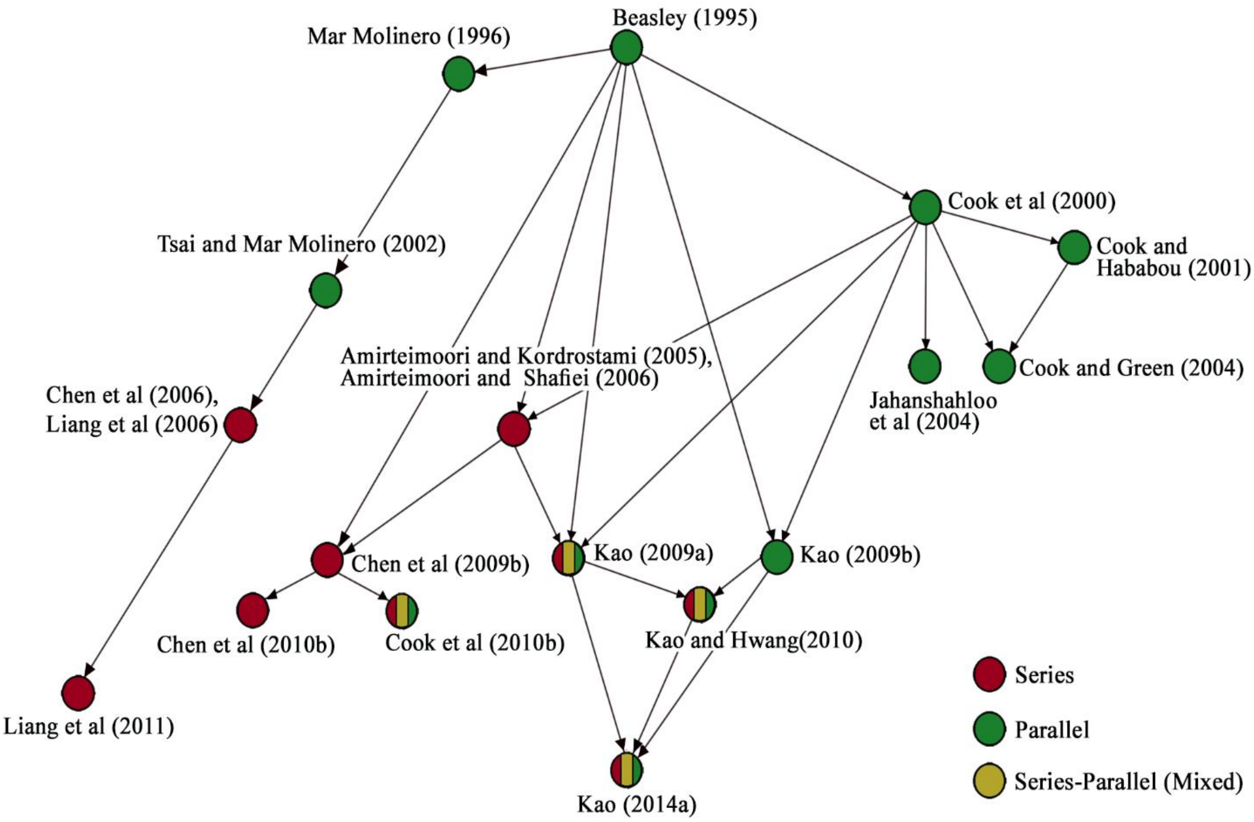


Fig. 3.11: Evolution of the additive efficiency decomposition method

The applications of the multiplicative efficiency decomposition method are presented in Table 3.4. The fifth column (*Model*) of Table 3.4 provides the model used in each study.

Table 3.4: Applications of the multiplicative efficiency decomposition method

Reference	Network Structure	No of Stages	PPS - Returns to Scale	Model	Application Field
Liu and Wang (2009)	Series –Type I	2	CRS	Kao and Hwang (2008)	Printed circuit board industry in Taiwan
Guan and Chen (2010)	Series –Type II	2	CRS	Kao (2009a)	High-tech innovations in Chinese provinces
Hsieh and Lin (2010)	Mixed	4	CRS	Kao (2009a)	International hotels in Taiwan
Cao and Yang (2011)	Series –Type I	2	CRS	Kao and Hwang (2008)	Internet companies
Zhu (2011)	Series –Type I	2	CRS	Kao and Hwang (2008)	Airlines
Lee and Johnson (2011)	Series –Type I	3	CRS	Kao (2009a)	Firms of semiconductor manufacturing industry
Lee and Johnson (2012)	Mixed	4	VRS	Kao (2009a)	US airlines

Reference	Network Structure	No of Stages	PPS - Returns to Scale	Model	Application Field
Chen et al (2012)	Series –Type I	2	CRS	Kao and Hwang (2008)	Automotive industry
Limaei (2013)	Series –Type I	2	CRS	Kao and Hwang (2008)	Iranian forests
Wanke (2013)	Series –Type I	2	CRS	Kao and Hwang (2008)	Brazilian ports
Wanke and Barros (2014)	Series –Type I	2	CRS	Kao and Hwang (2008)	Brazilian banks
Wanke et al (2016)	Series –Type I	2	CRS	Kao and Hwang (2008)	Australian public schools

Table 3.5 presents the applications of the additive efficiency decomposition method. Similar to Table 3.4, in Table 3.5 the fifth column (*Model*) reports the model used in each study.

Table 3.5: Applications of additive efficiency decomposition method

Reference	Network Structure	No of Stages	PPS - Returns to Scale	Model	Application Field
Diez-Ticio and Mancebon (2002)	Parallel	2	VRS	Tsai and Mar Molinero (2002)	Spanish police service
Yu (2008)	Parallel	2	VRS	Tsai and Mar Molinero (2002)	Taiwan's bus transit system
Yu and Fan (2009)	Mixed	3	CRS	Mar Molinero (1996)	Taiwan's bus transit system
Liu (2011)	Series –Type I	2	CRS	Chen et al (2009b)	Taiwanese financial holding companies
Guan and Chen (2012)	Series –Type IV	2	CRS / VRS	Chen et al (2009b)	Innovation activities of OECD countries
Premachandra et al (2012)	Series –Type II	2	VRS	Chen et al (2009b)	US mutual funds
Lu et al (2012)	Series –Type I	2	VRS	Chen et al (2009b)	US airlines
Kao (2012)	Parallel	2	CRS / VRS	Beasley (1995), Kao (2009b) and Kao (2010)	UK Chemistry and Physics departments
Rogge and Jaeger (2012)	Parallel	6	CRS	Mar Molinero (1996)	Solid waste in municipalities
Da Cruz et al (2013)	Parallel	2	CRS	Rogge and Jaeger (2012)	Water utilities
Amirteimoori (2013)	Series –Type IV	2	VRS	Chen et al (2009b)	Car distribution and service
Wang et al (2014)	Series –Type I	2	VRS	Chen et al (2009b)	Chinese commercial banks
Yang et al (2014)	Series –Type I	2	VRS	Chen et al (2009b)	National Basketball Association (NBA) teams
Toloo et al (2015)	Series –Type II	2	CRS	Chen et al (2006)	IT on banks / UK Chemistry and Physics departments

Reference	Network Structure	No of Stages	PPS - Returns to Scale	Model	Application Field
Halkos et al (2015a)	Series –Type I	2	VRS	Chen et al (2009b)	Secondary education in 65 countries
Halkos et al (2015b)	Series –Type I	2	VRS	Chen et al (2009b)	Sustainability of European regions
Halkos et al (2016)	Series –Type I	2	VRS	Chen et al (2009b)	Sustainable development of countries with advanced economies
Guo et al (2017)	Series –Type I / Type II	2	CRS	Liang et al (2006)	Non-life Insurance Companies in Taiwan / Regional R&D in China

The following table reports the studies that are based on the slacks-based measure approach. The fifth column (*Model*) of Table 3.6 below indicates whether the study is seminal, extension or application of an existing one.

Table 3.6: Studies based on the slacks-based measure approach

Reference	Network Structure	No of Stages	PPS - Returns to Scale	Model	Application Field
Tone and Tsutsui (2009)	Generalized Series / Mixed	ν>2	CRS / VRS	Seminal Study	Electric power companies
Avkiran (2009)	Mixed	3	VRS	Tone and Tsutsui (2009)	UAE domestic commercial banks
Yu (2010)	Mixed	3	CRS	Tone and Tsutsui (2009)	Domestic airports of Taiwan
Fukuyama and Weber (2010)	Series –Type I	2	CRS	Extension of Tone and Tsutsui (2009)	Japanese banks
Matthews (2013)	Generalized Series	3	VRS	Tone and Tsutsui (2009)	Chinese banks
Lin and Chiu (2013)	Mixed	4	VRS	Tone and Tsutsui (2009)	Taiwanese domestic banks
Akther et al (2013)	Series –Type I	2	CRS	Fukuyama and Weber (2010)	Bangladeshi banks
Lu et al (2014)	Series –Type I	2	VRS	Tone and Tsutsui (2009)	National innovation system among countries
Chang et al (2017)	Series –Type IV	2	VRS	Tone and Tsutsui (2009)	International cruise lines

Table 3.7 summarizes the studies that are characterized as system-centric. The fifth column (*Model*) of Table 3.7 indicates whether the study is seminal, extension or application of an existing one.

Table 3.7: Studies based on the system-centric approach

Reference	**Network Structure**	**No of Stages**	**PPS - Returns to Scale**	**Model**	**Application Field**
Fare (1991)	Series –Type IV	2	CRS	Seminal Study	-
Fare and Whittaker (1995)	Series –Type IV	2	VRS	Fare (1991)	US dairy farms
Fare and Grosskopf (1996)	Series –Type IV	2	CRS	Extension of Fare (1991)	-
Fare et al (1997)	Parallel	4	CRS	Seminal Study	US grain farms
Lothgren and Tambour (1999)	Series –Type IV	2	CRS	Fare and Grosskopf (1996)	Swedish pharmacies
Fare and Grosskopf (2000)	Generalized Series / Parallel / Mixed	$\nu \geq 2$	CRS	Unification of Fare (1991), Fare and Whittaker (1995), Fare and Grosskopf (1996) and Fare et al (1997)	-
Prieto and Zofio (2007)	Mixed	4	CRS	Fare and Grosskopf (2000)	OECD countries
Sheth et al (2007)	Series –Type I	2	VRS	Fare and Grosskopf (2000)	Bus routes in Virginia State of USA
Vaz et al (2010)	Parallel	5	VRS	Fare et al (1997)	Portuguese retail stores
Yang et al (2011)	Series	2	CRS	Seminal Study	Supply Chains / Branches of China Construction Bank
Chen and Yan (2011)	Series / Mixed	2 / 3	CRS	Seminal Study	Illustrative data on Supply Chains
Lozano et al (2013)	Series –Type IV	2	VRS	Fare and Grosskopf (2000)	Spanish airports
Wu et al (2016)	Series –Type IV	2	CRS	Seminal Study	Industrial production and pollution treatment of Chinese regions

Conclusion

The current chapter provides a detailed survey of the network DEA studies and reveals that a great volume of network DEA literature exists. We demonstrated the usefulness of network DEA and its advantages over the standard DEA for the assessment of multi-stage processes. The most influential network DEA approaches are the *efficiency decomposition approach* and the *slacks-based measure approach*. Thus, we presented in detail the most important network DEA methods of those categories and we discussed their extensions and modifications. We also reported their limitations concerning the returns to scale, the inconsistency between the multiplier and the envelopment models as well as the inadequate information that provide for the calculation of efficient projections. In addition, we reported most of the studies that apply the existing network DEA methods to real word problems. The network DEA studies were classified according to the model developed or used. We will revisit in the following chapters the additive and the multiplicative decomposition methods to show that the former yields biased efficiency results whereas they both provide non-unique stage efficiency scores. Then we will develop novel methods capable of overcoming these drawbacks.

Chapter 4

Composition versus decomposition in two-stage Network DEA: a reverse approach

Based on a reverse perspective on how to obtain and aggregate the stage efficiencies, we develop in this chapter the *composition approach* as opposed to the *decomposition approach* discussed in the previous chapter. Our novel approach overcomes the deficiencies of the decomposition methods, i.e. non-uniqueness of efficiency decomposition and bias. It is developed for the elementary two-stage process of Type I (Fig. 3.1a), whereas extensions of our concepts to more complex two-stage network processes will be presented in the next chapter.

Estimating the stage efficiencies of multi-stage processes simultaneously can be considered as a multi-objective optimization problem where the efficiency of each stage is treated as a separate objective function with their contradictory nature being taken into account. In section 4.1, we provide the basic concepts of multi-objective programming. In section 4.2, we discuss the major shortcoming of the multiplicative (Kao and Hwang, 2008) and the additive (Chen et al, 2009b) decomposition methods of providing non-unique efficiency scores. Also, we revisit the latter to show that the efficiency estimates are biased by unduly favoring one stage against the other. In section 4.3 we develop in detail our novel approach and we show that it effectively overcomes the shortcomings of the decomposition methods, i.e. it provides unique and unbiased stage efficiency scores. In section 4.4, we provide the results derived from our approach and we draw extensive comparisons with those obtained by the well-known methods on the literature. We apply our approach to experimental data as well as to test data drawn from the literature. We give also rigorous justifications for the similarities and the differentiations observed in the results.

As the conflicting role of the intermediate measures gives a peculiar character to two-stage processes that obscures the standard DEA premises, we introduce, in section 4.5, an envelopment model to derive the efficient frontier in two-stage DEA. It is linked to - and developed on the basis of - our primal multiplier efficiency assessment model. Furthermore, we propose an alternative two-phase method that projects the inefficient units on the frontier at a minimum distortion of the observed intermediate measures. The rationale for such a treatment is that the intermediate measures, conceived as a hidden layer in the production process, are the less controlled dimensions that should undergo changes at a minimum deviation from their observed values. Finally, concluding remarks are drawn in the last section of the chapter.

4.1 Basic concepts of multi-objective linear programming

Multi-objective programming problems are concerned with the optimization of multiple conflicting *objectives (criteria)*. When both the objective functions to be optimized and the constraints are linear then the multi-objective programming problem is called linear, MOLP in brief. MOLP and DEA are similar in structure, the relationships between them are explored, among others, by Golany (1988), Charnes et al (1989), Kornbluth (1991), Stewart (1996), Joro et al (1998) and Cooper (2005). The base of combining MOLP and DEA is the concept of Pareto efficiency which is present in both methods. Let the MOLP problem be given as follows:

$$\begin{aligned} & min\, f(a) = [f_1(a), \dots, f_k(a)] \\ & s.t. \\ & a \in A \end{aligned} \qquad (4.1)$$

where $f_h(a)$, h=1,..,k, are linear objective functions to be minimized and $A \neq \emptyset$ is a convex polyhedron that denotes the set of all feasible solutions in *decision (variables) space*. Let $C \subset R^n$, the image of *A*, denote the feasible region in *objective functions (criterion)* space, where $c \in C$ if and only if there exists $a \in A$ such that $c=(f_1(a),\dots,f_k(a))$. Single objective programming is studied in *decision space,* whereas in MOLP the attention is mostly focused in *objective functions* space. This is because the *objective functions* space usually is considerably smaller than the dimension of the *decision space* and the decision makers are mainly interested in the objective values. A MOLP problem rarely has a single optimal solution that simultaneously minimizes all objectives but possibly there exist an infinite number of optimal solutions called *efficient*, *non-dominated* or *Pareto optimal*. A solution of a MOLP problem is called *efficient*, *non-dominated* or *Pareto optimal* if there does not exist another feasible solution that improves the value of at least one objective function without deteriorating any other objective. In particular, a solution $a' \in A$ is *Pareto optimal* if and only if there does not exist another $a \in A$ such that $f_h(a) \leq f_h(a')$ for all h=1,..,k and $f_j(a) < f_j(a')$ for at least one objective *j*. Otherwise, if and only if there does not exist another $a \in A$ such that $f_h(a) < f_h(a')$ for all h=1,..,k, then the solution $a' \in A$ is *weakly Pareto optimal*.

4.1.1 Solution methods

Multi-objective programming methods are categorized according to the participation of the decision maker in the solution process (cf. Hwang and Masud, 1979). The methods that articulate preference information from the decision maker are classified as a priori, a posteriori and interactive. Also, there are methods that do not articulate preference information called no preference methods, whereby a neutral solution is generated. Most of them however, are simplifications of the a priori methods by typically excluding the parameters imposed by the articulation of preferences. A large variety of methods have been developed within the aforementioned classes of methods for solving multi-objective programming problems, see Steuer (1986) and Kaliszewski (2004). In our context though, we selected the *scalarization* method to solve MOLP problems. *Scalarization* means to convert the MOLP problem to a single objective LP, whose single objective function is termed *scalarizing function* (cf. Miettinen and Makela, 2002). The aim is to establish relations between the set of optimal solutions of the scalarized problem and the set of Pareto solutions of the MOLP. The most widely used scalarization methods are the additive aggregation of the objective functions (*weighted sum* method) and the *weighted Tchebycheff* or *weighted min-max* method. In the *weighted Tchebycheff* method the distance between some reference point and the feasible objective functions (criterion) space is minimized using the $L\infty$ *norm*.

Weighted Sum method

The MOLP (4.1) can be transformed to the following single objective LP (4.2) via the *weighted sum* methodology which is introduced by Gass and Saaty (1955). The single objective function of model (4.2) is constructed by the sum of the objective functions $f_h(a)$, h=1,..,k, multiplied by the weighting factor t_h, h=1,..,k, that reflects the relative importance of each objective. The theory of the *weighted sum* method is covered in detail by Ehrgott (2005).

$$\begin{aligned} & \min \sum_{h=1}^{k} t_h f_h(a) \\ & s.t. \\ & a \in A \end{aligned} \tag{4.2}$$

An optimal solution α' of the scalar LP model (4.2) is a Pareto optimal (non-dominated) solution to MOLP (4.1), if and only if there are $\{t_h > 0, h = 1, \dots, k \,/\, \sum_{h=1}^{k} t_h\}$, setting one or more of the weights to zero may result in weak Pareto optimal solution. The relations between nonnegative weights and Pareto optimality are examined by Lin (1976). Alternative Pareto optimal solutions can be obtained by changing the weights systematically. However, varying the weights will not necessarily change the solution since altering the weights will only provide extreme points (vertices) on the Pareto front, i.e. the solution jumps from one extreme point to another. The means to generate the whole Pareto optimal set are explored comprehensively by Censor (1977) and Chankong and Haimes (1983).

A special case of the *weighted sum* method results when equal importance is given to the objective functions or equivalently no preference among them exists, i.e. t_h=1, h=1,..,k. In this case, the *preference-free scalarizing* function is simply built by the sum of the objective functions.

Weighted Tchebycheff method

The *weighted Tchebycheff* or *weighted min-max scalarization* method belongs to the class of compromise programming methods (cf. Zeleny, 1973), it can be originally found in Bowman (1976) and it is also utilized in the milestone methods of Choo and Atkins (1980) and Steuer and Choo (1983). The *weighted Tchebycheff* method is based on the concept of minimizing the distance to a given reference point utilizing the $L\infty$ norm. In particular, the Tchebycheff norm minimization chooses the corner closest to the given reference point and still in contact with the feasible region. A reference point (cf. Wierzbicki, 1980) consists of aspiration levels (objective function values) that are desirable for the decision maker or can be any reasonable point in the objective space. These points can be reservation points that must be attained or exceeded so as to be considered acceptable (cf. Reeves and MacLeod, 1999) or worst outcome points that should be avoided (cf. Michalowski and Szapiro, 1992). In the frame of our approach, we use the *ideal point* as a reference point which represents, in the objective functions space, the ideal solution that simultaneously optimizes each objective separately. The ideal solution (best possible attainment) is obtained by optimizing each of the objective functions individually subject to the feasible region. In principle, it is reasonable when forming a measure of distance to seek for a point that is as close as possible to the *ideal* one. The MOLP (4.1) is *scalarized* via the *weighted min-max* methodology using the ideal point as follows:

$$\begin{aligned} & \min \max_{h=1,\ldots,k} \; [t_h(|f_h(a) - f_h^*|)] \\ & s.t. \\ & a \in A \end{aligned} \tag{4.3}$$

where t_h, h=1,…,k, is the vector of the weights that reflect the relative importance of each objective and f_h^*, h=1,…,k, are the components of the ideal objective vector that constitutes the *ideal point* in the objective functions space. Model (4.3) yields at least one solution that is Pareto optimal for the MOLP (4.1). If the optimal solution of model (4.3) is unique then it is a Pareto optimal solution to MOLP (4.1). In general, every optimal solution of model (4.3), with positive weights ($t_h>0$, h=1,...,k), is *weakly* Pareto optimal to MOLP (4.1) (cf. Yu, 1973 and Kaliszewski, 1994). However, in the case of two objectives ($k=2$), given a set of positive weights ($t_h>0$, h=1,…,k), the optimal solution of model (4.3) is unique and thus a Pareto optimal (non-dominated) solution to MOLP (4.1) (cf. Ballestero and Romero, 1991). The *weighted Tchebycheff* method, contrary to the *weighted sum* method, can generate the entire Pareto optimal set of the MOLP (4.1) with variation of the weights t_h, i.e. it also provides the non-extreme points on the non-dominated surface (cf. Olson, 1993).

When all the objective functions are thought to be equally important or equivalently *no preference* among them exists, i.e. t_h=1, h=1,…,k, then a special case of the *weighted Tchebycheff* method occurs, namely the *unweighted Tchebycheff* method. Under this assumption, problem (4.3) is also called *method of the global criterion*.

4.2 Criticism of the efficiency decomposition methods

In order to discuss and reveal the deficiencies of the decomposition methods it is sufficient to refer to the elementary two-stage process of Type I. In this two-stage process some external inputs *X* are transformed to final outputs *Y* via the intermediate measures *Z*.

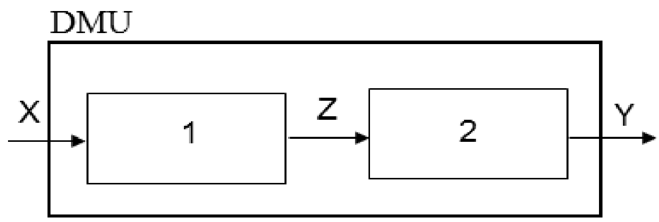


Fig. 4.1: The two-stage process of Type I

Assume n DMUs (j=1,...,n), each using m external inputs x_{ij}, $i=1,\dots,m$ in the first stage to produce q outputs z_{pj}, p=1,...,q from that stage. The outputs obtained from the first stage are then used as inputs to the second stage to produce s final outputs y_{rj}, $r=1,\dots,s$. In this basic setting, nothing but the external inputs to the first stage enters the system and nothing but the outputs of the second stage leaves the system.

Throughout this chapter we use the following notation:

$j \in J = \{1, \dots, n\}$: The index set of the n DMUs.

$j_0 \in J$: Denotes the evaluated DMU.

$X_j = (x_{ij}, i = 1, \dots, m)$: The vector of stage-1 external inputs used by DMU$_j$.

$Z_j = (z_{pj}, p = 1, \dots, q)$: The vector of intermediate measures for DMU$_j$.

$Y_j = (y_{rj}, r = 1, \dots, s)$: The vector of stage-2 final outputs produced by DMU$_j$.

$\eta = (\eta_1, \dots, \eta_m)$: The vector of weights for the stage-1 external inputs in the fractional model.

$v = (v_1, \dots, v_m)$: The vector of weights for the stage-1 external inputs in the linear model.

$\varphi = (\varphi_1, \dots, \varphi_q)$: The vector of weights for the intermediate measures in the fractional model.

$w = (w_1, \dots, w_q)$: The vector of weights for the intermediate measures in the linear model.

$\omega = (\omega_1, \dots, \omega_s)$: The vector of weights for the stage-2 outputs in the fractional model.

$u = (u_1, \dots, u_s)$: The vector of weights for the stage-2 outputs in the linear model.

e_j^o: The overall efficiency of DMU$_j$.

e_j^1: The efficiency of the first stage for DMU$_j$.

e_j^2: The efficiency of the second stage for DMU$_j$.

E_j^1: The independent efficiency score of the first stage for DMU$_j$.

E_j^2: The independent efficiency score of the first stage for DMU$_j$.

λ: The intensity vector for the first stage.

μ: The intensity vector for the second stage.

s^-: The vector of the input excesses.

s^+: The vector of the output shortfalls.

4.2.1 Non-unique efficiency scores

Both the multiplicative (Kao and Hwang, 2008) and the additive (Chen et al, 2009b) decomposition methods are developed on the basis of the two-stage process of Type I. As noticed in Chapter 3, in both methods the decomposition of the overall efficiency to the stage efficiencies is non-unique. Thus in both studies, similar post-optimality procedures were developed to derive extreme stage efficiency scores, maintaining the overall efficiency obtained from the decomposition models. Kao (2016) noticed that is critical to "*identify the most influential divisions that have decisive effects on the overall efficiency of the system*" because by improving these stages the system efficiency will be improved. However, this cannot be exercised safely due to the non-uniqueness of the efficiency scores, i.e. alternative efficiency decompositions deem different stage as influential. The shortcoming of non-unique efficiency scores of the decomposition methods may occur in the assessment of any type of network structure.

4.2.2 Biased efficiency scores

The additive efficiency decomposition method also suffers from biased efficiency results as we show below. We recall that the modelling approach of Chen et al (2009b) facilitates the linearization of a non-linear mathematical program by assuming that the weights of the two stages derive endogenously by the optimization process. We provide again the definitions of the overall and stage efficiencies of the two-stage process of Type I (Fig. 4.1).

$$e_j^o = \frac{uY_j + wZ_j}{vX_j + wZ_j}, e_j^1 = \frac{wZ_j}{vX_j}, e_j^2 = \frac{uY_j}{wZ_j} \quad (4.4)$$

In additive efficiency decomposition method the overall efficiency is estimated as a weighted average of the stage efficiencies. The additive decomposition model and the definitions of the weights of the stages are expressed as follows:

$$e_j^o = \frac{uY_j + wZ_j}{vX_j + wZ_j} = t_j^1 \frac{wZ_j}{vX_j} + t_j^2 \frac{uY_j}{wZ_j},\ t_j^1 + t_j^2 = 1$$

$$t_j^1 = \frac{vX_j}{vX_j + wZ_j}, t_j^2 = \frac{wZ_j}{vX_j + wZ_j}$$

Given the above definitions, the model below assesses the overall efficiency of the evaluated unit j_0:

$$\begin{aligned} & e_{jo}^o = max\, uY_{j_o} + wZ_{j_o} \\ & s.t. \\ & vX_{j_o} + wZ_{j_o} = 1 \\ & wZ_j - vX_j \leq 0,\ \ j = 1, \dots, n \\ & uY_j - wZ_j \leq 0,\ \ j = 1, \dots, n \\ & v \geq 0, w \geq 0, u \geq 0 \end{aligned} \quad (4.5)$$

Once an optimal solution (v^*, w^*, u^*) of model (4.5) is obtained, the overall efficiency and the stage efficiencies are calculated by the following relationships:

$$e_{jo}^{o} = u^{*}Y_{j_o} + w^{*}Z_{j_o}$$

$$t_{jo}^{1} = v^{*}X_{j_o}\,,\; t_{jo}^{2} = w^{*}Z_{j_o} \qquad (4.6)$$

$$e_{jo}^{1} = \frac{w^{*}Z_{j_o}}{v^{*}X_{j_o}}\,,\; e_{jo}^{2} = \frac{e_{jo}^{0} - t_{jo}^{1}\,e_{jo}^{1}}{t_{jo}^{2}} = \frac{u^{*}Y_{j_o}}{w^{*}Z_{j_o}}$$

The argument given in Chen et al (2009b) for the weights t_j^1 and t_j^2 is that they represent the relative contribution of the two stages to the overall performance of the DMU. The “size” of each stage, as measured by the portion of total resources devoted to each stage, is assumed to reflect their relative contribution to the overall efficiency of the DMU. However, as long as the weights derive from the optimization process, they depend on the DMU being evaluated and, generally, they are different for different DMUs. Thus, the “size” of a stage is not an objective reality, as it is viewed differently from each DMU. But this is not the only peculiarity emerging from the definition of the weights. Indeed, from the above relationships and the definition of the weights we derive that the additive decomposition method biases the efficiency assessments in favor of the second stage:

$$\frac{t_j^2}{t_j^1} = \frac{uY_j}{wZ_j} = e_j^1 \leq 1$$

i.e. $t_j^2 \leq t_j^1$, which is a major a shortcoming. Indeed, the maximum value that t_j^2 can attain is 0.5 and e_j^2 increases (e_j^1 decreases) as t_j^2 decreases. As long as the individual efficiency scores are biased, the overall efficiency score is biased as well. In conclusion, the endogenous weights assumed in Chen et al (2009b) for the individual stages favor the second stage against the first one.

Notice, that the above finding is based upon an input oriented framework, though it is still valid in the output oriented case. Also, this conclusion can be easily drawn for other types of series multi-stage processes, regardless of the number of stages. Specifically, when the additive decomposition method is applied to multi-stage processes of Type I, under both input and output orientations, then it suffers from biased efficiency assessments. Also, when it is applied to multi-stage processes of Type III (Fig. 3.1c), then the aforementioned shortcoming is reported only if input orientation is chosen. When the additive method is applied to parallel network structures (Fig. 3.3) or under the VRS assumption of any type of network structures, then we cannot predetermine the relationship of the weights of the stages.

4.3 The composition approach to two-stage network DEA

Unlike the decomposition methods presented in the previous chapter, our method does not require an a priori definition of the overall efficiency. This grants our approach the flexibility to select the aggregation method a posteriori. Thus, we call our method the "*composition approach*" as opposed to the "*decomposition approach*". Similarly to the other methods, we assume that the weights associated with the intermediate measures are the same and we define the efficiency of the two stages as follows:

$$\hat{e}_j^1 = \frac{\varphi Z_j}{\eta X_j}, \hat{e}_j^2 = \frac{\omega Y_j}{\varphi Z_j}$$

4.3.1 Constant returns to scale

Consider the reciprocal of model (3.1) that is the output-oriented CRS-DEA model for the first-stage and the input-oriented CRS-DEA model (3.2) for the second-stage, where the same intermediate weights are assumed for both stages:

Stage I: Output-oriented

$$\begin{aligned} & min \frac{\eta X_{j_0}}{\varphi Z_{j_0}} \\ & s.t. \\ & \frac{\eta X_j}{\varphi Z_j} \geq 1, \ \ j = 1, \dots, n \\ & \eta \geq 0, \varphi \geq 0 \end{aligned} \tag{4.7}$$

Stage II: Input-oriented

$$\begin{aligned} & max \frac{\omega Y_{j_0}}{\varphi Z_{j_0}} \\ & s.t. \\ & \frac{\omega Y_j}{\varphi Z_j} \leq 1, \ \ j = 1, \dots, n \\ & \varphi \geq 0, \omega \geq 0 \end{aligned} \tag{4.8}$$

As mentioned earlier, models (4.7) and (4.8) provide the independent efficiency scores $1/E_{jo}^1, E_{jo}^2$ for the first and the second stage respectively. Appending the constraints of model (4.7) to model (4.8) and vice versa we get the following augmented models (4.9) and (4.10) for the first and the second stage respectively:

Stage I: Output-oriented

$$\begin{aligned} & min\frac{\eta X_{j_0}}{\varphi Z_{j_0}} \\ & s.t. \\ & \frac{\eta X_j}{\varphi Z_j} \geq 1,\ \ j=1,\dots,n \\ & \frac{\omega Y_j}{\varphi Z_j} \leq 1,\ \ j=1,\dots,n \\ & \eta \geq 0, \varphi \geq 0, \omega \geq 0 \end{aligned} \qquad (4.9)$$

Stage II: Input-oriented

$$\begin{aligned} & max\frac{\omega Y_{j_0}}{\varphi Z_{j_0}} \\ & s.t. \\ & \frac{\omega Y_j}{\varphi Z_j} \leq 1,\ \ j=1,\dots,n \\ & \frac{\eta X_j}{\varphi Z_j} \geq 1,\ \ j=1,\dots,n \\ & \eta \geq 0, \varphi \geq 0, \omega \geq 0 \end{aligned} \qquad (4.10)$$

Notice that an optimal solution of model (4.7) is also optimal in model (4.9). Indeed, one can always choose small enough values for ω in model (4.9) to make any optimal solution of model (4.7) feasible, yet optimal, in model (4.9). Analogously, an optimal solution of model (4.8) is also optimal in model (4.10), as one can choose large enough values for η in model (4.10) to make any optimal solution of model (4.8) feasible, yet optimal, in model (4.10).

Theorem 4.1: An optimal solution of model (4.7) is also optimal in model (4.9).

Proof:

Let η^* and φ^* be an optimal solution of (4.7). First we will show that this solution is feasible in (4.9). Indeed, it satisfies the first set of constraints of (4.9), as it is identical to the constraints in (4.7). Notice that the first set of constraints of (4.9) is independent of the variables ω, which appear only in the second set of constraints. Then,

(a) If the number of outputs (Y) is lower or equal to the number of intermediate measures (Z), i.e. $s \leq q$, then the second set of constraints of (4.9) is satisfied for

$$\omega_r = \frac{\varphi_r^* z_r^{min}}{y_r^{max}} \geq 0, r = 1, \dots, s$$

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

(b) If the number of outputs (Y) is greater than the number of intermediate measures (Z), i.e. $s>q$, the second set of constraints of (4.9) is satisfied for

$$\omega_p = \frac{\varphi_p^* z_p^{min}}{y_p^{max}} \geq 0, p = 1, \ldots, q, \omega_{r=}0, r = q+1, \ldots, s$$

Thus, the optimal solution η^* and φ^* of (4.7) is a feasible solution of (4.9). Moreover, as the objective functions in both the (4.7) and (4.9) are independent of ω, the above solution is optimal in (4.9) as well. □

Theorem 4.2: An optimal solution of model (4.8) is also optimal in model (4.10).

Proof:

Let ω^* and φ^* be an optimal solution of (4.8). First we will show that this solution is feasible in (4.10). Indeed, it satisfies the first set of constraints of (4.10), as it is identical to the constraints in (4.8). Notice that the first set of constraints of (4.10) are independent of the variables η, which appear only in the second set of constraints. Then,

(a) If the number of intermediate measures (Z) is lower or equal to the number of inputs (X), i.e. $q \leq m$, the second set of constraints of (4.10) is satisfied for

$$\eta_p = \frac{\varphi_p^* z_p^{max}}{x_p^{min}} \geq 0, p = 1, \ldots, q, \eta_{i\geq}0, i = q+1, \ldots, m$$

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

(b) If the number of intermediate measures (Z) is greater than the number of inputs (X), i.e. $q>m$, the second set of constraints of (4.10) is satisfied for

$$\eta_i = \frac{\varphi_i^* z_i^{max}}{x_i^{min}}, i = 1, \ldots, m-1$$

$$\eta_m = \frac{\varphi_m^* z_m^{max}}{x_m^{min}} + \sum_{p=m+1}^{q} \frac{\varphi_m^* z_m^{max}}{x_m^{min}}$$

Thus, the optimal solution ω^* and φ^* of (4.8) is a feasible solution of (4.10). Moreover, as the objective functions in both the (4.8) and (4.10) are independent of η, the above solution is optimal in (4.10) as well. □

Models (4.9) and (4.10) have common constraints and, thus, can be jointly considered as a bi-objective program:

$$
\begin{aligned}
& min \frac{\eta X_{j_0}}{\varphi Z_{j_0}} \\
& max \frac{\omega Y_{j_0}}{\varphi Z_{j_0}} \\
& s.t. \\
& \frac{\eta X_j}{\varphi Z_j} \geq 1, \;\; j = 1, \dots, n \\
& \frac{\omega Y_j}{\varphi Z_j} \leq 1, \;\; j = 1, \dots, n \\
& \eta \geq 0, \varphi \geq 0, \omega \geq 0
\end{aligned}
\tag{4.11}
$$

Applying the C-C transformation, model (4.11) can be formulated and solved as a MOLP. The correspondence of variables is $v=\tau\eta$, $u=\tau\omega$, $w=\tau\varphi$ where τ is a scalar variable such that $\tau\varphi Z_{j_0} = 1$.

$$
\begin{aligned}
& E_{jo}^1 = min\, vX_{j_0} \\
& E_{jo}^2 = max\, uY_{j_0} \\
& s.t. \\
& wZ_{j_0} = 1 \\
& wZ_j - vX_j \leq 0, \;\; j = 1, \dots, n \\
& uY_j - wZ_j \leq 0, \;\; j = 1, \dots, n \\
& v \geq 0, w \geq 0, u \geq 0
\end{aligned}
\tag{4.12}
$$

Optimizing the first and the second objective function separately one gets the independent efficiency scores of the two stages $\left(1/E_{jo}^1 \leq 1, E_{jo}^2 \leq 1\right)$. In terms of MOLP, the vector $\left(E_{jo}^1 \geq 1, E_{jo}^2 \leq 1\right)$ constitutes the ideal point of the bi-objective program (4.12) in the objective functions space. Thus, the efficiencies of the two stages can be obtained by solving the MOLP (4.12). However, as the ideal point is not generally attainable, solving a MOLP means finding non-dominated feasible solutions in the variable space that are mapped on the

Pareto front in the objective functions space, i.e. solutions that they cannot be altered to increase the value of one objective function without decreasing the value of at least one other objective function. As already noticed, a usual approach in solving a MOLP is the scalarizing approach, which transforms the MOLP in a single objective LP, whose optimal solution is a Pareto optimal (non-dominated) solution of the MOLP. Aggregating additively the objective functions and introducing a distance function are two alternative methods to build the scalarizing function. We present both cases in the following, as they possess different properties.

Firstly, by aggregating the two objective functions of MOLP (4.12) additively without giving any priority (no preference) to the objectives we derive the following single objective LP. In particular, we employ the special case of the *weighted sum* scalarizing method with all weights equal to one, i.e. $t_h=1$, $h=1,...,k$.

$$\begin{aligned} & min\, vX_{j_0} - uY_{j_0} \\ & s.t. \\ & wZ_{j_0} = 1 \\ & wZ_j - vX_j \leq 0, \;\; j = 1, \ldots, n \\ & uY_j - wZ_j \leq 0, \;\; j = 1, \ldots, n \\ & v \geq 0, w \geq 0, u \geq 0 \end{aligned} \qquad (4.13)$$

Once an optimal solution (v^*, w^*, u^*) of model (4.13) is obtained, the efficiency scores for unit j_0 in the first and the second stage are respectively:

$$\hat{e}^1_{j_0} = \frac{w^* Z_{j_0}}{v^* X_{j_0}} = \frac{1}{v^* X_{j_0}}, \hat{e}^2_{j_0} = \frac{u^* Y_{j_0}}{w^* Z_{j_0}} = u^* Y_{j_0} \qquad (4.14)$$

The optimal value of the objective function in (4.13) is $v^* X_{j_0} - u^* Y_{j_0} \geq 0$. The unit j_0 is efficient in both stages and, thus, overall efficient, if and only if the optimal value of the objective function is zero. Otherwise it is overall inefficient. Indeed, if $v^* X_{j_0} - u^* Y_{j_0} = 0$ then, as $w^* Z_{j_0} = 1$ and $uY_j \leq wZ_j \leq vX_j$ for every j, it holds that $v^* X_{j_0} = w^* Z_{j_0} = u^* Y_{j_0} = 1$, i.e. $\hat{e}^1_{j_0} = 1, \hat{e}^2_{j_0} = 1$. Model (4.13) does not provide a direct measure of the overall efficiency, as it is the case in the multiplicative model (3.7) and the additive model (4.5), but it does discriminate among overall efficient and inefficient units, a property that is closely related to the standard additive DEA model. However, it is the normalization constraint $wZ_{j_0} = 1$, on the intermediate measures in (4.13), that allows us to infer on the efficiency scores of the individual stages, as given in (4.14). This is the key that enables us to assess the efficiencies of the two stages simultaneously without the need to assume weights for the two stages.

The optimal solution (v^*, w^*, u^*) of model (4.13) is a Pareto optimal solution of the MOLP (4.12) and the optimal vector $\left(v^* X_{j_o}, u^* Y_{j_o}\right)$ is a non-dominated point on the Pareto front in the objective functions space of (4.12). This is a direct implication of the Geoffrion's (1968) theorem, which states that: given a multi-objective LP model $\{min\, f_h(a), h = 1, \dots, k\ /\ a \in A, a \geq 0\}$, a^* is a Pareto-optimal (efficient) solution for this model if and only if there are $\left\{t_h > 0, h = 1, \dots, k\ /\ \sum_{h=1}^{k} t_h = 1\right\}$ such that a^* is optimal for the scalar LP model $\left\{min \sum_{h=1}^{k} t_h\, f_h(a)\ /\ a \in A, a \geq 0\right\}$. Getting advantage of this property, one can scan the Pareto front and get alternative Pareto optimal solutions by solving model (4.15), i.e. the weighted counterpart of model (4.13), for different values of the parameter *t* with 0 < *t* < 1:

$$
\begin{aligned}
& min\, tvX_{j_0} - (1-t)uY_{j_0} \\
& s.t. \\
& wZ_{j_0} = 1 \\
& wZ_j - vX_j \leq 0, \ \ j = 1, \dots, n \\
& uY_j - wZ_j \leq 0, \ \ j = 1, \dots, n \\
& v \geq 0, w \geq 0, u \geq 0
\end{aligned}
\tag{4.15}
$$

We note that model (4.15) provides only extreme points on the Pareto front i.e. the optimal solutions are confined to vertices of the efficient region only. Notice also that the same Pareto optimal point can be obtained for a range of values of *t*, the so called indifference range. Thus, the solution obtained from model (4.13) by way of its unweighted scalar objective function can be obtained as well by giving different priorities (weights) to the two terms of the objective function within their indifference range (Steuer, 1986). Figure 4.2 below, is a general representation of the objective functions space of the MOLP (4.12) for an evaluated unit (X_o, Z_o, Y_o). Actually, it is the plane in the three-dimensional space (*vX*, *wZ*, *uY*) that is vertical to the axis *wZ* at wZ_o=1. The point (E^1, E^2) represents the ideal point, whereas the points *A*, *B*, *C* and *D* are the alternative Pareto optimal extreme points derived by the parametric model (4.15) for different values of the parameter *t*. The crooked line ABCD represents the Pareto front in the objective functions space. The dotted line passing from the point B has slope 1 and depicts the objective function of model (4.13), which, when minimized for the optimal solution (v^*, w^*, u^*), takes the non-negative value $v^*X_{j_o} - u^*Y_{j_o} =$ b > 0 and locates the point B on the Pareto front.

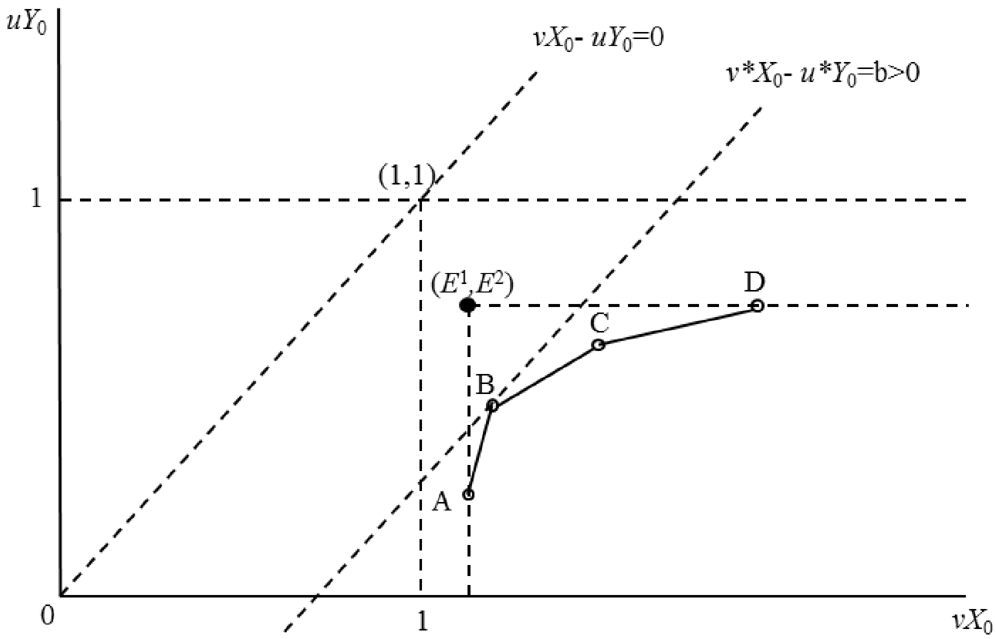


Fig. 4.2: The Pareto front of MOLP (4.12) and the optimal solution of model (4.13)

Although it is not very likely to occur in practice, the Pareto optimal point derived by model (4.13) and, thus, the efficiency scores of the two stages might be non-unique. This is the case where a segment of the Pareto front has slope 1, i.e. when it is parallel to the objective function line. For example, if the segment BC defined by the two successive Pareto optimal points B and C was parallel to the objective function line, then B, C and any convex combination of them would be optimal in terms of model (4.13). The uniqueness of the Pareto optimal point $\left(v^*X_{j_o}, u^*Y_{j_o}\right)$ and, thus, the uniqueness of the optimal efficiency scores of the two stages derived by model (4.13), can be tested by minimizing vX_0 and maximizing uY_0 subject to the constraints of (4.13) plus the constraint $vX_{j_o} - uY_{j_o} \leq v^*X_{j_o} - u^*Y_{j_o}$.

Model (4.13) is equivalent to finding an optimal solution that locates a point on the Pareto front at a minimum sum of the deviations $vX_{j_o} - 1$ and $1 - uY_{j_o}$ (L_1 norm) of (vX_{j_o}, uY_{j_o}) from the boundary point (1,1) in the objective functions space. Next, we employ the unweighted Tchebycheff norm (L_∞ norm) to locate a unique solution on the Pareto front by minimizing the maximum of the deviations $vX_{j_o} - E^1_{j_o}$ and $E^2_{j_o} - uY_{j_o}$ of (vX_{j_o}, uY_{j_o}) from the ideal point $\left(E_{j_o}^1, E_{j_o}^2\right)$. This is accomplished by the following min-max model, where δ denotes the largest deviation:

$$
\begin{aligned}
& min\,\delta \\
& s.t. \\
& vX_{j_0} - \delta \leq E_{jo}^1 \\
& uY_{j_0} + \delta \geq E_{jo}^2 \\
& wZ_{j_0} = 1 \\
& wZ_j - vX_j \leq 0, \;\; j = 1, \ldots, n \\
& uY_j - wZ_j \leq 0, \;\; j = 1, \ldots, n \\
& v \geq 0, w \geq 0, u \geq 0, \delta \geq 0
\end{aligned}
\qquad (4.16)
$$

Solving model (4.16) means searching for a solution where the deviations from the ideal point are equal and minimized. As depicted in Fig.4.3, the min-max solution is point *D*, being the intersection of the Pareto front and a ray from the ideal point (E^1,E^2) with slope (-1). The main advantage of model (4.16) over model (4.13) and the decomposition models (3.7) and (4.5) is that it provides a unique point, not necessarily extreme (vertex), on the Pareto front, i.e. unique efficiency scores for the two stages. Once an optimal solution (v^*, w^*, u^*) of model (4.16) is obtained, the stage efficiency scores for unit j_0 are as in (4.14).

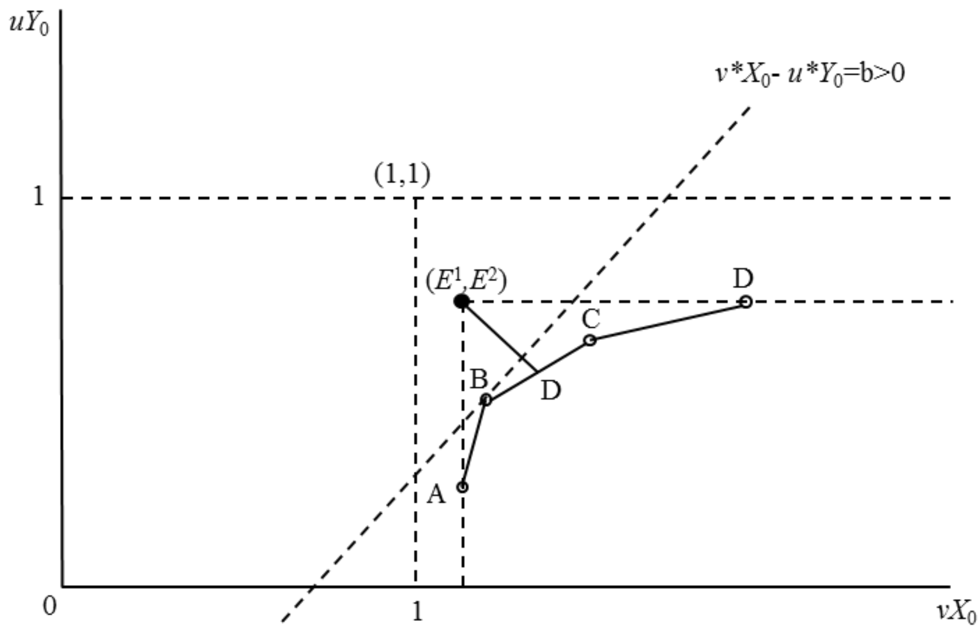


Fig. 4.3: The Pareto front of MOLP (4.12) and the optimal solution of model (4.16)

Considering the weighted Tchebycheff distance, the following parametric min-max model searches for a solution where the weighted deviations $t\left(vX_{j_o} - E_{j_o}^1\right)$ and $(1-t)\left(E_{j_o}^2 - uY_{j_o}\right)$with 0 < t < 1, are equal and minimized.

$$
\begin{aligned}
& min\,\delta \\
& s.t. \\
& tvX_{j_0} - \delta \leq tE_{jo}^1 \\
& (1-t)uY_{j_0} + \delta \geq (1-t)E_{jo}^2 \\
& wZ_{j_0} = 1 \\
& wZ_j - vX_j \leq 0, \;\; j = 1, \dots, n \\
& uY_j - wZ_j \leq 0, \;\; j = 1, \dots, n \\
& v \geq 0, w \geq 0, u \geq 0, \delta \geq 0
\end{aligned}
\qquad (4.17)
$$

Unlike the parametric model (4.15), the above min-max formulation (4.17) gives continuous changes on the location of the Pareto optimal point for continuous changes of the parameter *t*. Thus, the optimal solution of (4.17) responds accurately to any given set of weights that gives priority to one stage over the other. In this sense, the unweighted min-max model (4.16) aligns more effectively with the notion of "neutrality" in the efficiency assessments than model (4.13) does and provides, thus, more balanced efficiency scores for the two stages.

Aggregation of the individual efficiencies

As noticed in Cook et al (2010a), it is reasonable to define the overall efficiency of the two-stage process either as the average (arithmetic mean) of the efficiencies of the two individual stages or as their product. Liang et al (2006) and Chen et al (2006) propose the use of the arithmetic mean, in this line of thought, the overall efficiency of unit j_0 is defined as:

$$\hat{e}_{j_o}^o = \frac{1}{2}\left(\hat{e}_{j_o}^1 + \hat{e}_{j_o}^2\right)$$

As the stage efficiencies are assumption-free, i.e. their assessment does not depend on any a priori definition of the overall efficiency, alternatively, they can be aggregated multiplicatively to get the overall efficiency as follows:

$$\hat{e}^o_{j_o} = \hat{e}^1_{j_o} \cdot \hat{e}^2_{j_o} = \frac{1}{v^* X_{j_o}} \cdot u^* Y_{j_o} = \frac{u^* Y_{j_o}}{v^* X_{j_o}}$$

In the following, we compare the overall and the stage efficiencies obtained from our approach with those obtained from the additive and the multiplicative decomposition methods. Although the overall efficiency scores $\hat{e}^o$ and e^o obtained respectively by our aggregation method (simple arithmetic average) and the additive decomposition model (4.5) are not comparable, because of the endogenous weights assumed for the two stages in the latter, in the case of the multiplicative decomposition model (3.7) the following hold:

Theorem 4.3: If $\hat{e}^o_{j_o} = \hat{e}^1_{j_o} \cdot \hat{e}^2_{j_o}$ is the overall efficiency score of the evaluated unit j_0, with $\hat{e}^1_{j_o}$, $\hat{e}^2_{j_o}$ as derived by model (4.13), and $e^o_{j_o}$ is its overall efficiency score obtained from model (3.7) then $\hat{e}^o_{j_o} \leq e^o_{j_o}$.

Proof:

Let (v', w', u') be an optimal solution of model (3.7) with $e^o_{j_o} = u' Y_{j_o}$ and (v^*, w^*, u^*) an optimal solution of model (4.13) with $\hat{e}^o_{j_o} = u^* Y_{j_o} / v^* X_{j_o}$. The following hold:

(a) (v', w', u') is an optimal solution in model (3.6). This is a direct implication of the C-C transformation.

(b) (v^*, w^*, u^*) is a feasible solution in (3.6). Indeed, (v^*, w^*, u^*) is optimal in the following ratio model:

$$\begin{aligned}
& min \frac{\eta X_{j_0} - \omega Y_{j_0}}{\varphi Z_{j_0}} \\
& s.t. \\
& \varphi Z_j - \eta X_j \leq 0, \;\; j = 1, \dots, n \\
& \omega Y_j - \varphi Z_j \leq 0, \;\; j = 1, \dots, n \\
& \eta \geq 0, \varphi \geq 0, \omega \geq 0
\end{aligned}$$

which derives from (4.13) by applying the inverse C-C transformation: $\eta = v/\tau$, $\varphi = w/\tau$, $\omega = u/\tau$ with $\tau > 0$ such that $\tau \varphi Z_o = 1$. As the above model and model (3.6) have the same feasible regions, (v^*, w^*, u^*) is feasible in (3.6). From a) and b) derives that $\hat{e}^o_{j_o} = u^*Y_{j_o}/v^*X_{j_o} \leq u'Y_{j_o} = e^o_{j_o}$, which completes the proof. □

Theorem 4.4: If $\hat{e}^o_{j_o} = \hat{e}^1_{j_o} \cdot \hat{e}^2_{j_o}$ is the overall efficiency score of the evaluated unit j_0, with $\hat{e}^1_{j_o}$, $\hat{e}^2_{j_o}$ as derived by model (4.16), and $e^o_{j_o}$ is its overall efficiency score obtained from model (3.7) then $\hat{e}^o_{j_o} \leq e^o_{j_o}$.

Proof:

Let (v', w', u') be an optimal solution of model (3.7) with $e^o_{j_o} = u'Y_{j_o}$ and (v^*, w^*, u^*, δ^*) an optimal solution of model (4.16) with $\hat{e}^o_{j_o} = u^*Y_{j_o}/v^*X_{j_o}$. The following hold:

(a) The sub-vector (v^*, w^*, u^*) is a feasible solution of model (4.13). Indeed, given the optimal δ^*, the optimal sub-vector (v^*, w^*, u^*) satisfies the three last constraints of (4.16), which define the feasible region of (4.13).

(b) (v^*, w^*, u^*) is a feasible *solution* in (3.6) as well. The proof is as in Theorem 4.3(b).

Given (a) and (b), $\hat{e}^o_{j_o} \leq e^o_{j_o}$ is direct implication of Theorem 4.3. □

The case of a single intermediate measure

The following theorems complement the findings of Liang et al (2008), where it is shown that the multiplicative approach provides the independent stage efficiencies when a single intermediate measure is assumed in the two-stage process of Type I. In the Appendix, we provide an illustrative example of this two-stage process with a single intermediate measure, originally studied by Wang et al (1997).

Theorem 4.5: In a two-stage production process of Type I with a single intermediate measure, the efficiency scores derived for the two stages by model (4.13) are identical to the independent efficiency scores.

Proof:

Assuming different weights for the intermediate measures in each stage, model (4.13) can be written as follows:

$$\begin{aligned}
& min\, vX_{j_0} - uY_{j_0} \\
& s.t. \\
& wZ_{j_0} = 1 \\
& \widehat{w}Z_{j_0} = 1 \\
& wZ_j - vX_j \leq 0, \;\; j = 1, \dots, n \\
& uY_j - \widehat{w}Z_j \leq 0, \;\; j = 1, \dots, n \\
& v \geq 0, w \geq 0, \widehat{w} \geq 0, u \geq 0
\end{aligned} \qquad (4.18)$$

where w and $\widehat{w}$ are the weight variables associated to the intermediate measures for the first and the second stage respectively. It derives straightforwardly that if (u^*, v^*, w^*, $\widehat{w}^*$) is an optimal solution of model (4.18), then (v^*, w^*) is optimal in the linear equivalent of model (4.7) and (u^*, $\widehat{w}^*$) is optimal in the linear equivalent of model (4.8). Thus, model (4.18) can be used to estimate the independent efficiency scores of the two stages in one run for each evaluated unit. Obviously, models (4.13) and (4.18) are equivalent if $w = \widehat{w}$, with $w = (w_1, \dots, w_q), \widehat{w} = (\widehat{w}_1, \dots, \widehat{w}_q)$.This naturally holds in the case of a two-stage production process with a single intermediate measure. Indeed, from model (4.18) derives that the weights $(w, \widehat{w})$, associated with the single intermediate measure z, coincide and for each evaluated unit j_0 can be obtained as $w^* = \widehat{w}^* = 1/z_{j_0}$. □

Theorem 4.6: In a two-stage production process with a single intermediate measure, the efficiency scores derived for the two stages by model (4.16) are identical to the independent efficiency scores.

Proof:

Similar to the proof of Theorem 4.5, assuming different weights for the intermediate measures in each stage, the model (4.16) can be written as follows:

$$
\begin{aligned}
& min\, \delta \\
& s.t. \\
& vX_{j_0} - \delta \le E^1_{jo} \\
& uY_{j_0} + \delta \ge E^2_{jo} \\
& wZ_{j_0} = 1 \\
& \widehat{w}Z_{j_0} = 1 \\
& wZ_j - vX_j \le 0, \;\; j = 1, \dots, n \\
& uY_j - \widehat{w}Z_j \le 0, \;\; j = 1, \dots, n \\
& v \ge 0, w \ge 0, \widehat{w} \ge 0, u \ge 0, \delta \ge 0
\end{aligned}
\qquad (4.19)
$$

In model (4.19) the weights w and $\widehat{w}$ are the variables associated to the intermediate measures for the first and the second stage respectively. Models (4.16) and (4.19) are equivalent if $w = \widehat{w}$, with $w = (w_1, \dots, w_q), \widehat{w} = (\widehat{w}_1, \dots, \widehat{w}_q)$. As occurs in model (4.18), in model (4.19) as well, the weights $(w, \widehat{w})$, associated with the single intermediate measure z, coincide and for each evaluated unit j_0 are calculated as $w^* = \widehat{w}^* = 1/z_{j_0}$. To put it differently, the competition between the stages over the intermediate measures is cancelled and their single value is derived directly from the normalization constraints. The purpose of model (4.19) is to minimize the deviations of the two objectives from their ideal values. If at the given optimal solution $(v^*, u^*, w^*, \widehat{w}^*, \delta^*)$ of (4.19) the optimal value of δ is zero (δ^*=0), then the two objectives achieve their ideal values. This happens because the first two constraints of model (4.19) are always binding (cf. Ballestero and Romero, 1991, Tamiz et al, 1998 and Ogryczak, 2001). As a result, it straightforwardly derives that (v^*, w^*) are optimal in the linear equivalent of model (4.7) and $(u^*, \widehat{w}^*)$ are optimal in the linear equivalent of model (4.8). From the above, we conclude that model (4.19) can be used to estimate the independent efficiency scores of the two stages in one run for each evaluated unit. □

4.3.2 Variable returns to scale

Our approach enables us to extend our developments under the variable returns-to-scale (VRS) assumption by considering the VRS variants of models (4.7) and (4.8).

Stage I - VRS variant of (4.7)

$$\begin{aligned} & min \frac{\eta X_{j_0} - \psi_1}{\varphi Z_{j_0}} \\ & s.t. \\ & \frac{\eta X_j - \psi_1}{\varphi Z_j} \geq 1, \;\; j = 1, \ldots, n \\ & \eta \geq 0, \varphi \geq 0 \end{aligned} \qquad (4.20)$$

Stage II - VRS variant of (4.8)

$$\begin{aligned} & max \frac{\omega Y_{j_0} - \psi_2}{\varphi Z_{j_0}} \\ & s.t. \\ & \frac{\omega Y_j - \psi_2}{\varphi Z_j} \leq 1, \;\; j = 1, \ldots, n \\ & \varphi \geq 0, \omega \geq 0 \end{aligned} \qquad (4.21)$$

Models (4.20) and (4.21) yield the independent VRS efficiency scores for the two stages. Working similarly to the CRS case, we formulate the augmented models (4.22) and (4.23), for the first and the second stage respectively, by appending the constraints of model (4.20) to model (4.21) and vice versa.

Stage I

$$\begin{aligned} & min \frac{\eta X_{j_0} - \psi_1}{\varphi Z_{j_0}} \\ & s.t. \\ & \frac{\eta X_j - \psi_1}{\varphi Z_j} \geq 1, \;\; j = 1, \ldots, n \\ & \frac{\omega Y_j - \psi_2}{\varphi Z_j} \leq 1, \;\; j = 1, \ldots, n \\ & \eta \geq 0, \varphi \geq 0, \omega \geq 0 \end{aligned} \qquad (4.22)$$

Stage II

$$\begin{aligned} & max \frac{\omega Y_{j_0} - \psi_2}{\varphi Z_{j_0}} \\ & s.t. \\ & \frac{\omega Y_j - \psi_2}{\varphi Z_j} \leq 1, \;\; j = 1, \ldots, n \\ & \frac{\eta X_j - \psi_1}{\varphi Z_j} \geq 1, \;\; j = 1, \ldots, n \\ & \eta \geq 0, \varphi \geq 0, \omega \geq 0 \end{aligned} \qquad (4.23)$$

Since models (4.22) and (4.23) have common constraints can be jointly considered as a bi-objective program:

$$
\begin{aligned}
& min \frac{\eta X_{j_0} - \psi_1}{\varphi Z_{j_0}} \\
& max \frac{\omega Y_{j_0} - \psi_2}{\varphi Z_{j_0}} \\
& s.t. \\
& \frac{\eta X_j - \psi_1}{\varphi Z_j} \geq 1, \;\; j = 1, \ldots, n \\
& \frac{\omega Y_j - \psi_2}{\varphi Z_j} \leq 1, \;\; j = 1, \ldots, n \\
& \eta \geq 0, \varphi \geq 0, \omega \geq 0
\end{aligned}
\tag{4.24}
$$

Applying the C-C transformation to model (4.24) can be formulated and solved as a MOLP. The correspondence of variables is $v=\tau\eta$, $u=\tau\omega$, $w=\tau\varphi$, $\xi_1=\tau\psi_1$, $\xi_2=\tau\psi_2$ where τ is a scalar variable such that $\tau\varphi Z_o = 1$. Below we give the VRS variant of the *weighted sum* model (4.13):

$$
\begin{aligned}
& min\, vX_{j_0} - \xi_1 - uY_{j_0} + \xi_2 \\
& s.t. \\
& wZ_{j_0} = 1 \\
& wZ_j - vX_j + \xi_1 \leq 0, \;\; j = 1, \ldots, n \\
& uY_j - wZ_j - \xi_2 \leq 0, \;\; j = 1, \ldots, n \\
& v \geq 0, w \geq 0, u \geq 0
\end{aligned}
\tag{4.25}
$$

As noticed in previous chapter the additive decomposition approach of Chen et al (2009b) is extendable to VRS situations as well. Notably however, the principle that the VRS efficiency scores are not less than their CRS counterparts does not generally hold in neither the additive model or in our model (4.25) above. This irregularity can be attributed to the conflicting nature of the intermediate measures, which have different interpretations in the two stages. Adding however, the constraints $vX_{j_0} - \xi_1 \leq 1/\hat{e}^1_{CRS}$ and $uY_{j_0} - \xi_2 \geq \hat{e}^2_{CRS}$ in model (4.25), where $\hat{e}^1_{CRS}$ and $\hat{e}^2_{CRS}$ are the CRS efficiency scores obtained by model (4.13), rectifies this irregularity for the units where it is observed, without affecting the efficiency scores of the other units.

The VRS variant of the *min-max* model (4.16) is given below:

$$
\begin{aligned}
& \min \delta \\
& s.t. \\
& vX_{j_0} - \xi_1 - \delta \le E^1_{jo} \\
& uY_{j_0} - \xi_2 + \delta \ge E^2_{jo} \\
& wZ_{j_0} = 1 \\
& wZ_j - vX_j + \xi_1 \le 0, \;\; j = 1, \dots, n \\
& uY_j - wZ_j - \xi_2 \le 0, \;\; j = 1, \dots, n \\
& v \ge 0, w \ge 0, u \ge 0, \delta \ge 0
\end{aligned}
\tag{4.26}
$$

Once an optimal solution of models (4.25) or (4.26) is obtained, the VRS efficiency scores for unit j_0 in the first and the second stage are respectively:

$$\hat{e}^1_{j_0} = \frac{w^* Z_{j_0}}{v^* X_{j_0} - \xi^*_1} = \frac{1}{v^* X_{j_0} - \xi^*_1}, \hat{e}^2_{j_0} = \frac{u^* Y_{j_0} - \xi^*_2}{w^* Z_{j_0}} = u^* Y_{j_0} - \xi^*_2$$

We note that in a two-stage production process with the structure of Figure 4.1, in case of a single intermediate measure, the VRS efficiency scores derived for the two stages by models (4.25) and (4.26) are identical to the VRS independent efficiency scores. The proofs are similar to the proofs in Theorems 4.5 and 4.6.

4.4 Illustration and experimentation

We apply our approach to the 24 Taiwanese non-life insurance companies originally studied in Kao and Hwang (2008). The authors noted that the production process of the non-life insurance companies in Taiwan resembles the two-stage process that illustrated in Figure 4.1. In the first stage (*marketing of the insurance*) were utilized two inputs (Operation expenses-X1 and Insurance expenses-X2) in order to produce two intermediate measures (Direct written premiums-Z1 and Reinsurance premiums-Z2). The direct written premiums are obtained from the payments of the clients while the reinsurance premiums are received from other insurance companies. Subsequently, in the second stage (*investment*) the intermediate measures are used for the production of the two final outputs (Underwriting profit-Y1 and

Investment profit-Y2). The collected premiums are invested in a portfolio, in accordance with the Insurance Law of Taiwan, that includes bank deposits, marketable securities, real estate and mortgage loans. Table 4.1 exhibits the data set.

Table 4.1: Taiwanese non-life insurance companies (source: Kao and Hwang, 2008)

DMU		X1	X2	Z1	Z2	Y1	Y2
1	Taiwan Fire	1,178,744	673,512	7,451,757	856,735	984,143	681,687
2	Chung Kuo	1,381,822	1,352,755	10,020,274	1,812,894	1,228,502	834,754
3	Tai Ping	1,177,494	592,790	4,776,548	560,244	293,613	658,428
4	China Mariners	601,320	594,259	3,174,851	371,863	248,709	177,331
5	Fubon	6,699,063	3,531,614	37,392,862	1,753,794	7,851,229	3,925,272
6	Zurich	2,627,707	668,363	9,747,908	952,326	1,713,598	415,058
7	Taian	1,942,833	1,443,100	10,685,457	643,412	2,239,593	439,039
8	Ming Tai	3,789,001	1,873,530	17,267,266	1,134,600	3,899,530	622,868
9	Central	1,567,746	950,432	11,473,162	546,337	1,043,778	264,098
10	The First	1,303,249	1,298,470	8,210,389	504,528	1,697,941	554,806
11	Kuo Hua	1,962,448	672,414	7,222,378	643,178	1,486,014	18,259
12	Union	2,592,790	650,952	9,434,406	1,118,489	1,574,191	909,295
13	Shingkong	2,609,941	1,368,802	13,921,464	811,343	3,609,236	223,047
14	South China	1,396,002	988,888	7,396,396	465,509	1,401,200	332,283
15	Cathay Century	2,184,944	651,063	10,422,297	749,893	3,355,197	555,482
16	Allianz President	1,211,716	415,071	5,606,013	402,881	854,054	197,947
17	Newa	1,453,797	1,085,019	7,695,461	342,489	3,144,484	371,984
18	AIU	757,515	547,997	3,631,484	995,620	692,731	163,927
19	North America	159,422	182,338	1,141,950	483,291	519,121	46,857
20	Federal	145,442	53,518	316,829	131,920	355,624	26,537
21	Royal & Sunalliance	84,171	26,224	225,888	40,542	51,950	6,491
22	Aisa	15,993	10,502	52,063	14,574	82,141	4,181
23	AXA	54,693	28,408	245,910	49,864	0.1	18,980
24	Mitsui Sumitomo	163,297	235,094	476,419	644,816	142,370	16,976

CRS Case

Table 4.2 displays the independent CRS efficiency scores (columns 2-3) of the two stages, as obtained from models (4.7) and (4.8), and the efficiency scores (columns 4-7) obtained by applying our model (4.13) on the data of Table 4.1. Also, Table 4.3 presents the ideal values (columns 2-3) of vX_{j_0} and uY_{j_0} in the bi-objective LP (4.12) and the results (columns 4-8) obtained by applying our model (4.16) on the data of Table 4.1.

Table 4.2: Independent efficiencies and results from our model (4.13)

DMU	$1/E_1$	E_2	$\hat{e}^1$	$\hat{e}^2$	$\hat{e}^o=(\hat{e}^1+\hat{e}^2)/2$	$\hat{e}^o=\hat{e}^1\cdot\hat{e}^2$
1	0.9926	0.7134	0.9926	0.7045	0.8485	0.6992
2	0.9985	0.6275	0.9985	0.6257	0.8121	0.6248
3	0.6900	1	0.6900	1	0.8450	0.6900
4	0.7244	0.4323	0.7243	0.4200	0.5722	0.3042
5	0.8375	1	0.8307	0.9233	0.8770	0.7670
6	0.9637	0.4057	0.9606	0.4057	0.6831	0.3897
7	0.7521	0.5378	0.7521	0.3522	0.5521	0.2649
8	0.7256	0.5113	0.7256	0.3780	0.5518	0.2743
9	1	0.2920	1	0.2233	0.6116	0.2233
10	0.8615	0.6736	0.8615	0.5408	0.7012	0.4660
11	0.7405	0.3267	0.7292	0.2066	0.4679	0.1507
12	1	0.7596	1	0.7596	0.8798	0.7596
13	0.8107	0.5435	0.8107	0.2431	0.5269	0.1970
14	0.7246	0.5178	0.7246	0.3740	0.5493	0.2710
15	1	0.7047	1	0.6138	0.8069	0.6138
16	0.9072	0.3847	0.9072	0.3356	0.6214	0.3044
17	0.7233	1	0.7232	0.4597	0.5914	0.3325
18	0.7935	0.3737	0.7935	0.3262	0.5599	0.2588
19	1	0.4158	1	0.4112	0.7056	0.4112
20	0.9332	0.9014	0.9332	0.5857	0.7594	0.5465
21	0.7505	0.2795	0.7505	0.2623	0.5064	0.1969
22	0.5895	1	0.5895	1	0.7948	0.5895
23	0.8501	0.5599	0.8426	0.4989	0.6707	0.4203
24	1	0.3351	1	0.0870	0.5435	0.0870

Table 4.3: Ideal values and results from our model (4.16)

DMU	E_1	E_2	δ	$\hat{e}^1$	$\hat{e}^2$	$\hat{e}^o=(\hat{e}^1+\hat{e}^2)/2$	$\hat{e}^o=\hat{e}^1\cdot\hat{e}^2$
1	1.0075	0.7134	0.0079	0.9848	0.7054	0.8451	0.6947
2	1.0015	0.6275	0.0014	0.9971	0.6260	0.8116	0.6242
3	1.4492	1	0	0.6900	1	0.8450	0.6900
4	1.3805	0.4323	0.0121	0.7181	0.4202	0.5692	0.3018
5	1.1940	1	0.0543	0.8011	0.9457	0.8734	0.7577
6	1.0377	0.4057	0.0019	0.9619	0.4037	0.6828	0.3883
7	1.3296	0.5378	0.1352	0.6827	0.4026	0.5426	0.2748
8	1.3782	0.5113	0.1038	0.6748	0.4076	0.5412	0.2750
9	1	0.2920	0.0597	0.9437	0.2323	0.5880	0.2192
10	1.1607	0.6736	0.1139	0.7845	0.5597	0.6721	0.4391
11	1.3504	0.3267	0.0991	0.6899	0.2276	0.4587	0.1570
12	1	0.7596	0	1	0.7596	0.8798	0.7596
13	1.2335	0.5435	0.2383	0.6794	0.3052	0.4923	0.2073
14	1.3800	0.5178	0.0956	0.6777	0.4222	0.5500	0.2861
15	1	0.7047	0.0671	0.9371	0.6376	0.7874	0.5976
16	1.1023	0.3847	0.0250	0.8871	0.3597	0.6234	0.3191
17	1.3825	1	0.3817	0.5668	0.6183	0.5925	0.3504
18	1.2602	0.3737	0.0401	0.7691	0.3335	0.5513	0.2565
19	1	0.4158	0.0038	0.9962	0.4120	0.7041	0.4104
20	1.0716	0.9014	0.2251	0.7712	0.6763	0.7238	0.5216
21	1.3324	0.2795	0.0127	0.7434	0.2668	0.5051	0.1984
22	1.6963	1	0	0.5895	1	0.7948	0.5895
23	1.1764	0.5599	0.0520	0.8141	0.5079	0.6610	0.4135
24	1	0.3351	0.2096	0.8267	0.1255	0.4761	0.1037

For comparison purposes, we give in Table 4.4 the results obtained from the additive decomposition model (4.5) of Chen et al (2009b) along with the weights (columns 2-6) and the corresponding results obtained from the multiplicative decomposition model (3.7) of Kao and Hwang (2008) (columns 7-9).

Table 4.4: Results from models (4.5) and (3.7)

	Chen et al (2009b)					Kao and Hwang (2008)		
DMU	e^1	e^2	e^o	t^1	t^2	e^1	e^2	e^o
1	0.9926	0.7045	0.8491	0.502	0.498	0.9926	0.7045	0.6992
2	0.9985	0.6257	0.8122	0.500	0.500	0.9985	0.6257	0.6248
3	0.6900	1	0.8166	0.592	0.408	0.6900	1	0.6900
4	0.7243	0.4200	0.5965	0.580	0.420	0.7243	0.4200	0.3042
5	0.8307	0.9233	0.8727	0.546	0.454	0.8307	0.9233	0.7670
6	0.9606	0.4057	0.6887	0.510	0.490	0.9606	0.4057	0.3897
7	0.7521	0.3522	0.5804	0.571	0.429	0.6706	0.4124	0.2766
8	0.7256	0.3780	0.5795	0.580	0.420	0.6630	0.4150	0.2752
9	1	0.2233	0.6116	0.500	0.500	1	0.2233	0.2233
10	0.8615	0.5408	0.7131	0.537	0.463	0.8615	0.5408	0.4660
11	0.7291	0.2068	0.5088	0.578	0.422	0.6468	0.2534	0.1639
12	1	0.7596	0.8798	0.500	0.500	1	0.7596	0.7596
13	0.8107	0.2431	0.5565	0.552	0.448	0.6720	0.3093	0.2078
14	0.7246	0.3740	0.5773	0.580	0.420	0.6699	0.4309	0.2886
15	1	0.6138	0.8069	0.500	0.500	1	0.6138	0.6138
16	0.8856	0.3615	0.6395	0.530	0.470	0.8856	0.3615	0.3202
17	0.7232	0.4597	0.6126	0.580	0.420	0.6276	0.5736	0.3600
18	0.7935	0.3262	0.5868	0.558	0.442	0.7935	0.3262	0.2588
19	1	0.4112	0.7056	0.500	0.500	1	0.4112	0.4112
20	0.9332	0.5857	0.7654	0.517	0.483	0.9332	0.5857	0.5465
21	0.7505	0.2623	0.5412	0.571	0.429	0.7321	0.2743	0.2008
22	0.5895	1	0.7418	0.629	0.371	0.5895	1	0.5895
23	0.8426	0.4989	0.6854	0.543	0.457	0.8426	0.4989	0.4203
24	1	0.0870	0.5435	0.500	0.500	0.4287	0.3145	0.1348

Although one can spot only a few differences among the individual efficiency scores obtained by model (4.13) and those obtained by models (4.5) and (3.7), in general, our approach does not yield the same efficiency scores for the individual stages with the other two methods. For instance, the stage-1 and stage-2 efficiency scores for DMU 16 (Allianz President) differ substantially from those obtained from the additive decomposition method. As regards the results obtained from the multiplicative decomposition method, the individual efficiency scores are different for 9 of the 24 units. Our experiments with different randomly generated data sets (100 data sets drawn from a uniform distribution, with 50 DMUs, 2 external inputs, 3 intermediate measures and 2 final outputs) revealed significant differentiation in the efficiency results between the three methods.

Figure 4.4 depicts the percentage of units in each run that showed different stage efficiency scores, with respect to model (4.13) and the additive model (4.5). The range of differences varies from 0% to 82%. In only one case the efficiency scores were identical for all the units.

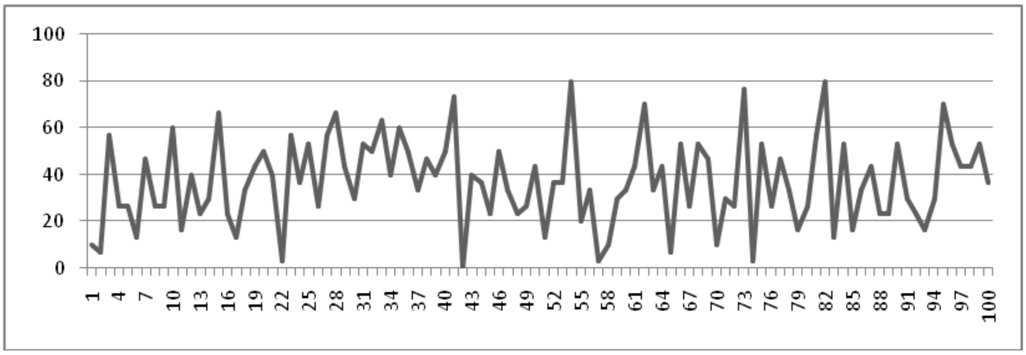


Fig. 4.4: Percentage of units showing different stage efficiencies: model (4.13) vs. model (4.5)

Analogously, Figure 4.5 depicts the percentage of units in each run that showed different individual efficiency scores, with respect to model (4.13) and the multiplicative model (3.7). The range of differences varies from 23% to 97%. None case was spotted with identical efficiency scores for all the units.

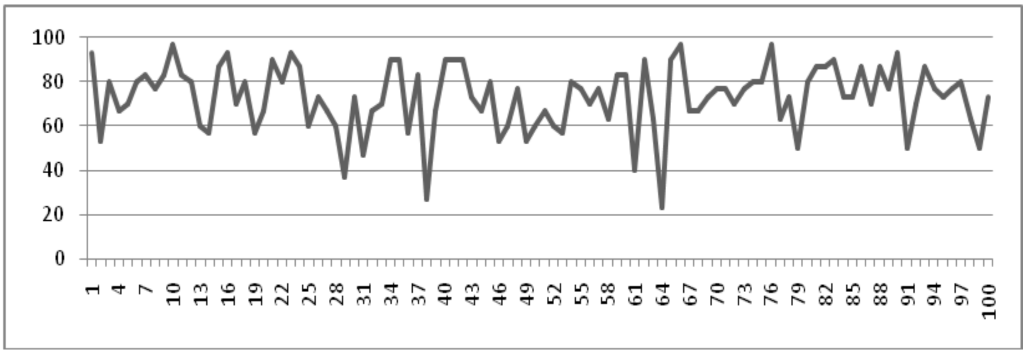


Fig. 4.5: Percentage of units showing different stage efficiencies: model (4.13) vs. model (3.7)

For the scores obtained from model (4.13), one can see that $\hat{e}^1 \geq e^1$ and $\hat{e}^2 \leq e^2$ where e^1 and e^2 are the stage-1 and stage-2 efficiency scores derived by either the additive or the multiplicative models. These relations are completely verified throughout our experiments mentioned above. As concerns the additive decomposition model (4.5), it is empirical evidence that the efficiency assessments are biased in favor of the second stage. As noted earlier, in reference to the results obtained by models (4.13) and (4.5), all the units but one (DMU 16) show identical individual scores for the two stages. A rigorous justification of both the similarities and the dissimilarities in the results can be given by solving model (4.15) for different values of the parameter t; $0<t<1$.Table 4.5 exhibits, for a limited number of DMUs, the different efficiency scores with the indifference ranges of the parameter t. Due to space limitations, we have omitted most of the DMUs that show identical results for all the models. Column two shows the indifference ranges of the parameter t, within which the efficiency scores remain the same. Columns four and five present the efficiency scores for the two stages supported by the corresponding t-range in line. These scores correspond to successive extreme points (vertices) on the Pareto front generated by model (4.15). The asterisks in the last three columns indicate, among the alternative efficiency scores, those derived by the additive decomposition model (4.5) of Chen et al (2009b), our model (4.13) and the multiplicative model (3.7) of Kao and Hwang (2008), respectively. Column three depicts the endogenous weight t^2 assumed for the second stage in model (4.5). As illustrated above, the additive decomposition model (4.5) biases the efficiency assessments in favor of the second stage, since the maximum value that t^2 can attain is 0.5 and e^2 increases (e^1 decreases) as t^2 decreases. Also, because the parametric model (4.15) is a composition rather than a decomposition model, the effect of changing the parameter t is strictly interpreted in relation to the weight t^2. The coinciding efficiency scores derived by models (4.5) and (4.13), for all the units but one (DMU 16) can now be rigorously justified by the fact that the supporting t-ranges contain both the weight values for t^2 assumed by model (4.5) as well as t=0.5, which reflects the neutral (unweighted) character of model (4.13). As concerns the DMU 16, the t-range supporting the efficiency scores obtained by model (4.5) does not include the parameter value t=0.5. This is exactly the source of differentiation in the results for DMU 16. In addition, Table 4.5 shows that the parametric version of our model (4.13) can effectively locate the individual efficiency scores obtained from both the additive and the multiplicative decomposition methods.

Table 4.5: Efficiency scores obtained by model (4.15) for different values of *t*

DMU	*t* (indifference ranges)	t^2	e^1	e^2	Model (4.5)	Model (4.13)	Model (3.7)
3	(0,1)	0.408	0.690	1	*	*	*
5	(0, 0.3355)		0.738	1			
	[0.3355, 0.9228)	0.454	0.831	0.923	*	*	*
	[0.9228, 1)		0.837	0.806			
7	(0, 0.048)		0.300	0.538			
	[0.048, 0.0528)		0.382	0.502			
	[0.0528, 0.0575)		0.514	0.464			
	[0.0575, 0.1368)		0.575	0.452			
	[0.1368, 0.2718)		0.671	0.412			*
	[0.2718, 1)	0.429	0.752	0.352	*	*	
8	(0, 0.0702)		0.390	0.511			
	[0.0702, 0.0907)		0.491	0.472			
	[0.0907, 0.1192)		0.619	0.430			
	[0.1192, 0.2215)		0.663	0.415			*
	[0.2215, 1)	0.420	0.726	0.378	*	*	
11	(0, 0.1133)		0.472	0.327			
	[0.1133, 0.2114)		0.647	0.253			*
	[0.2114, 0.651)	0.422	0.729	0.207	*	*	
	[0.651, 1)		0.741	0.168			
13	(0, 0.1148)		0.338	0.543			
	[0.1148, 0.1355)		0.405	0.480			
	[0.1355, 0.1647)		0.519	0.395			
	[0.1647, 0.2007)		0.672	0.309			*
	[0.2007, 0.211)		0.729	0.280			
	[0.211, 1)	0.448	0.811	0.243	*	*	
14	(0, 0.0298)		0.310	0.518			
	[0.0298, 0.0334)		0.392	0.497			
	[0.0334, 0.0371)		0.521	0.475			
	[0.0371, 0.1367)		0.579	0.468			
	[0.1367, 0.3356)		0.670	0.431			*
	[0.3356, 1)		0.725	0.374	*	*	
16	(0, 0.0281)		0.599	0.385			
	[0.0281, 0.0504)		0.744	0.375			
	[0.0504, 0.1406)		0.869	0.365			
	[0.1406, 0.491)	0.470	0.886	0.362	*		*
	[0.491, 1)		0.907	0.336		*	
17	(0, 0.1358)		0.251	1			
	[0.1358, 0.1461)		0.333	0.845			
	[0.1461, 0.1564)		0.466	0.698			
	[0.1564, 0.2071)		0.529	0.651			
	[0.2071, 0.3511)		0.628	0.574			*
	[0.3511, 0.9451)	0.420	0.723	0.460	*	*	
	[0.9451, 1)		0.723	0.455			
21	(0, 0.0619)		0.692	0.280			
	[0.0619, 0.2625)		0.732	0.274			*
	[0.2625, 1)	0.429	0.751	0.262	*	*	
24	(0, 0.1051)		0.399	0.335			
	[0.1051, 0.1441)		0.429	0.314			*
	[0.1441, 0.1663)		0.908	0.107			
	[0.1663, 1)	0.500	1.000	0.087	*	*	

Concerning the model (4.5) of Chen et al (2009b), the fact that the weight given to the second stage is always at least as much as the weight given to the first stage, i.e. $t_j^2 \leq t_j^1$, the additive decomposition method biases the efficiency assessments for the individual stages. Thus, the overall efficiency score is biased as well. Indeed, although each DMU is free to choose its own multipliers so as to maximize its efficiency score, the freedom in selecting the weights t^1 and t^2 is structurally limited by $t_j^2 \leq t_j^1$. The case of unit #3 in Tables 4.2, 4.3 and 4.4 is indicative. By selecting the weights t^1=0.592 and t^2=0.408 for the two stages, the stage efficiencies and the overall efficiency score obtained by the additive decomposition method are respectively e^1=0.690, e^2=1 and e^o=0.817 (=0.592×0.690+0.408×1). We get the same stage efficiencies by our model (4.13). This is due to the fact that these scores are maintained for any value of the parameter $t \in (0,1)$ in model (4.15) (see Table 4.5). However, taking the simple (unweighted) average of the same individual scores gives an overall efficiency score 0.845, which is greater than the optimal overall efficiency obtained by the additive decomposition method. The same holds for units #5 and #22.

As concerns the results obtained by the min-max model (4.16), one can see that the efficiency scores of the individual stages are more balanced than those obtained by all the other models. The fact that three units, namely units 3, 12 and 22, show identical efficiency scores for the two stages across all models is justified by the fact that, for these units, the ideal point is attainable and thus the Pareto front degenerates in this single point. Figure 4.6 depicts the Pareto front ABDE for unit 11, the Pareto optimal point B(1.3713,0.2066) derived from model (4.13) that gives the optimal stage efficiency scores (0.7292,0.2066) as well as the Pareto optimal point C(1.4495,0.2276) derived by the model (4.16) that gives the unique optimal stage efficiencies (0.6899, 0.2276).

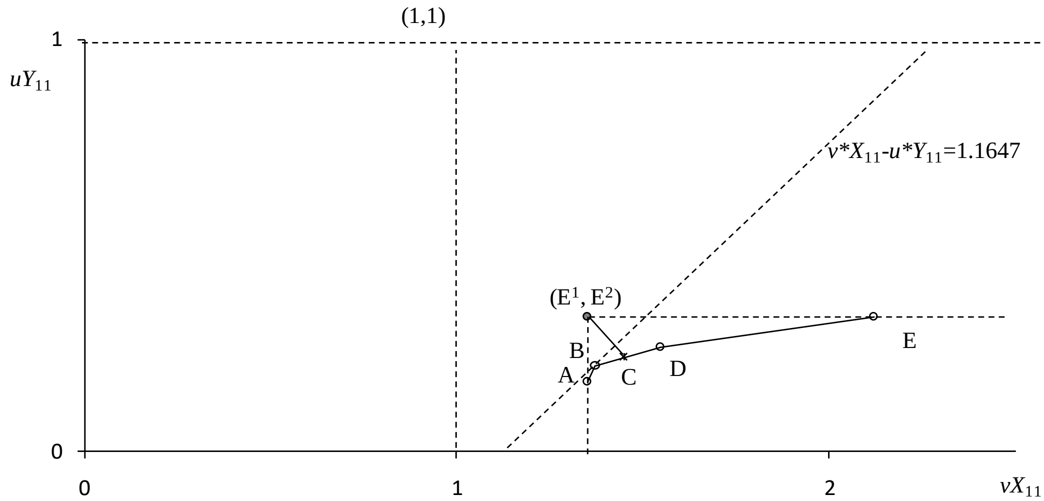


Fig. 4.6:The Pareto front of unit 11 and the Pareto points derived by models (4.13) and (4.16)

VRS Case

Table 4.6 summarizes the results obtained from our VRS model (4.25) and the corresponding results given in Chen et al (2009b) under the VRS assumption. The free of sign variables ξ_1 and ξ_2 which are related to the first and the second stage respectively, verify that variable returns to scale exist in each stage.

Table 4.6: Results from our VRS model (4.25) compared to VRS model (3.36) of Chen *et al* (2009b)

Our VRS model (4.25)						Chen et al (2009b) – VRS model(3.36)				
DMU	ξ_1	ξ_2	$\hat{e}^1$	$\hat{e}^2$	$\hat{e}^o = (\hat{e}^1+\hat{e}^2)/2$	e^1	e^2	e^0	t^1	t^2
1	-0.012	0.004	1	0.736	0.868	0.990	0.743	0.867	0.503	0.497
2	-0.001	0.029	1	0.711	0.856	1	0.711	0.856	0.500	0.500
3	-0.164	-0.010	0.700	1	0.850	0.690	1	0.818	0.587	0.413
4	-0.002	0.003	0.724	0.425	0.575	0.726	0.424	0.599	0.581	0.419
5	-0.065	0.008	1	1	1	1	1	1	0.483	0.517
6	-0.466	0.018	0.975	0.490	0.733	0.964	0.490	0.732	0.511	0.489
7	-0.331	0.044	0.803	0.592	0.698	0.752	0.593	0.684	0.571	0.429
8	-0.146	0.676	0.838	0.687	0.762	0.783	0.722	0.754	0.523	0.477
9	-0.145	0.014	1	0.285	0.643	1	0.276	0.639	0.501	0.499
10	-0.001	0.062	0.862	0.727	0.794	0.862	0.727	0.780	0.538	0.462
11	-0.207	0.033	0.750	0.432	0.591	0.741	0.443	0.614	0.576	0.424
12	0.010	0.004	0.968	0.803	0.885	0.968	0.803	0.887	0.511	0.489
13	-0.176	0.872	0.869	0.763	0.816	0.846	0.763	0.804	0.494	0.506
14	-0.001	0.069	0.725	0.555	0.640	0.725	0.555	0.654	0.581	0.419
15	0.011	0.050	1	0.880	0.940	1	0.880	0.940	0.503	0.497
16	0.009	0.055	0.910	0.417	0.663	0.911	0.417	0.676	0.526	0.474
17	0.004	0.081	0.723	1	0.862	0.724	1	0.840	0.581	0.419
18	-0.161	0.007	0.974	0.278	0.626	0.850	0.369	0.618	0.517	0.483
19	0.058	0.457	1	0.657	0.828	1	0.657	0.833	0.515	0.485
20	0.095	0.133	0.894	1	0.947	0.902	1	0.946	0.548	0.452
21	0.233	-0.209	0.895	0.362	0.628	0.913	0.362	0.679	0.575	0.425
22	1.022	-1.000	1	1	1	1	1	1	0.634	0.366
23	0.177	-0.122	0.972	0.620	0.796	0.976	0.620	0.815	0.547	0.453
24	-0.407	0.026	1	0.101	0.551	1	0.098	0.564	0.517	0.483

In the standard DEA approach, the efficiency scores obtained under the VRS assumption are not less than their counterparts under the CRS assumption (columns 2-3 in Table 4.2). Although this is true in our additive two-stage DEA models for the overall efficiency scores, the results reveal that the efficiency scores of some units, for the individual stages, do not comply with this conventional principle. This is the case for the DMUs 12 and 20, with respect to their first stage efficiency scores e^1, and for DMU 18 with respect to the second stage efficiency e^2. A similar irregularity has been spotted in Chen et al (2009b), where experimentation with the same date set indicated that the stage-1 VRS-efficiency scores of

DMUs 1, 12 and 20 are less than their CRS counterparts, hence, not complying with the standard DEA principles.

This irregularity is also observed in the results obtained from the VRS min-max model (4.26). As it can be seen from Table 4.7, in the first stage the VRS efficiency scores of DMUs 1, 2, 4, 6 and 12 are lower than their CRS counterparts.

Table 4.7: Results from min-max model (4.26) under VRS assumption

DMU	ξ_1	ξ_2	$\hat{e}^1$	$\hat{e}^2$	$\hat{e}^o=(\hat{e}^1+\hat{e}^2)/2$
1	0.007	0.011	0.973	0.749	0.861
2	0.001	0.030	0.997	0.713	0.855
3	-0.164	0	0.700	1	0.850
4	-0.011	0.004	0.716	0.426	0.571
5	-0.095	0.008	1	1	1.000
6	-0.457	0.033	0.952	0.507	0.730
7	-0.332	0.045	0.802	0.592	0.697
8	-0.144	0.687	0.826	0.701	0.763
9	-0.143	0.013	0.991	0.286	0.639
10	-0.001	0.062	0.862	0.727	0.794
11	-0.101	0.042	0.746	0.437	0.591
12	0.011	0.009	0.955	0.814	0.885
13	-0.180	0.863	0.878	0.752	0.815
14	-0.007	0.066	0.722	0.556	0.639
15	0	0.050	1	0.880	0.940
16	0.010	0.046	0.902	0.420	0.661
17	0.004	0.081	0.723	1	0.862
18	-0.107	0.029	0.884	0.354	0.619
19	0.024	0.457	1	0.657	0.828
20	0.094	0.113	0.927	0.922	0.925
21	0.241	-0.217	0.878	0.363	0.621
22	0.759	0	1	1	1.000
23	0.185	-0.120	0.962	0.622	0.792
24	-0.344	0.027	0.829	0.143	0.486

However, as noted above, the aforementioned irregularities can be rectified by adding the constraints $vX_{j_o} - \xi_1 \leq 1/\ \hat{e}^1_{CRS}$ and $uY_{j_o} - \xi_2 \geq\ \hat{e}^2_{CRS}$ in our VRS models (4.25) and (4.26), where $\hat{e}^1_{CRS}$ and $\hat{e}^2_{CRS}$ are the CRS efficiency scores obtained by model (4.13) and (4.16) accordingly. Table 4.8 presents the results for the problematic units that derived after appending the aforementioned constraints to models (4.25) and (4.26). We point out that the efficiency scores for the rest units remained unchanged.

Table 4.8: VRS results for the problematic DMUs after appending the additional constraints to models (4.25) and (4.26)

Model	DMU	ξ_1	ξ_2	$\hat{e}^1$	$\hat{e}^2$	$\hat{e}^o=(\hat{e}^1+\hat{e}^2)/2$
(4.25)	12	0.009	0.001	1.000	0.769	0.885
	18	-0.131	0.011	0.922	0.326	0.624
	20	0.094	0.109	0.933	0.908	0.921
(4.26)	1	0.007	0.007	0.985	0.745	0.865
	2	0.001	0.029	0.997	0.713	0.855
	4	-0.008	0.004	0.718	0.426	0.572
	6	-0.460	0.027	0.962	0.500	0.731
	12	0.009	0.001	1	0.769	0.885

4.5 Deriving the efficient frontier

A peculiarity of the two-stage DEA models, resulting from the conflicting nature of the intermediate measures, is that they are not capable of providing sufficient information to derive the efficient frontier, as it is with the standard DEA models. Chen et al (2010a) observed that the usual procedure of adjusting the inputs and outputs by the efficiency scores is not sufficient to yield a frontier projection neither in the additive nor in the multiplicative decomposition models. To overcome this inability, they proposed an envelopment model to locate the efficient frontier in the Kao and Hwang's (2008) multiplicative framework, by setting the intermediate measures as variables to be estimated. This approach enabled them to compute new levels for the inputs, the outputs and the intermediate measures that constitute efficient projections. These projections depend on the orientation selected. Actually, if an input orientation is assumed, new levels of inputs and intermediate measures are computed, while the original levels of outputs are maintained. Accordingly, assuming an output orientation, new levels of outputs and intermediate measures are obtained that maintain the original input levels. However, the levels of intermediate measures in these two cases differ substantially. Unfortunately, this technique cannot be applied in the additive decomposition framework. Chen et al (2013) pointed out that the envelopment and the multiplier forms are two types of network DEA models, which use different concepts of efficiency; the former is developed explicitly on the basis of the production possibility set whereas the latter under the standard DEA ratio efficiency. Unlike the standard DEA, network DEA duality may not lead to a particular pair of network multiplier and envelopment models. Hence, Chen et al (2010a, 2013) proposed that the multiplier models should be used only for estimating the efficiency scores, while modified envelopment forms should be used for determining the frontier projections of the inefficient DMUs.

In the following, we formulate the envelopment form of model (4.13) and we use it as the basis to derive the efficient frontier of the two-stage process. To this end, consider the following model:

$$\begin{aligned}
&g_{j_0} = min\, vX_{j_0} - uY_{j_0} \\
&s.t. \\
&wZ_{j_0} = 1 \\
&-\widehat{w}Z_{j_0} = -1 \\
&-wZ_j + vX_j \geq 0,\ \ j = 1, \dots, n \\
&-uY_j + \widehat{w}Z_j \geq 0,\ \ j = 1, \dots, n \\
&w - \widehat{w} = 0 \\
&v \geq 0, w \geq 0, \widehat{w} \geq 0, u \geq 0
\end{aligned} \tag{4.27}$$

Model (4.27) is strictly equivalent to model (4.13). The difference in the formulation is that, in (4.27), different weight variables are used for the intermediate measures in the first and the second stage, which then are explicitly equalized in the last constraint. The dual of (4.27) is given below:

$$\begin{aligned}
&max\, \theta_1 - \theta_2 \\
&s.t. \\
&X\lambda + s^- = X_o \\
&Y\mu - s^+ = Y_o \\
&Z\lambda \geq \theta_1 Z_o + d \\
&Z\mu \leq \theta_2 Z_o + d \\
&\lambda \geq 0, \mu \geq 0, s^- \geq 0, s^+ \geq 0
\end{aligned} \tag{4.28}$$

where θ_1 and θ_2 are free in sign scalar variables and $d = (d_1, \dots, d_q)$ is a vector of free in sign variables. We note that in case the variables d were omitted from model (4.28), i.e. if the constraint $w - \widehat{w} = 0$ in model (4.27) was eliminated, then the optimal values of θ_1 and θ_2 in model (4.28) are the independent efficiency scores of stage-1 and stage-2 respectively $\left(\theta_1 = 1/E^1_{j_o} \geq 1, \theta_2 = E^2_{j_o} \leq 1\right)$. However, in the optimal solution of (4.28), it is $\theta_1^* - \theta_2^* = g^*_{j_o} \geq 0$, where $g^*_{j_o}$ denotes the optimal objective value in (4.27), or equivalently, in model (4.13). If $\theta_1^* - \theta_2^* = g^*_{j_o} = 0$, then the evaluated unit is overall efficient. If $\theta_1^* - \theta_2^* = g^*_{j_o} > 0$, the unit it is overall inefficient. The interpretation of model (4.28) is straightforward if we take into account the double and conflicting role of the intermediate measures and the way that the primal model (4.13) was derived. With respect to the overall efficiency, whose components of vX_{j_0}, uY_{j_0} appear in the objective function of (4.13) and (4.27), the model is of the non-oriented additive form and is capable of discriminating among overall efficient and overall inefficient units. With respect to the individual stages, the model simultaneously encompasses an output orientation for stage-1, expressed by the constraint $Z\lambda \geq \theta_1 Z_O + d$, and an input orientation for the stage-2, expressed by the constraint $Z\mu \leq \theta_2 Z_O + d$. Model (4.28) provides a dichotomic characterization of overall efficiency of the evaluated unit but not the individual efficiency scores. This limitation, however, is in analogy with the relevant limitation of Chen's et al (2010a) oriented envelopment model developed for the multiplicative decomposition method. Indeed, both the Chen's et al (2010a) model and our model (4.28) they provide the overall efficiency characterization they are structurally designed for, i.e. the former provides the overall efficiency score, as it is based on an oriented formulation, whereas the latter provides the overall efficiency status (efficient or inefficient) of the units being evaluated, as it is based on an non-oriented additive formulation with respect to external inputs and the final outputs. The analogy is completed by the fact that none of the above provides the efficiency scores for the individual stages. Although $\theta_1^* = 1/\hat{e}^1$, $\theta_2^* = \hat{e}^2$ are feasible, yet optimal values of the variables θ_1 and θ_2, it is unlikely that they will be obtained by solving (4.28). In fact, in the optimal solution $(\lambda^*, \mu^*, \theta_1^*, \theta_2^*, d^*)$ of (4.28), θ_1^* and θ_2^* can take any values by adjusting accordingly the values of d^*, such that $\theta_1^* - \theta_2^* = g_{j_o}^*$, which prohibits the model (4.28) from providing the individual efficiency scores. This is because the variables θ_1, θ_2 and d are free in sign and unbounded. Thus, the optimal λ^* and μ^* as well as the optimal value of the objective function are not affected if we require $\theta_1 \geq 1$ and $\theta_2 \leq 1$, which reflect the output and the input orientation assumed for stage-1 and stage-2 respectively. Similar remarks were made by Chen et al (2013), where they demonstrated that, under network DEA, the envelopment models do not yield the stage efficiencies but the overall efficiency.

Another issue with model (4.28) is that the divergent orientations imposed by the constraints $Z\lambda \geq \theta_1 Z_o + d$ and $Z\mu \leq \theta_2 Z_o + d$ on the intermediate measures do not allow it to provide correct projections of the inefficient units on the efficient frontier. Chen et al (2010a) overcame an analogous issue in their developments by solving a modified model, where the observed values of the intermediate measures Z_O in the constraints $Z\lambda \geq Z_o$ and $Z\mu \leq Z_o$ are replaced with variables $\tilde{Z}_o$that represent the projections for the intermediate measures. The transition of our basic envelopment model (4.28), in a form capable to derive the efficient frontier, has exactly the same rationale: to make the right hand sides of the above two constraints coincide. Setting $\theta_1 = \theta_2 = 1$, i.e. at the value where $\theta_1 \geq 1$ and $\theta_2 \leq 1$ meet, the right hand sides of the last two constraints in (4.28) become $Z_o + d = \tilde{Z}_o$, where the variables $\tilde{Z}_o$ represent, as in Chen et al (2010a), the targets for the intermediate measures. Hence, the following model is solved to obtain the projections of the inefficient units on the frontier:

$$\begin{aligned}
& max\, es^- + es^+ \\
& s.t. \\
& X\lambda + s^- = X_o \\
& Y\mu - s^+ = Y_o \\
& Z\lambda \geq \tilde{Z}_o \\
& Z\mu \leq \tilde{Z}_o \\
& \lambda \geq 0, \mu \geq 0, s^- \geq 0, s^+ \geq 0
\end{aligned} \tag{4.29}$$

where the variables $\tilde{Z}_o$ are left free in sign. Actually, as $\tilde{Z}_o$ will never take negative values because of the last constraint, the natural restrictions $\tilde{Z}_o \geq 0$ are redundant and, thus, omitted. Once an optimal solution $(\lambda^*, \mu^*, \tilde{Z}_o^*, s^{-*}, s^{+*})$ of model (4.29) is obtained, the evaluated unit is overall efficient if $s^{-*} = s^{+*} = 0$. The efficient projections of the inefficient units are calculated by the following relationships:

$$\hat{X}_o = X_o - s^{-*} \qquad \hat{Y}_o = Y_o + s^{+*} \qquad \hat{Z}_o = \tilde{Z}_o^*$$

Thus, an inefficient DMU (X_o, Z_o, Y_o) is projected onto the efficient frontier at the point $(\hat{X}_o, \hat{Z}_o, \hat{Y}_o)$. Model (4.29) is now in a pure additive form. Indeed, the dual of (4.29) is as follows:

$$\begin{aligned}
& min\, vX_{j_0} - uY_{j_0} \\
& s.t. \\
& -wZ_j + vX_j \geq 0, \;\; j = 1, \dots, n \\
& -uY_j + \hat{w}Z_j \geq 0, \;\; j = 1, \dots, n \\
& v \geq e \\
& u \geq e \\
& w - \hat{w} = 0 \\
& w \geq 0, \hat{w} \geq 0
\end{aligned} \tag{4.30}$$

or

$$
\begin{aligned}
& \min vX_{j_0} - uY_{j_0} \\
& s.t. \\
& -wZ_j + vX_j \geq 0, \quad j = 1, \dots, n \\
& -uY_j + wZ_j \geq 0, \quad j = 1, \dots, n \\
& v \geq e \\
& u \geq e \\
& w \geq 0
\end{aligned}
\tag{4.31}
$$

Table 4.9 below exhibits the projections obtained by applying model (4.29) to the data of Table 4.1.

Table 4.9: Projections for Taiwanese non-life insurance companies by model (4.29)

DMU	X1	X2	Z1	Z2	Y1	Y2
1	1,178,744.00	673,512.00	6,574,261.72	1,405,802.31	6,321,322.37	681,687.00
2	1,381,822.00	1,282,484.97	9,984,108.87	2,701,675.41	14,883,490.16	834,754.00
3	1,177,494.00	592,790.00	6,301,206.75	1,173,093.25	5,222,222.92	658,428.00
4	601,320.00	566,085.63	4,342,422.69	1,215,574.75	6,851,140.78	348,724.99
5	6,699,063.00	3,531,614.00	36,299,271.58	7,179,671.93	30,095,254.05	3,925,272.00
6	1,160,508.24	668,363.00	5,711,178.98	1,598,730.82	9,010,659.25	458,645.09
7	1,942,833.00	1,443,100.00	11,869,185.17	3,322,542.01	18,726,288.14	953,173.33
8	3,253,093.02	1,873,530.00	16,009,361.92	4,481,502.04	25,258,340.80	1,285,656.65
9	1,567,746.00	950,432.00	8,068,966.39	2,258,746.45	12,730,595.02	647,990.87
10	1,303,249.00	1,226,885.07	9,411,391.65	2,634,531.66	14,848,570.42	755,796.41
11	1,167,542.17	672,414.00	5,745,794.88	1,608,420.85	9,065,273.56	461,424.97
12	2,563,321.18	650,952.00	9,356,387.29	1,127,326.43	6,005,636.18	909,295.00
13	2,376,711.46	1,368,802.00	11,696,448.21	3,274,187.74	18,453,757.03	939,301.42
14	1,396,002.00	988,888.00	8,245,849.34	2,308,261.30	13,009,667.34	662,195.73
15	1,467,228.98	651,063.00	6,454,920.94	1,456,357.91	8,126,418.64	555,482.00
16	720,706.14	415,071.00	3,546,792.34	992,853.88	5,595,856.36	284,830.66
17	1,453,797.00	1,085,019.00	8,910,486.74	2,494,313.31	14,058,281.15	715,570.46
18	757,515.00	547,997.00	4,545,666.61	1,272,468.84	7,171,803.41	365,046.81
19	159,422.00	150,080.66	1,151,263.40	322,273.26	1,816,374.92	92,453.99
20	92,925.67	53,518.00	457,312.68	128,015.58	721,512.80	36,725.20
21	45,533.89	26,224.00	224,084.75	62,728.06	353,543.70	17,995.47
22	15,993.00	10,502.00	88,313.63	24,721.64	139,334.46	7,092.16
23	49,326.07	28,408.00	242,747.09	67,952.21	382,987.70	19,494.18
24	163,297.00	153,728.61	1,179,246.65	330,106.62	1,860,524.74	94,701.23

The efficiency status of these projections is verified by applying models (4.13) and (4.16) to an expanded data set that contains both the original DMUs (Table 4.1) and their projections (Table 4.9). Indeed, the results showed that all the projected units are rendered efficient in both stages as well as in the overall sense, while the efficiency scores of the original units remained unchanged. This confirms that our approach accurately determines the improvement targets on the efficient frontier. We extended our calculations by adding in the expanded data set the projections derived by the other approaches. Particularly, we incorporated the projections of Chen et al (2010a) and those obtained by the non-oriented envelopment network DEA model (Chen et al, 2013), which is a modification of the slacks based measure (SBM) model of Tone and Tsutsui (2009). The results showed that all the projections, no matter the method that they derive from, are efficient. This verifies that our models (4.13) and (4.16) maintain the efficiency status of alternative projections obtained by the other methods. Notably, our projections were deemed efficient as well, when tested with the decomposition models (3.7) and (4.5).

Minimizing the distortion of intermediate measures

The fact that the intermediate measures are outputs of the first stage and, simultaneously, inputs to the second stage imposes that higher levels are desirable with respect to the first stage, whereas lower levels are desirable with respect to the second stage. In addition, the intermediate measures are conceived somehow as a “hidden layer” in the production process and they are or should be the less controlled dimensions. Thus, we argue that improvement suggestions with target setting should mainly involve the external inputs and the final outputs, with the changes on the observed status of the intermediate measures being kept at a minimum distortion of their original values. Based on this rationale we develop another two-phase method that provides targets on the efficient frontier at a minimum distortion of the observed status of the intermediate measures. Such an issue is not taken into account in other projection methods (Chen et al, 2010a and Chen et al, 2013), where the projected levels of the intermediate measures differ substantially from their original values and, moreover, depend on the orientation assumed.

To develop our two-phase procedure, we select an input orientation for the first stage and an output orientation for the second stage as follows:

Stage I: Input-oriented

$$\begin{aligned} &\min \theta_1 \\ &s.t. \\ &X\lambda \leq \theta_1 X_o \\ &Z\lambda \geq Z_o \\ &\lambda \geq 0 \end{aligned} \tag{4.32}$$

Stage II: Output-oriented

$$\begin{aligned} &\max \theta_2 \\ &s.t. \\ &Y\mu \geq \theta_2 Y_o \\ &Z\mu \leq Z_o \\ &\mu \geq 0 \end{aligned} \tag{4.33}$$

Appending the constraints of model (4.32) to model (4.33), and vice versa, we derive the following two augmented models for the first and the second stage respectively:

Stage I: Input-oriented

$$\begin{aligned}
&\min \theta_1 \\
&s.t. \\
&X\lambda \leq \theta_1 X_o \\
&Z\lambda \geq Z_o \\
&Y\mu \geq \theta_2 Y_o \\
&Z\mu \leq Z_o \\
&\lambda \geq 0\,, \mu \geq 0
\end{aligned} \tag{4.34}$$

Stage II: Output-oriented

$$\begin{aligned}
&\max \theta_2 \\
&s.t. \\
&Y\mu \geq \theta_2 Y_o \\
&Z\mu \leq Z_o \\
&X\lambda \leq \theta_1 X_o \\
&Z\lambda \geq Z_o \\
&\lambda \geq 0\,, \mu \geq 0
\end{aligned} \tag{4.35}$$

Since models (4.34) and (4.35) have common constraints, they enable us to jointly consider them as a bi-objective linear program. By aggregating the two objective functions, we derive a single-objective linear program, which is solved in the first phase. Actually, in phase-I we obtain the independent efficiency scores θ_1^*, θ_2^* of stage-1 and stage-2 respectively, which then are passed in phase-II. The optimal solution obtained by solving the phase-II linear program provides all necessary information to derive the efficient frontier.

Phase-I

$$\begin{aligned}
&\min \theta_1 - \theta_2 \\
&s.t. \\
&X\lambda \leq \theta_1 X_o \\
&Y\mu \geq \theta_2 Y_o \\
&Z\lambda \geq Z_o \\
&Z\mu \leq Z_o \\
&\lambda \geq 0\,, \mu \geq 0 \\
&\theta_1 \leq 1, \theta_2 \geq 1
\end{aligned} \tag{4.36}$$

Phase-II

$$\begin{aligned}
&\max M(es^- + es^+) - (e\tilde{\alpha} + e\tilde{\beta}) \\
&s.t. \\
&X\lambda + s^- = \theta_1^* X_o \\
&Y\mu - s^+ = \theta_2^* Y_o \\
&Z\lambda + \tilde{\alpha} - \tilde{\beta} \geq Z_o \\
&Z\mu + \tilde{\alpha} - \tilde{\beta} \leq Z_o \\
&\lambda \geq 0\,, \mu \geq 0, s^- \geq 0, s^+ \geq 0, \tilde{\alpha} \geq 0, \tilde{\beta} \geq 0
\end{aligned} \tag{4.37}$$

Notably, model (4.37) encompasses both orientations (i.e. an input orientation for the first stage and an output orientation for the second stage). Introducing the slacks with respect to the external inputs and the final outputs in phase-II, we grant priority in the first term of the objective function for defining the max-slack solution through a large positive number *M*. The vectors of variables $\tilde{\alpha}$ and $\tilde{\beta}$ express the deviations of the projected intermediate measures from their original values and are minimized in the second term of the objective function. Once an optimal solution (λ^*, μ^*, $\tilde{\alpha}^*$, $\tilde{\beta}^*$, s^{-*}, s^{+*}) of model (4.37) is obtained, the efficient projections are as follows:

$$\hat{X}_{j_o} = \sum_{j \in J} X_j \lambda_j^* \qquad \hat{Y}_{j_o} = \sum_{j \in J} Y_j \mu_j^* \qquad \hat{Z}_{j_o} = Z_{j_o} - \tilde{\alpha}^* + \tilde{\beta}^*$$

The non-zero λ's ($\lambda_j^* > 0$) and μ's ($\mu_j^* > 0$) define the reference sets of the evaluated unit j_0 with respect to the first and the second stage respectively. Hence, an inefficient DMU j_0 is projected onto the efficient frontier at the point $\left(\hat{X}_{j_o}, \hat{Z}_{j_o}, \hat{Y}_{j_o}\right)$. Table 4.10 shows the projections obtained by applying our two-phase approach. We verified the efficiency status of these projections by applying models (4.13) and (4.16) to an expanded data set that contains both the original DMUs (Table 4.1) and their projections (Table 4.10). The results showed that the projected DMUs are deemed efficient in each stage, as well as in overall sense. The efficiency scores of the original DMUs remained unaffected. Hence, we conclude that our method sufficiently determines the improvement targets on the efficient frontier.

Table 4.10: Projections for Taiwanese non-life insurance companies' data set

DMU	X1	X2	Z1	Z2	Y1	Y2
1	1,169,991.48	668,510.98	7,249,893.91	974,113.79	1,599,526.87	955,580.96
2	1,379,744.88	1,007,235.23	10,048,015.08	1,333,283.51	2,060,313.97	1,330,331.72
3	812,494.68	409,037.09	4,776,548	560,244	293,613	658,428
4	435,565.74	323,350.63	3,170,461.17	447,758.03	902,420.22	410,185.60
5	5,610,572.87	2,957,783.46	31,960,484.12	5,106,415.77	14,622,638.43	3,925,272
6	2,394,561.62	644,095.31	9,747,908	952,326	4,224,257.62	1,023,175.75
7	1,461,171.13	1,085,330.57	9,184,851.39	2,348,664.39	12,417,281.80	816,303.89
8	2,749,308.73	1,359,438.12	13,828,017.05	2,854,279.04	15,849,864.27	1,218,092.77
9	1,567,746	950,432	8,980,647.15	1,995,677.64	9,337,201.67	904,563.02
10	1,122,800.66	962,118.19	8,135,684.23	1,796,085.06	8,348,448.83	823,638.86
11	864,619.42	497,953.92	4,255,029.05	1,191,110.64	6,713,257.80	341,706.71
12	2,592,790	650,952	9,434,406	1,118,489	2,072,447.87	1,197,101.55
13	1,926,762.67	1,109,666.29	9,482,126.92	2,654,332.59	14,960,171.08	761,476.92
14	1,011,570.85	716,567.94	6,403,267.63	1,432,154.48	6,743,464.76	641,697.63
15	2,184,944	651,063	8,362,846.66	1,239,274.13	6,727,614.13	788,216.23
16	1,099,271.91	376,553.49	5,228,225.16	569,607.31	3,001,776.66	514,488
17	1,051,543.44	784,803.25	6,445,029.02	1,804,157.52	10,168,471.44	517,578.06
18	601,119.25	434,858.11	3,994,189.62	784,714.94	3,193,705.33	438,704.08
19	159,422	138,944.61	1,154,478.63	266,685.84	1,294,827.15	112,696.62
20	86,719.86	49,943.93	426,772.20	119,466.38	673,328.37	34,272.60
21	63,172.11	19,681.67	247,590.63	37,755.61	205,432.59	23,223.15
22	9,428.26	6,191.18	52,063	14,574	82,141	4,181
23	46,491.97	24,148.32	266,799.30	39,432.18	92,279.54	33,897.71
24	163,297	153,728.61	1,179,246.65	330,106.62	1,860,524.74	94,701.23

As above, we extended our calculations by incorporating into the expanded data set the projections derived by the other approaches. We used again the projections of Chen et al (2010a) and those obtained by the modified version of the non-oriented SBM model used by Chen et al (2013). The results of applying the models (4.13) and (4.16) to the expanded data, revealed that all the projected DMUs were deemed as efficient. Also, our projections were tested using the decomposition models (3.7) and (4.5) and they were deemed as efficient. Consequently, all the projections, no matter the method that they derive from, are efficient. However, unlike the other methods, our two-phase approach minimizes the deviation of the projected intermediate measures from their original counterparts. We examined this diversifying issue further by conducting experiments with 100 randomly generated data sets drawn by the normal distribution (mean 0.5 and standard deviation 0.1). In each data set, 100 DMUs were considered with 2 external inputs, 2 intermediate measures and 2 external outputs. We calculate the mean square distance (MSD) between the original intermediate measures and the projected ones derived by our two-phase approach, the Chen's et al (2010a) input and output oriented models and the modified SBM model presented in Chen et al (2013). The descriptive statistics that facilitate the comparisons are exhibited in Table 4.11 and Figure 4.7.

Table 4.11: MSD between the original and the projected intermediate measures

	2 Phase		Chen et al (2010a) - Input Oriented		Chen et al (2010a) - Output Oriented		Chen et al (2013) - modified SBM	
	Z1	Z2	Z1	Z2	Z1	Z2	Z1	Z2
Min	0.0008	0.0018	0.0218	0.0201	0.0210	0.0422	0.0177	0.0400
Max	0.0251	0.0396	0.0672	0.0668	0.3615	0.4873	0.3586	0.4873
Average	0.0092	0.0105	0.0417	0.0415	0.1263	0.1283	0.1230	0.1272
St. Dev.	0.0059	0.0076	0.0095	0.0124	0.0751	0.0950	0.0759	0.0873

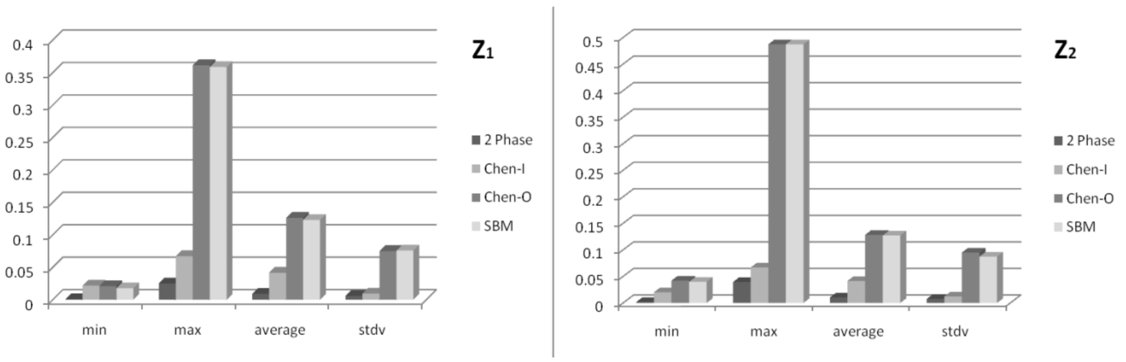


Fig. 4.7: MSD between the original and the projected intermediate measures (Z_1, Z_2)

The results obtained from another experiment with 100 data sets generated by the same distribution and parameters, but with 3 intermediate measures, are shown in Table 4.12 and Figure 4.8.

Table 4.12: MSD between the original and the projected intermediate measures

	2 Phase			Chen et al (2010a)-Input Oriented			Chen et al (2010a)-Output Oriented			SBM		
	Z1	Z2	Z3	Z1	Z2	Z3	Z1	Z2	Z3	Z1	Z2	Z3
Min	,	0.0014	0.0044	0.0197	0.0203	0.0210	0.0289	0.0403	0.0343	0.0288	0.0209	0.0404
Max	0.0531	0.0325	0.0469	0.0647	0.0646	0.0563	0.5840	0.4903	0.6087	0.5870	0.4875	0.6128
Average	0.0141	0.0120	0.0127	0.0344	0.0361	0.0355	0.1087	0.0964	0.1016	0.1064	0.0953	0.0997
St. Dev.	0.0082	0.0070	0.0083	0.0096	0.0094	0.0090	0.0876	0.0728	0.0853	0.0889	0.0750	0.0853

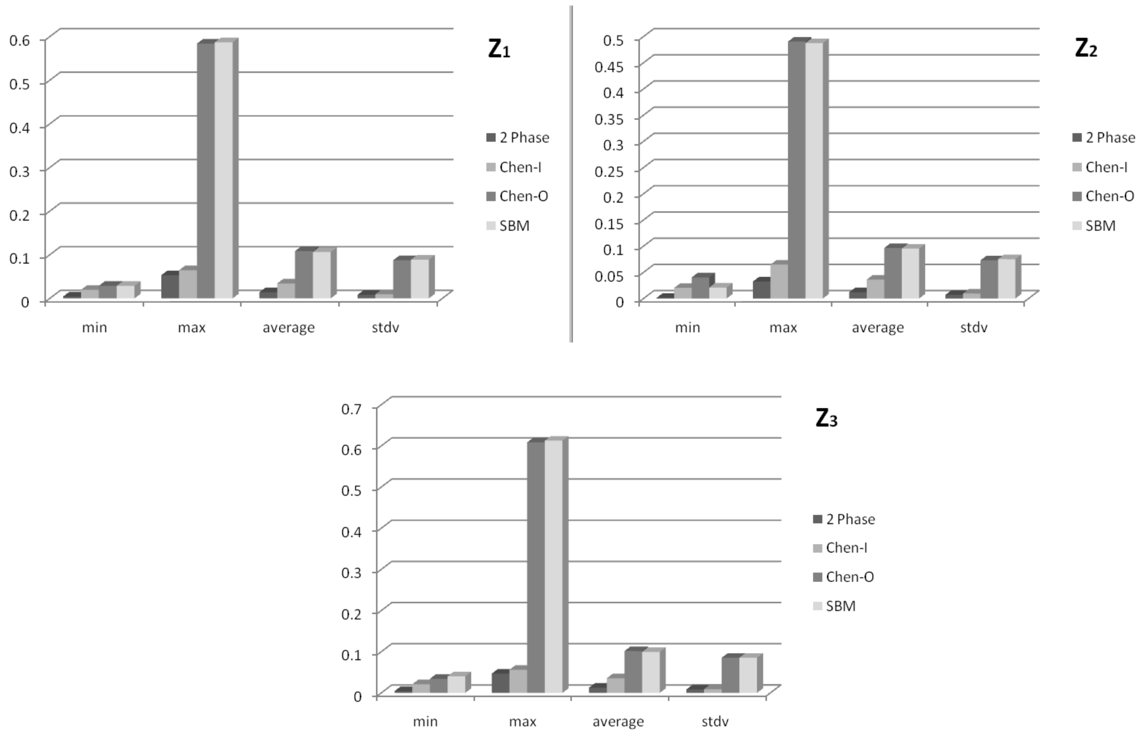


Fig. 4.8: MSD between the original and the projected intermediate measures (Z_1, Z_2, Z_3)

It is clear from the above results that our two-phase approach outperforms the other relative approaches reported in the literature, with respect to the deviations of the projected from the original intermediate measures.

Conclusion

In this chapter, we introduced a novel *composition* approach to assess the individual and the overall efficiencies for series two-stage processes of Type I. Our approach effectively overcomes the shortcomings highlighted for the additive and the multiplicative decomposition methods, by providing unique and unbiased efficiency scores for the two stages. Based on a reverse perspective in aggregating the individual efficiency scores, i.e. the composition as opposed to the decomposition approach, we estimate first the individual efficiencies for the two stages, which then can be aggregated in either an additive or a multiplicative form to obtain the overall efficiency. Our modelling approach is based on MOLP techniques and on the selection of an output orientation for the first stage and an input orientation for the second stage, with respect to the standard DEA ratio models. In this manner, the intermediate measures are used as the basis to link the efficiency assessment models for the two stages in a single linear program. The proposed approach is straightforwardly extended to fit VRS situations. Acknowledging the inadequacies observed for the envelopment network DEA models, we presented two methods to derive the efficient frontier in two-stage DEA. The first stems from the envelopment form of our basic multiplier model while the second aims to adjust the levels for the intermediate measures at a minimum distortion of their original levels.

Chapter 5

The "weak-link" approach to network DEA for two-stage processes

In the previous chapter we introduced the composition approach in two-stage network DEA. We used bi-objective linear programming to derive the efficiency scores for the individual stages, which they were aggregated ex post to obtain the overall system efficiency. Our approach provides unique and unbiased efficiency scores, in contrast to the additive and the multiplicative decomposition methods. A limitation of this approach is that it is restricted to the simple two-stage process of Type I, portrayed in Figure 3.1a. This is an effect of the different orientations selected for the first and the second stage, which in fact was made to simplify the models and keep them within the field of linear programming (simplicity at the expense of generality).

In this chapter, we extend the composition paradigm to two-stage processes of varying complexity and we introduce a novel definition of the system efficiency in two-stage processes, inspired by the "weak link" notion in supply chains and the maximum-flow/minimum-cut problem in networks (cf. Bazaraa et al, 2011). The natural representation of the supply chain operations as a multi-stage process is indicative of the synergy of supply chain management with network DEA, as they benefit mutually from the development of methodological tools for performance measurement. Adapting the "weak link" notion to fit the multi-stage processes dealt with in network DEA, we develop a max-min optimization model to estimate the stage efficiencies and the overall efficiency simultaneously in a multi-objective programming framework. This is accomplished by a two-phase procedure that provides Pareto optimal solutions and secures the uniqueness of the efficiency scores for the two-stages. We also bring into light the advantages of our method by drawing comparisons with the multiplicative method of Kao and Hwang (2008) and by identifying effectively the source of inefficiency for the DMUs.

5.1 The "weak-link" approach

We extend our composition approach presented in Chapter 4 by revisiting initially the two-stage process of Type I, where each DMU transforms some external inputs X to final outputs Y via the intermediate measures Z with a two-stage process, as depicted in Figure 5.1.

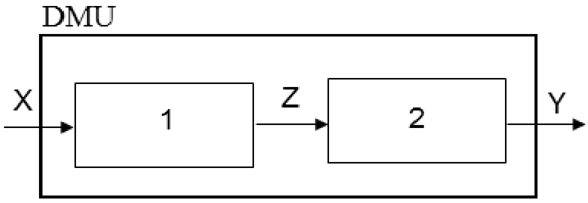


Fig. 5.1: The two-stage process of Type I

We introduce at this point some notation that will be used henceforth:

$j \in J = \{1, \dots, n\}$: The index set of the n DMUs.

$j_0 \in J$: Denotes the evaluated DMU.

$X_j = (x_{ij}, i = 1, \dots, m)$: The vector of stage-1 external inputs used by DMU$_j$.

$Z_j = (z_{pj}, p = 1, \dots, q)$: The vector of intermediate measures for DMU$_j$.

$Y_j = (y_{rj}, r = 1, \dots, s)$: The vector of stage-2 final outputs produced by DMU$_j$.

$L_j = (l_{dj}, d = 1, \dots, a)$: The vector of stage-2 extra in structures of Type I and IV.

$K_j = (k_{cj}, c = 1, \dots, b)$: The vector of stage-1 extra outputs in structures of Types III and IV.

$\eta = (\eta_1, \dots, \eta_m)$: The vector of weights for the stage-1 external inputs in the fractional model.

$v = (v_1, \dots, v_m)$: The vector of weights for the stage-1 external inputs in the linear model.

$\varphi = (\varphi_1, \dots, \varphi_q)$: The vector of weights for the intermediate measures in the fractional model.

$w = (w_1, \dots, w_q)$: The vector of weights for the intermediate measures in the linear model.

$\omega = (\omega_1, \dots, \omega_s)$: The vector of weights for the stage-2 outputs in the fractional model.

$u = (u_1, \dots, u_s)$: The vector of weights for the stage-2 outputs in the linear model.

$g = (g_1, \dots, g_a)$: The vector of weights for the stage-2 extra inputs in the fractional model.

$\gamma = (\gamma_1, \dots, \gamma_a)$: The vector of weights for the stage-2 extra inputs in the linear model.

$h = (h_1, \dots, h_b)$: The vector of weights for the stage-1 extra outputs in the fractional model.

$\mu = (\mu_1, \dots, \mu_b)$: The vector of weights for the stage-1 extra outputs in the linear model.

e_j^o: The overall efficiency of DMU$_j$.

e_j^1: The efficiency of the first stage for DMU$_j$.

e_j^2: The efficiency of the second stage for DMU$_j$.

E_j^1: The independent efficiency score of the first stage for DMU$_j$.

E_j^2: The independent efficiency score of the first stage for DMU$_j$.

Typically, the efficiencies of the first and the second stage of a DMU j are defined as follows:

$$e_j^1 = \frac{\varphi Z_j}{\eta X_j}, \; e_j^2 = \frac{\omega Y_j}{\varphi Z_j} \tag{5.1}$$

The basic input-oriented CRS-DEA models that estimate the stage-1 and the stage-2 efficiencies for the evaluated unit j_0 independently are as follows:

$$\begin{aligned} E_{j_0}^1 &= max \frac{\varphi Z_{j_0}}{\eta X_{j_0}} \\ s.t. & \\ \varphi Z_j - \eta X_j &\leq 0, \;\; j = 1, \dots, n \\ \eta \geq \varepsilon, &\varphi \geq \varepsilon \end{aligned} \tag{5.2}$$

$$\begin{aligned} E_{j_0}^2 &= max \frac{\omega Y_{j_0}}{\varphi Z_{j_0}} \\ s.t. & \\ \omega Y_j - \varphi Z_j &\leq 0, \;\; j = 1, \dots, n \\ \varphi \geq \varepsilon, &\omega \geq \varepsilon \end{aligned} \tag{5.3}$$

Appending the constraints of model (5.2) to model (5.3), and vice versa, does not affect the optimal efficiency scores attained by the models (5.2) and (5.3), see Theorems 4.1 and 4.2. Thus, the models (5.2) and (5.3) can be written as follows:

$$\begin{aligned} E_{j_0}^1 &= max \frac{\varphi Z_{j_0}}{\eta X_{j_0}} \\ s.t.& \\ \varphi Z_j - \eta X_j &\leq 0, \;\; j = 1, \dots, n \\ \omega Y_j - \varphi Z_j &\leq 0, \;\; j = 1, \dots, n \\ \eta \geq \varepsilon, \varphi &\geq \varepsilon, \omega \geq \varepsilon \end{aligned} \quad (5.4)$$

$$\begin{aligned} E_{j_0}^2 &= max \frac{\omega Y_{j_0}}{\varphi Z_{j_0}} \\ s.t.& \\ \varphi Z_j - \eta X_j &\leq 0, \;\; j = 1, \dots, n \\ \omega Y_j - \varphi Z_j &\leq 0, \;\; j = 1, \dots, n \\ \eta \geq \varepsilon, \varphi &\geq \varepsilon, \omega \geq \varepsilon \end{aligned} \quad (5.5)$$

The optimal efficiency scores $E_{j_0}^1$ and $E_{j_0}^2$ for the two stages are obtained by solving the linear equivalents of models (5.4) and (5.5), derived by applying the C-C transformation.

The following bi-objective mathematical program, is used for the performance assessment of the elementary two-stage process of Type I (Fig. 5.1):

$$\begin{aligned} & max\, e^1 = \frac{\varphi Z_{j_0}}{\eta X_{j_0}} \\ & max\, e^2 = \frac{\omega Y_{j_0}}{\varphi Z_{j_0}} \\ & s.t. \\ & \varphi Z_j - \eta X_j \leq 0, \;\; j = 1, \dots, n \\ & \omega Y_j - \varphi Z_j \leq 0, \;\; j = 1, \dots, n \\ & \eta \geq \varepsilon, \varphi \geq \varepsilon, \omega \geq \varepsilon \end{aligned} \quad (5.6)$$

or, equivalently:

$$\begin{aligned} & max\, e^1 = w Z_{j_0} \\ & max\, e^2 = \frac{u Y_{j_0}}{w Z_{j_0}} \\ & s.t. \\ & v X_{j_0} = 1 \\ & w Z_j - v X_j \leq 0, \;\; j = 1, \dots, n \\ & u Y_j - w Z_j \leq 0, \;\; j = 1, \dots, n \\ & v \geq \varepsilon, w \geq \varepsilon, u \geq \varepsilon \end{aligned} \quad (5.7)$$

Model (5.7) derives from (5.6) by applying the C-C transformation with respect to the first objective function, i.e. multiplying all the terms of the fractional objective functions and the constraints by t>0, such that $t\eta X_{j_0} = 1$ and setting $t\eta$=v, $t\omega$=u, $t\varphi$=w. Notice that in model (5.7) the second objective function is still in fractional form.

The multiplicative decomposition method estimates the overall system efficiency of the evaluated unit as the squared geometric average of the stage efficiencies. Given that $e_{j_o}^1 \leq 1$ and $e_{j_o}^2 \leq 1$ it is $e_{j_o}^o \leq min\{e_{j_o}^1, e_{j_o}^2\}$. The latter holds as equality if and only if at least one of the two stages is efficient, i.e. if $e_{j_o}^1 = 1$ and/or $e_{j_o}^2 = 1$. This property declares that the less efficient stage is determinant of the overall system efficiency. This is a natural property that can be easily identified in multi-stage processes, such as in supply chains. In such a context, the less efficient stage is called the "weak link" of the supply chain. In this line of thought, Kao (2014a) states that "*Efficiency decomposition enables decision makers to identify the stages that cause the inefficiency of the system, and to effectively improve the performance of the system*". However, in order to draw safe conclusions about the system efficiency, the identification of the weak link should meet two properties: a) uniqueness and b) being supported by a reasonable and meaningful search orientation. As mentioned in Kao and Hwang (2008), the stage efficiency scores obtained by the multiplicative method are not unique in general, i.e. different efficiency scores can be obtained for the two stages that maintain the same overall efficiency. Consequently, the weak link might be interchanged between the two stages, depending on the decomposition selected. Thus, as already remarked, the uniqueness property is not met by the model (3.7). We will give such an example below to illustrate our approach in comparison with the multiplicative method.

We define the system efficiency as the minimum of the stage efficiencies, i.e.

$$e^o = min\{e^1, e^2\} \tag{5.8}$$

To conceptualize our argument, let us resort to a *max flow-min cut* analogue. Figure 5.2 provides an alternative representation of the basic two-stage process, where the role of nodes and links is interchanged. The link that connects the nodes X and Z

represents the first stage of the process, with the nodes *X* and *Z* representing the inputs to and the outputs from the first stage. The second link represents the second stage, whose inputs and outputs are *Z* and *Y* respectively. At the DMU level, *X* and *Y* are the external inputs and outputs respectively and Z represents the intermediate measures linking the two-stages. The labels e^1 and e^2 assigned to the links represent their capacity, that is the efficiencies of the stages. Given the stage efficiencies e^1 and e^2, the system efficiency e^o can be viewed as the maximum flow through the two-stage network and can be estimated as the min-cut of the network, which, in the case of the simple network of Figure 5.2, is given by the minimal of the capacities of the two links, i.e. $e^o = min\{e^1, e^2\}$.

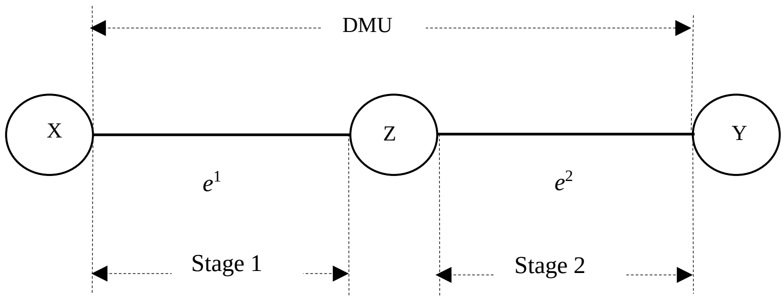


Fig. 5.2: An alternative representation of the two-stage process of Figure 5.1.

We focus on estimating the capacities (individual efficiency scores) of the two stages in a manner that the minimal capacity (the weak link) and, thus, the overall system efficiency gets the maximum possible value. The mathematical representation of this notion is expressed by the weighted max-min formulation which seeks to maximize the minimum weighted achievement from zero-level efficiency:

$$e_{j_o}^o = max_{v,w,u}[min\{q_1 e_{j_o}^1, q_2 e_{j_o}^2\}] \quad (5.9)$$

where q_1>0 and q_2>0 are strictly positive parameters (weights). A reasonable pair of values for these parameters is $q_1 = 1/E_{j_0}^1$ and $q_2 = 1/E_{j_0}^2$ (cf. Lightner and Director, 1981 and Buchanan and Gardiner, 2003). This implies that the estimated stage efficiency scores will be proportional to their independent counterparts.

5.1.1 Elementary two-stage process (Type I)

In this section we develop our approach for the elementary two-stage process of Type I. With respect to the bi-objective mathematical program (5.7), the search for the individual scores of the two stages is made in two phases: *Phase* I locates a point on the upper-right boundary of the feasible region in the objective functions space of (5.7) by means of the following max-min model (5.10), which maximizes the minimal efficiency score, whereas Phase II provides a Pareto optimal solution.

Phase I:

$$
\begin{aligned}
& max\,\theta \\
& s.t. \\
& wZ_{j_0} \geq \theta E^1_{j_0} \\
& \frac{uY_{j_0}}{wZ_{j_0}} \geq \theta E^2_{j_0} \\
& vX_{j_0} = 1 \\
& wZ_j - vX_j \leq 0,\ \ j = 1, \dots, n \\
& uY_j - wZ_j \leq 0,\ \ j = 1, \dots, n \\
& v \geq \varepsilon, w \geq \varepsilon, u \geq \varepsilon, \theta \geq 0
\end{aligned}
\tag{5.10}
$$

Model (5.10) is the canonical form of the max-min model (5.9). The solution of a weighted max-min problem, such as model (5.10), is weakly Pareto optimal (see, e.g. theorem 5.7.1 in Miettinen, 1999, pp.171). At optimality, at least one of the first two constraints in (5.10) will be binding. Although model (5.10) is non-linear, it can be solved by bisection search (known also as dichotomy method) in terms of θ in the bounded interval [0,1], since $0 \leq \theta \leq 1$ (cf. Despotis, 1996). Let $\underline{\theta}$ be a lower bound of θ for which the constraints of (5.10) are consistent (initially $\underline{\theta} = 0$) and $\overline{\theta}$ an upper bound of θ for which the constraints are not consistent (initially $\overline{\theta} = 1 + \varepsilon$), where ε is a very small positive number. Then the consistency of the constraints is tested for $\theta' = (\underline{\theta} + \overline{\theta})/2$. If they are consistent, θ' will replace $\underline{\theta}$; if they are not it will replace $\overline{\theta}$. The bisection search continues until both bounds come sufficiently close to each other. Below we provide the bisection method in algorithmic form.

Initialization

tolerance $\varepsilon > 0$, lower bound $\underline{\theta} = 0$, upper bound $\overline{\theta} = 1 + \varepsilon$, $\underline{\theta} \leq \theta \leq \overline{\theta}$.

Loop

Step 1: $= (\underline{\theta} + \overline{\theta})/2$.

Step 2: find a feasible solution for model (5.10) for the given value of θ.

Step 3: ***if*** there is a feasible solution for model (5.10) ***then*** $\underline{\theta} = \theta$ ***else*** $\overline{\theta} = \theta$.

Until $|\underline{\theta} - \overline{\theta}| < \varepsilon$

Let (θ^*, v^*, w^*, u^*) be an optimal solution of (5.10) the stage efficiencies are calculated as follows:

$$e_{j_0}^{1*} = \frac{w^* Z_{j_0}}{v^* X_{j_0}} = w^* Z_{j_0} \ , \ e_{j_0}^{2*} = \frac{u^* Y_{j_0}}{w^* Z_{j_0}}$$

The underlying idea in model (5.10) is to locate a point on the upper-right boundary of the feasible region in the objective functions space of (5.7), which is formed by the intersection of the boundary with a ray from the origin (0,0) to the ideal point $\left(E_{j_0}^1, E_{j_0}^2\right)$. However, as the *weak Pareto optimality* is a weaker property than *Pareto optimality*, it is not unlikely that the solution of (5.10) is Pareto optimal. This depends on the shape of the boundary of the objective functions space where the Pareto front is located, on the position of the ideal point relatively to that boundary and on the weights used to drive the search, which in our case are related to the components of the ideal point. This radial search approach for obtaining the individual efficiency scores will be visualized in the illustration given below. The Phase II model below provides a Pareto optimal solution to (5.7). The model (5.11) is equivalent to employing lexicographically the L_1 norm on the set of optimal solutions of (5.10) (for example Steuer and Choo, 1983).

Phase II:

$$
\begin{aligned}
& max\, s_1 + s_2 \\
& s.t. \\
& wZ_{j_0} - s_1 = e_{j_0}^{1*} \\
& uY_{j_0} - s_2 w^* Z_{j_0} - e_{j_0}^{2*} wZ_{j_0} = 0 \\
& vX_{j_0} = 1 \\
& wZ_j - vX_j \leq 0,\;\; j = 1, \dots, n \\
& uY_j - wZ_j \leq 0,\;\; j = 1, \dots, n \\
& 0 \leq s_1 \leq E_{j_0}^1, 0 \leq s_2 \leq E_{j_0}^2 \\
& v \geq \varepsilon, w \geq \varepsilon, u \geq \varepsilon
\end{aligned}
\tag{5.11}
$$

In fact, the phase II seeks for non-radial improvements of the efficiency scores on the boundary of the objective functions space. As long as the solution obtained in Phase I is weakly Pareto, in Phase II at most one of the two optimal values of the variables $\hat{s}_1$ and $\hat{s}_2$ will be strictly positive (i.e. $\hat{s}_1\hat{s}_2 = 0$). If $\hat{s}_1 = 0$ and $\hat{s}_2 = 0$, then the Phase I solution is Pareto optimal.

In (5.11), the second constraint originates from its original non-linear form $\left(uY_{j_0}/wZ_{j_0}\right) - s_2 = e_{j_0}^{2*}$ or $uY_{j_0} - s_2 wZ_{j_0} - e_{j_0}^{2*} wZ_{j_0} = 0$, where the virtual measure wZ_{j_0} in the second term is replaced by $s_1 + e_{j_0}^{1*} = s_1 + w^* Z_{j_0}$, as per the first constraint, to get $uY_{j_0} - s_1 s_2 - s_2 w^* Z_{j_0} - e_{j_0}^{2*} wZ_{j_0} = 0$. At optimality, it is $s_1 s_2 = 0$, because at least one of the two variables will be zero (see above). So the non-linear term $s_1 s_2$ can be omitted without altering the optimal solution, to get the linear form $uY_{j_0} - s_2 w^* Z_{j_0} - e_{j_0}^{2*} wZ_{j_0} = 0$. Once the optimal solution $(\hat{v}, \hat{w}, \hat{u})$ of (5.11) is obtained, the individual stage efficiency scores for unit *j0* as well as the overall efficiency of the system, according to the definition (5.8), are respectively:

$$
\hat{e}_{j_0}^1 = \frac{\hat{w}Z_{j_0}}{\hat{v}X_{j_0}} = \hat{w}Z_{j_0},\; \hat{e}_{j_0}^2 = \frac{\hat{u}Y_{j_0}}{\hat{w}Z_{j_0}},\; \hat{e}_{j_0}^o = min\left\{\hat{e}_{j_0}^1, \hat{e}_{j_0}^2\right\}
$$

Illustration

To illustrate our approach we use a synthetic case with 30 DMUs, two inputs (X1, X2), two intermediate measures (Z1, Z2) and two outputs (Y1, Y2). The data are drawn from a uniform distribution in the interval [10,100] and are presented in Table 5.1.

Table 5.1: Synthetic data

DMU	X1	X2	Z1	Z2	Y1	Y2
1	69.5	68.6	56.6	84.4	48.7	62.8
2	40.2	66.2	88	47.2	85.8	28.3
3	81.3	89.8	44.4	18.4	38.3	20.7
4	55	97.9	28.7	41.6	38.2	10.3
5	56.2	59.1	26.5	52.7	44.2	17.4
6	64.8	64.4	14.7	70.5	86.6	22.9
7	79.2	68.1	63.5	39.3	47.6	35
8	36	74.3	66.6	57.4	40.3	94.8
9	10.8	10.3	46.5	47.9	57.5	95.2
10	17.7	93.6	35.9	58.7	45.9	12
11	38.8	97.5	55.2	41.7	60.5	82.7
12	60.9	96.4	86	28.9	93.1	72.3
13	70.3	45.8	65.3	35.3	34.3	98.8
14	20.5	75.6	13.1	60	53.3	18.3
15	17.9	74.8	54.2	66.7	52.1	15.8
16	51.8	19.8	52.3	74.2	73.6	84.7
17	11.3	27.3	42.7	72.3	68.9	37.4
18	58.7	42.1	95.9	26.6	51.6	96.4
19	41.4	51.6	83	75.4	20.5	72
20	99.7	87.1	87.5	96.9	58.6	39
21	25.6	14.6	52	19.1	44.3	51.3
22	65.1	97.3	79.4	68	53.8	55.5
23	40.4	33	74.5	21.7	13.9	55.7
24	19.4	20.1	77.5	74.1	60.9	71
25	54.2	99.3	20.8	69.9	47.8	12.2
26	80.1	27.5	51.3	95.8	21.7	12.6
27	82.9	38.1	43.3	75.3	16.8	26.6
28	98.6	81.8	93.8	15.9	40.3	35.8
29	77.3	40.3	95.6	52.5	96.1	44.2
30	38.6	58.3	37.8	66.1	16	69.9

Tables 5.2 exhibits the results obtained from the multiplicative approach. The figures in columns 2-4 are the overall and the stage efficiency scores obtained by the multiplicative model (3.7). Columns 5-6 provide the maximal and the minimal efficiencies for the first stage that maintain the overall efficiency score, whereas columns 7-8 present the corresponding efficiencies for the second stage. They are all

calculated by applying the uniqueness test proposed in Kao and Hwang (2008). The results show that the efficiency decompositions for the units 8, 13, 18, 19, 23 and 30 are not unique.

Table 5.2: Results obtained from the multiplicative approach (Kao and Hwang, 2008)

DMU	e^1	e^2	e^o	e^1_{max}	e^1_-	e^2_{max}	e^2_-
1	0.2316	0.5070	0.1174	0.2316	0.2316	0.5070	0.5070
2	0.3265	0.7508	0.2451	0.3265	0.3265	0.7508	0.7508
3	0.0866	0.6844	0.0593	0.0866	0.0866	0.6844	0.6844
4	0.1344	0.5830	0.0784	0.1344	0.1344	0.5830	0.5830
5	0.1688	0.5823	0.0983	0.1688	0.1688	0.5823	0.5823
6	0.1685	1	0.1685	0.1685	0.1685	1	1
7	0.1644	0.5613	0.0923	0.1644	0.1644	0.5613	0.5613
8	0.4297	0.6953	0.2987	0.4297	0.4145	0.7208	0.6953
9	1	1	1	1	1	1	1
10	0.5235	0.5149	0.2696	0.5235	0.5235	0.5149	0.5149
11	0.3113	0.8189	0.2550	0.3113	0.3113	0.8189	0.8189
12	0.2006	1	0.2006	0.2006	0.2006	1	1
13	0.3158	0.7390	0.2334	0.3158	0.2334	1	0.7390
14	0.3759	0.7190	0.2703	0.3759	0.3759	0.7190	0.7190
15	0.6443	0.4696	0.3026	0.6443	0.6443	0.4696	0.4696
16	0.6980	0.8163	0.5698	0.6980	0.6980	0.8163	0.8163
17	1	0.6694	0.6694	1	1	0.6694	0.6694
18	0.5046	0.4910	0.2477	0.5046	0.3021	0.8201	0.4910
19	0.4656	0.4237	0.1973	0.4656	0.4537	0.4348	0.4237
20	0.2311	0.3739	0.0864	0.2311	0.2311	0.3739	0.3739
21	0.5293	0.8807	0.4662	0.5293	0.5293	0.8807	0.8807
22	0.2438	0.4927	0.1201	0.2438	0.2438	0.4927	0.4927
23	0.5001	0.3652	0.1826	0.5001	0.3031	0.6025	0.3652
24	0.8855	0.5673	0.5023	0.8855	0.8855	0.5673	0.5673
25	0.1867	0.5291	0.0988	0.1867	0.1867	0.5291	0.5291
26	0.5850	0.1674	0.0979	0.5850	0.5850	0.1674	0.1674
27	0.3403	0.2273	0.0774	0.3403	0.3403	0.2273	0.2273
28	0.1455	0.4735	0.0689	0.1455	0.1455	0.4735	0.4735
29	0.3766	0.7660	0.2885	0.3766	0.3766	0.7660	0.7660
30	0.2274	0.9032	0.2054	0.2618	0.2274	0.9032	0.7847

The unit 18, for example, is characterized by two extreme efficiency decompositions ($e^1_{max} = 0.5046, e^2_- = 0.4910$) and ($e^1_- = 0.3021, e^2_{max} = 0.8201$) that maintain the optimal overall efficiency $e^O = 0.2477$ obtained by model (3.7). Consequently, the identification of the weak link for unit 18 is not unique. Indeed, according to the first decomposition, the second stage is identified as the weak link, whereas the second decomposition identifies the first stage as the weak link of the process. Table 5.3 presents the results obtained by applying our proposed two–phase approach. Columns 2-3 present the independent efficiency scores of the two stages obtained from models (5.4) and (5.5). Columns 4-7 present the optimal value of θ, the efficiency scores of the two stages obtained by model (5.11) and the overall system efficiency, in line with the definition (5.8). Notice that the phase II did not alter the efficiency scores obtained in phase I, for any of the DMUs. That is, the phase I solutions are Pareto optimal.

Table 5.3: Results obtained from models (5.10) and (5.11)

DMU	E^1	E^2	θ	$\hat{e}^1$	$\hat{e}^2$	$\hat{e}^o = \min\{\hat{e}^1, \hat{e}^2\}$
1	0.2711	0.5511	0.8750	0.2372	0.4822	0.2372
2	0.5084	0.7508	0.7484	0.3805	0.5619	0.3805
3	0.1268	0.7391	0.7551	0.0958	0.5581	0.0958
4	0.1364	0.5830	0.9856	0.1345	0.5746	0.1345
5	0.2053	0.5842	0.8601	0.1766	0.5025	0.1766
6	0.2424	1	0.7799	0.1890	0.7799	0.1890
7	0.2065	0.5613	0.8497	0.1755	0.4769	0.1755
8	0.4297	0.7638	0.9483	0.4075	0.7243	0.4075
9	1	1	1	1	1	1
10	0.5310	0.5149	0.9865	0.5238	0.5079	0.5079
11	0.3304	0.9481	0.8910	0.2944	0.8448	0.2944
12	0.3280	1	0.7515	0.2465	0.7515	0.2465
13	0.3158	1	0.8597	0.2715	0.8597	0.2715
14	0.4574	0.8381	0.8296	0.3795	0.6953	0.3795
15	0.7329	0.4696	0.9087	0.6660	0.4267	0.4267
16	0.8058	0.8163	0.9003	0.7254	0.7349	0.7254
17	1	0.6694	1	1	0.6694	0.6694
18	0.5046	1	0.7007	0.3536	0.7007	0.3536
19	0.4656	0.4531	0.9647	0.4492	0.4371	0.4371
20	0.2392	0.3739	0.9691	0.2318	0.3624	0.2318
21	0.7889	0.9788	0.7597	0.5993	0.7436	0.5993
22	0.2833	0.4927	0.9117	0.2582	0.4492	0.2582
23	0.5001	0.7244	0.7100	0.3550	0.5144	0.3550
24	0.9278	0.5673	0.9638	0.8943	0.5468	0.5468
25	0.2297	0.5291	0.8523	0.1958	0.4509	0.1958
26	0.7491	0.1674	0.8467	0.6342	0.1417	0.1417
27	0.4250	0.3009	0.7769	0.3301	0.2337	0.2337
28	0.2540	0.8400	0.5664	0.1439	0.4757	0.1439
29	0.5255	0.7660	0.8001	0.4204	0.6129	0.4204
30	0.3304	0.9032	0.8213	0.2714	0.7418	0.2714

The curve AE in Figure 5.3 depicts the Pareto front for unit 18. The coordinates of the point E(0.5046, 1) are the independent efficiency scores of the two stages. This is the ideal point in the multi-objective programming terminology. The points A(0.2378,1) and D(0.5046, 0.4910) are the extreme points on the upper-right boundary of the feasible set in the objective functions space. The LP models that provide these two points are given below, details about the bounds on the efficient set and the range of the values of the efficient points can be found in Ehrgott (2000). The points B(0.3021, 0.8201) and D represent the two extreme decompositions mentioned above. The segment BD of the Pareto front depicts alternative decompositions, all maintaining the same optimal overall efficiency $e^O = 0.2477$, as shown in Table 5.2. Among them, D is located as an optimal solution by the multiplicative decomposition model (3.7).

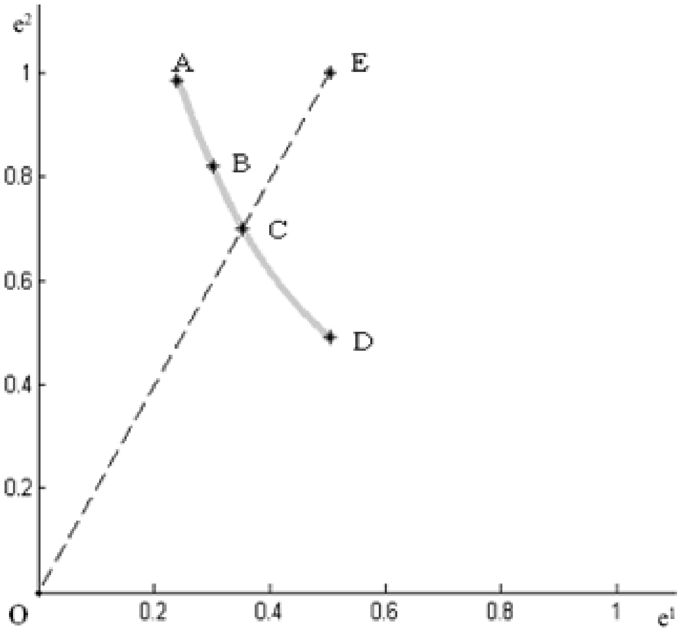


Fig. 5.3: The Pareto front for unit 18

The point C(0.3536, 0.7007), which is formed by the intersection of the dotted line OE with the Pareto front, depicts the unique Pareto optimal point obtained by model (5.11)-same as from model (5.10). Thus, the overall system efficiency of unit 18 is $\hat{e}_{18}^{o}$= min{0.3536, 0.7007}=0.3536.

Sensitivity Analysis of the weak link

The optimal value of the objective function in model (5.10) is the ratio $\theta^* = OC/OE$. That is, the search direction employed for locating the point C on the Pareto front assumes that the stage efficiency scores $\hat{e}_{j_0}^1$ and $\hat{e}_{j_0}^2$ are proportional to their independent counterparts $E_{j_0}^1$ and $E_{j_0}^2$. This is a reasonable assumption but not necessarily restrictive. For instance, one might weight differently the two stages in model (5.10), if it is to give some other priority to one stage over the other. In such a case any other point on the Pareto front could be located. In a real-world application, the management might be interested to investigate the range of weights, in which the status of the stage, characterized as weak link, is preserved. Table 5.4 provides such information concerning the sensitivity of the weak link when the weights given to the two stages vary. For each unit it is sufficient to calculate the coordinates of the two extreme points $A(N^1, E^2)$ and $D(E^1, N^2)$ on the Pareto front (columns 2-3 and 6-7 respectively) and the weights that drive the ray from the origin to these points (columns 4-5 and 8-9 respectively). Columns 10-11 show the ranges of the weight q_2, in which the status of the identified weak link for each unit is preserved.

Table 5.4: Sensitivity of the weak link

	Coordinates of $A(N^1,E^2)$		Weights pointing $A(N^1,E^2)$		Coordinates of $D(E^1,N^2)$		Weights pointing $D(E^1,N^2)$		Ranges of normalized weight q_2	
									Weak link	
DMU	N^1	E^2	q_1^A	q_2^A	E^1	N^2	q_1^D	q_2^D	Stage-1	Stage-2
1	0.1891	0.5511	0.7445	0.2555	0.2711	0.2347	0.4641	0.5359	[0.2555-0.5)	(0.5-0.5359]
2	0.3265	0.7508	0.6969	0.3031	0.5084	0.1948	0.2770	0.7230	[0.3031-0.5)	(0.5-0.723]
3	0.0791	0.7391	0.9034	0.0966	0.1268	0.2558	0.6685	0.3315	[0.0966-0.3315]	
4	0.1344	0.5830	0.8127	0.1873	0.1364	0.2850	0.6763	0.3237	[0.1873-0.3237]	
5	0.1598	0.5842	0.7852	0.2148	0.2053	0.2604	0.5591	0.4409	[0.2148-0.4409]	
6	0.1685	1	0.8558	0.1442	0.2424	0.3813	0.6114	0.3886	[0.1442-0.3886]	
7	0.1644	0.5613	0.7735	0.2265	0.2065	0.2893	0.5835	0.4165	[0.2265-0.4165]	
8	0.3423	0.7638	0.6905	0.3095	0.4297	0.6953	0.6180	0.3820	[0.3095-0.382]	
9	1	1	0.5000	0.5000	1	1	0.5000	0.5000		
10	0.5235	0.5149	0.4958	0.5042	0.5310	0.3450	0.3938	0.6062		[0.5042-0.6062]
11	0.2097	0.9481	0.8189	0.1811	0.3304	0.7410	0.6916	0.3084	[0.1811-0.3084]	
12	0.2006	1	0.8329	0.1671	0.3280	0.4384	0.5720	0.4280	[0.1671-0.428]	
13	0.2334	1	0.8108	0.1892	0.3158	0.7390	0.7006	0.2994	[0.1892-0.2994]	
14	0.1484	0.8381	0.8496	0.1504	0.4574	0.2758	0.3761	0.6239	[0.1504-0.5)	(0.5-0.6239]
15	0.6443	0.4696	0.4216	0.5784	0.7329	0.2935	0.2860	0.7140		[0.5784-0.714]
16	0.6980	0.8163	0.5391	0.4609	0.8058	0.3776	0.3191	0.6809	[0.4609-0.5)	(0.5-0.6808]
17	1	0.6694	0.4010	0.5990	1	0.6694	0.4010	0.5990		0.5990
18	0.2378	1	0.8079	0.1921	0.5046	0.4910	0.4932	0.5068	[0.1921-0.5)	(0.5-0.5068]
19	0.4197	0.4531	0.5192	0.4808	0.4656	0.4237	0.4764	0.5236	[0.4808-0.5)	(0.5-0.5236]
20	0.2311	0.3739	0.6181	0.3819	0.2392	0.1877	0.4397	0.5603	[0.3819-0.5)	(0.5-0.5603]
21	0.4217	0.9788	0.6989	0.3011	0.7889	0.4942	0.3852	0.6148	[0.3011-0.5)	(0.5-0.6148]
22	0.2438	0.4927	0.6689	0.3311	0.2833	0.3538	0.5553	0.4447	[0.3311-0.4447]	
23	0.2100	0.7244	0.7753	0.2247	0.5001	0.3652	0.4221	0.5779	[0.2247-0.5)	(0.5-0.5779]
24	0.8855	0.5673	0.3905	0.6095	0.9278	0.4587	0.3308	0.6692		[0.6095-0.6692]
25	0.1867	0.5291	0.7391	0.2609	0.2297	0.2123	0.4803	0.5197	[0.2609-0.5)	(0.5-0.5197]
26	0.5850	0.1674	0.2225	0.7775	0.7491	0.0703	0.0858	0.9142		[0.7775-0.9142]
27	0.2517	0.3009	0.5445	0.4555	0.4250	0.1031	0.1952	0.8048	[0.4555-0.5)	(0.5-0.8048]
28	0.0418	0.8400	0.9526	0.0474	0.2540	0.1960	0.4356	0.5644	[0.047-0.5)	(0.5-0.5644]
29	0.3766	0.7660	0.6704	0.3296	0.5255	0.2609	0.3318	0.6682	[0.3296-0.5)	(0.5-0.6682]
30	0.2274	0.9032	0.7988	0.2012	0.3304	0.2918	0.4690	0.5310	[0.2012-0.5)	(0.5-0.531]

Given the ideal point $(E^1_{j_0}, E^2_{j_0})$ defined by the independent stage efficiency scores of the evaluated unit j_0, the extreme points $A(N^1_{j_0}, E^2_{j_0})$ and $D(E^1_{j_0}, N^2_{j_0})$ on the upper-right boundary of the feasible set in the objective functions space of (5.7), as depicted in Figure 5.3, are obtained as follows:

For the point $A\left(N_{j_0}^1, E_{j_0}^2\right)$, get $N_{j_0}^1$ as the optimal value of the objective function in the following linear program:

$$
\begin{aligned}
& N_{j_0}^1 = max\ wZ_{j_0} \\
& s.t. \\
& vX_{j_0} = 1 \\
& \frac{uY_{j_0}}{wZ_{j_0}} \geq E_{j_0}^2 \\
& wZ_j - vX_j \leq 0,\ \ j = 1, \dots, n \\
& uY_j - wZ_j \leq 0,\ \ j = 1, \dots, n \\
& v \geq \varepsilon, w \geq \varepsilon, u \geq \varepsilon
\end{aligned}
\tag{5.12}
$$

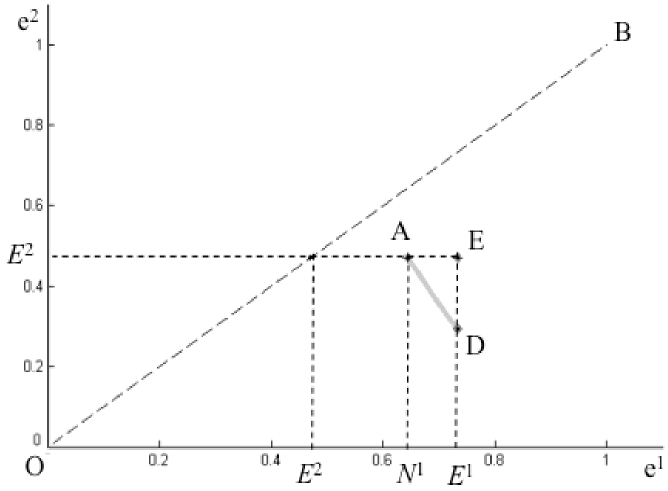
For the point $D\left(E_{j_0}^1, N_{j_0}^2\right)$, get $N_{j_0}^2$ as the optimal value of the objective function in the following linear program:

$$
\begin{aligned}
& N_{j_0}^2 = max\ uY_{j_0} \\
& s.t. \\
& wZ_{j_0} = 1 \\
& \frac{wZ_{j_0}}{vX_{j_0}} \geq E_{j_0}^1 \\
& wZ_j - vX_j \leq 0,\ \ j = 1, \dots, n \\
& uY_j - wZ_j \leq 0,\ \ j = 1, \dots, n \\
& v \geq \varepsilon, w \geq \varepsilon, u \geq \varepsilon
\end{aligned}
\tag{5.13}
$$

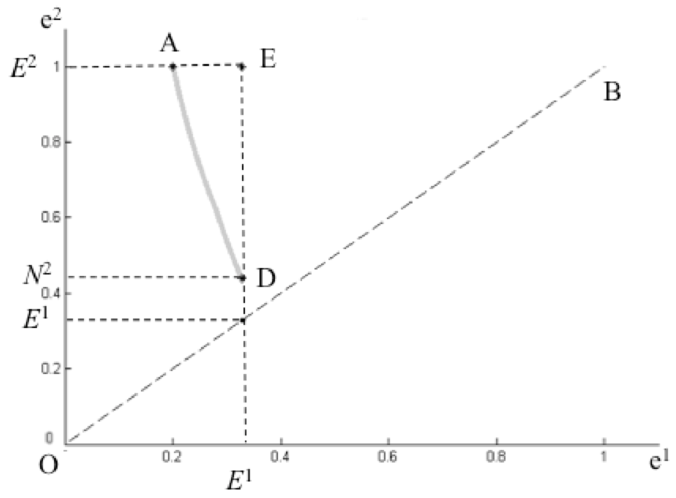
Then, the normalized weights are as follows:

For the point $A(N^1, E^2)$: $q_1^A = \frac{E^2}{N^1+E^2}$, $q_2^A = \frac{N^1}{N^1+E^2}$

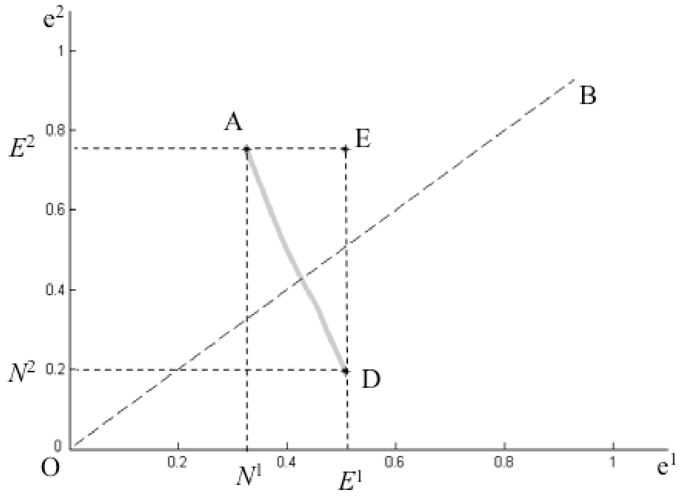
For the point $D(E^1, N^2)$: $q_1^D = \frac{N^2}{N^2+E^1}$, $q_2^D = \frac{E^1}{N^2+E^1}$



(a) Unit 15



(b) Unit 12



(c) Unit 2

Fig. 5.4: The location of the Pareto front

Figure 5.4 exhibits the three possible positions of the Pareto front AD with respect to the bisection line OB. The units depicted are No 15, 12 and 2 (Table 5.4). If $N^1 > E^2$, then the line AD lies on the right of OB and the second stage will be steadily the weak link (Fig. 5.4a). This is the case for units 10, 15, 24 and 26. If $N^2 > E^1$, then the line AD lies on the left of OB and the first stage will be steadily the weak link (Fig. 5.4b). This is the case for the units 3-8, 11-13 and 22. If none of the above holds, the line AD intersects with the bisection line OB and the weak link is differentiated on the left and on the right of the intersection point (Fig. 5.4c). This is the case for the units 1, 2, 14, 16, 18-21, 23, 25, 27-30. The units 9 and 17 are differentiated from the others, as they achieve their ideal efficiency scores and the Pareto front degenerates to a single point.

5.1.2 Two-stage process with extra inputs in the stage-2 (Type II)

In the structure of Type II, the second stage uses some extra external inputs L beyond the intermediate measures as depicted in Figure 5.5.

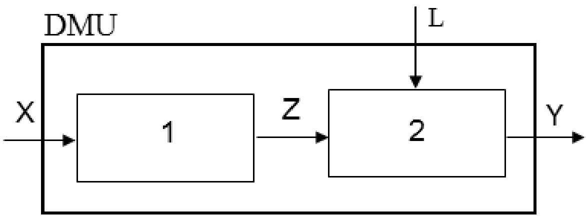


Fig. 5.5: The two-stage process of Type II

In view of the "weak-link" approach, the Fig. 5.5 can be alternatively represented by the Fig. 5.6 below.

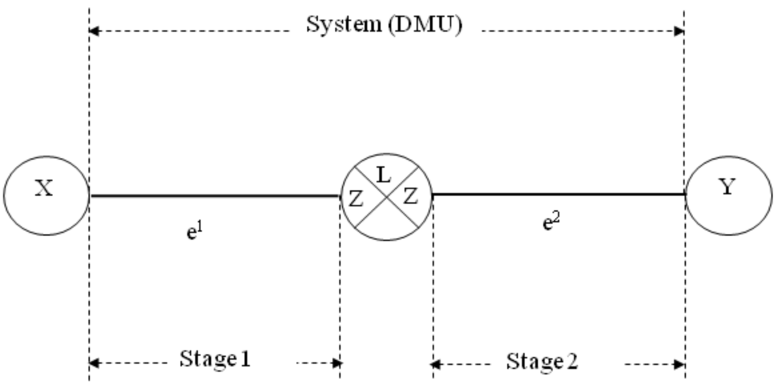


Fig. 5.6: An alternative representation of the two-stage process of Figure 5.5

In this case, the efficiency of the first and the second stage of DMU j are defined as follows:

$$e_j^1 = \frac{\varphi Z_j}{\eta X_j}, e_j^2 = \frac{\omega Y_j}{\varphi Z_j + gL_j}$$

Analogously to the simple process elaborated in the previous section, the two-phase procedure for estimating the stage efficiencies as well as the overall system efficiency, in line with model (5.9), is as follows:

$$\begin{aligned}
& max\, e^1 = \frac{\varphi Z_{j_0}}{\eta X_{j_0}} \\
& max\, e^2 = \frac{\omega Y_{j_0}}{\varphi Z_{j_0} + gL_{j_0}} \\
& s.t. \\
& \varphi Z_j - \eta X_j \leq 0,\ \ j = 1, \dots, n \\
& \omega Y_j - \varphi Z_j - gL_j \leq 0,\ \ j = 1, \dots, n \\
& \eta \geq \varepsilon, \varphi \geq \varepsilon, \omega \geq \varepsilon, g \geq \varepsilon
\end{aligned} \tag{5.14}$$

After applying the C-C transformation with respect to the first objective function we derive the following model:

$$\begin{aligned}
& max\, e^1 = wZ_{j_0} \\
& max\, e^2 = \frac{uY_{j_0}}{wZ_{j_0} + \gamma L_{j_0}} \\
& s.t. \\
& vX_{j_0} = 1 \\
& wZ_j - vX_j \leq 0,\ \ j = 1, \dots, n \\
& uY_j - wZ_j - \gamma L_j \leq 0,\ \ j = 1, \dots, n \\
& v \geq \varepsilon, w \geq \varepsilon, u \geq \varepsilon, \gamma \geq \varepsilon
\end{aligned} \tag{5.15}$$

Phase I:

$$
\begin{aligned}
& max\ \theta \\
& s.t. \\
& wZ_{j_0} \geq \theta E_{j_0}^1 \\
& \frac{uY_{j_0}}{wZ_{j_0} + \gamma L_{j_0}} \geq \theta E_{j_0}^2 \\
& vX_{j_0} = 1 \\
& wZ_j - vX_j \leq 0,\ \ j = 1, \dots, n \\
& uY_j - wZ_j - \gamma L_j \leq 0,\ \ j = 1, \dots, n \\
& v \geq \varepsilon, w \geq \varepsilon, u \geq \varepsilon, \gamma \geq \varepsilon, \theta \geq 0
\end{aligned}
\tag{5.16}
$$

Let (θ^*, v^*, w^*, u^*, γ^*) be an optimal solution of model (5.16) and the stage efficiencies as follows:

$$e_{j_o}^{1*} = \frac{w^* Z_{j_o}}{v^* X_{j_o}} = w^* Z_{j_0}\ ,\ \ e_{j_o}^{2*} = \frac{u^* Y_{j_o}}{w^* Z_{j_0} + \gamma^* L_{j_o}}$$

The phase I solution is weakly Pareto optimal to the bi-objective program (5.15).

Phase II:

Solve the following linear program:

$$
\begin{aligned}
& max\ s_1 \\
& s.t. \\
& wZ_{j_0} - s_1 = e_{j_0}^{1*} \\
& uY_{j_0} - e_{j_0}^{2*}\left(wZ_{j_0} + \gamma L_{j_0}\right) \geq 0 \\
& vX_{j_0} = 1 \\
& wZ_j - vX_j \leq 0,\ \ j = 1, \dots, n \\
& uY_j - wZ_j - \gamma L_j \leq 0,\ \ j = 1, \dots, n \\
& v \geq \varepsilon, w \geq \varepsilon, u \geq \varepsilon, \gamma \geq \varepsilon \\
& 0 \leq s_1 \leq E_{j_0}^1 - e_{j_0}^{1*}
\end{aligned}
\tag{5.17}
$$

Let $(\hat{s}_1, \hat{v}, \hat{w}, \hat{u}, \hat{\gamma})$ be the optimal solution of (5.17). If $\hat{s}_1 > 0$, then the solution is Pareto optimal and the stage efficiency scores are:

$$\hat{e}_{j_o}^1 = \frac{\hat{w} Z_{j_o}}{\hat{v} X_{j_o}}\ ,\ \hat{e}_{j_0}^2 = \frac{\hat{u} Y_{j_o}}{\hat{w} Z_{j_o} + \hat{\gamma} L_{j_o}}$$

Otherwise, solve the following program to get a Pareto optimal solution $(s'_2, v', w', u', \gamma')$ and the stage efficiency scores:

$$e'^{1}_{j_o} = \frac{w'Z_{j_o}}{v'X_{j_o}}, \; e'^{2}_{j_0} = \frac{u'Y_{j_o}}{w'Z_{j_o}+\gamma'L_{j_o}}$$

$$\begin{aligned}
& max \; s_2 \\
& s.t. \\
& wZ_{j_0} \geq e^{1*}_{j_0} \\
& \frac{uY_{j_0}}{wZ_{j_0} + \gamma L_{j_0}} - s_2 = e^{2*}_{j_0} \\
& vX_{j_0} = 1 \\
& wZ_j - vX_j \leq 0, \;\; j = 1, \dots, n \\
& uY_j - wZ_j - \gamma L_j \leq 0, \;\; j = 1, \dots, n \\
& v \geq \varepsilon, w \geq \varepsilon, u \geq \varepsilon, \gamma \geq \varepsilon \\
& 0 \leq s_2 \leq E^2_{j_0} - e^{2*}_{j_0}
\end{aligned} \tag{5.18}$$

Model (5.18) is non-linear but it can be solved by bisection search in terms of s_2 in the bounded interval $[0, E^2_{j_0} - e^{2*}_{j_0}]$. Obviously, in the case that the stage efficiency scores, derived from the phase I, they are both equal to the corresponding independent efficiency scores, there is no need to run the phase II, since the Pareto front degenerates to the point $\left(E^1_{j_0}, E^2_{j_0}\right)$.

Notice that in the two-stage process of Figure 5.5, choosing the minimum of e^1, e^2 as the system overall efficiency might underestimate it. To be more specific, if the extra inputs L in the second stage are partly a substitute for the intermediate measures Z to produce the final outputs Y, the *weak link* between the two stages can be remedied by L. For instance, if $e^1 < e^2$ and this difference is mainly attributed to the economical input on L, taking e^1 as the overall efficiency may be deceptive since the economical low level of input L may be a relative advantage for the stage under consideration among its competing peer on L and probably underestimates the overall system efficiency. In a two-stage process with extra inputs to the second stage, the aforementioned effect must be taken into account in order to transfer the advantages of stage-2, caused by the extra inputs L, to stage-1. An effective way to absorb these effects in favor of the system efficiency is proposed below. Given the optimal solution

$(\hat{v}, \hat{w}, \hat{u}, \hat{\gamma})$, as derived by the two-phase procedure above, we modify the stage efficiencies as follows:

$$e_{j_o}^{I} = \frac{\hat{w}Z_{j_o} + \hat{\gamma}L_{j_o}}{\hat{v}X_{j_o} + \hat{\gamma}L_{j_o}}, \; e_{j_o}^{II} = \frac{\hat{u}Y_{j_o}}{\hat{w}Z_{j_o} + \hat{\gamma}L_{j_o}} \tag{5.19}$$

In (5.19), the adjusted stage efficiencies originate from the network structure depicted in Figure 5.7a, which is a modification of the two-stage process of Figure 5.5. The valued input $\hat{\gamma}L$ to the second stage is used as an input and, at the same time, as an output from the first stage. So, the term $\hat{\gamma}L$ is added to the numerator and to the denominator of the original efficiency ratio to obtain the modified efficiency score $e_{j_o}^{I}$. A similar but equivalent representation to the modified structure of Figure 5.7a is the network structure depicted in Figure 5.7b. This representation technique is coined by Kao (2014a) to deal with general multi-stage systems. In the latter, the motivation was to adjust the conventional stage efficiency scores, so as they meet the property that the system efficiency is the product of the stage efficiencies. This is accomplished by introducing the dummy process labelled 3, in a parallel configuration with the original stage-1.

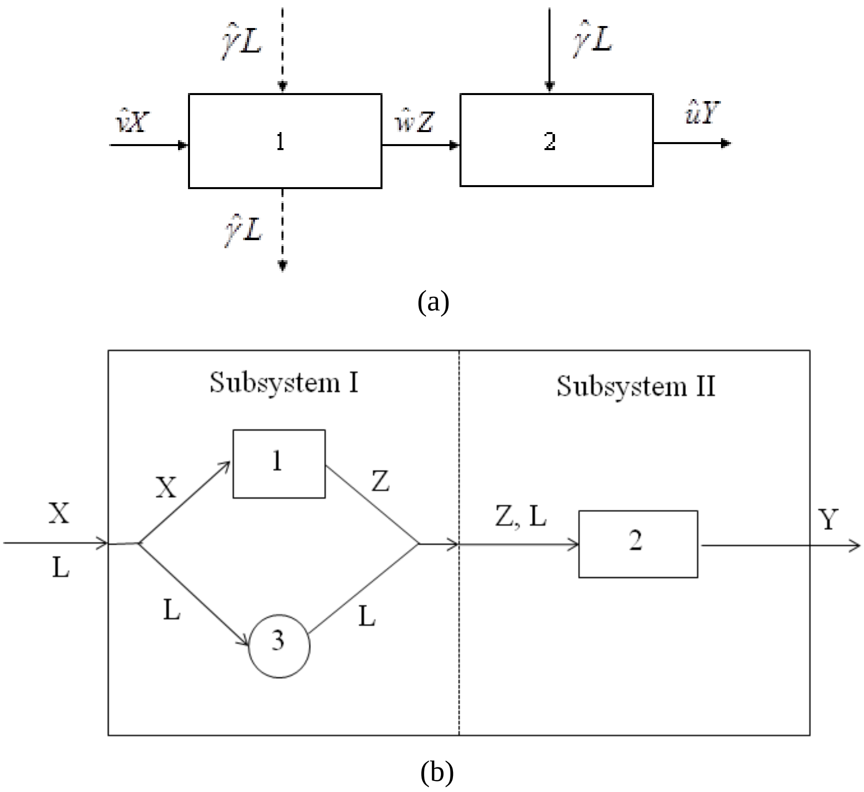


Fig. 5.7: The modified process of Figure 5.5

The adjusted efficiency scores $e_{j_0}^{I}$ and $e_{j_0}^{II}$ are related to their original counterparts $\hat{e}_{j_0}^{1}$ and $\hat{e}_{j_0}^{2}$ as follows:

$$e_{j_0}^{I} = \frac{\hat{v}X_{j_0}}{\hat{v}X_{j_0} + \hat{\gamma}L_{j_0}} \times \hat{e}_{j_0}^{1} + \frac{\hat{\gamma}L_{j_0}}{\hat{v}X_{j_0} + \hat{\gamma}L_{j_0}} \times 1$$
$$e_{j_0}^{II} = \hat{e}_{j_0}^{2} \tag{5.20}$$

That is, the adjusted efficiency $e_{j_0}^{I}$ is a weighted average of the original efficiency score $\hat{e}_{j_0}^{1}$ and the efficiency score of the dummy process 3 associated to the first stage, which is 1. Concerning the second stage, the corresponding efficiency $e_{j_0}^{II}$ is not being adjusted and remains $\hat{e}_{j_0}^{2}$. Once the adjusted efficiency scores are obtained, the overall efficiency of the evaluated unit is $e_{j_0}^{o} = min\{e_{j_0}^{I}, e_{j_0}^{II}\}$. Kao (2014a) employed a similar concept so as to decompose the overall system efficiency as the product of the subsystems efficiencies, where the efficiency of each subsystem is a weighted average of the efficiency of the corresponding process (stage) and the dummy process. However, our methodology differs substantially in the optimality criterion used. Indeed, our optimality criterion is to maximize the lowest of the stage efficiencies, whereas the optimality criterion in Kao (2014a) is to maximize the overall efficiency.

Illustration

We illustrate our two–phase procedure on a two-stage process of Type II with a real example about regional R&D process of 30 Provincial Level Regions in China, drawn from Li et al (2012). The first stage represents the technology development whereas the second one represents the economic application. The inputs of the first stage are three indices from science and technology activities and government’s support, namely the R&D personnel (X1), the R&D expenditure (X2) and the proportion of regional science and technology funds in regional total financial expenditure (X3). The intermediate measures (outputs of stage-1 and inputs to stage-2) are the number of patents (Z1) and the number of papers (Z2). The extra input to stage-2 is the contract value in technology market (L). The final outputs are four economic indicators, namely the GDP (Y1), the total exports (Y2), the urban per capita annual income (Y3) and the gross output of high-tech industry (Y4). The data are originally obtained from “China statistical yearbook, 2009” and “China science and technology statistical yearbook, 2009”. Table 5.5 below exhibits the dataset.

Table 5.5: Data of 30 Provincial Level Regions in China from Li et al (2012)

DMU	X1	X2	X3	Z1	Z2	L	Y1	Y2	Y3	Y4
1	10.34786	668.63510	5.44577	9157	65951	1236.245	12153.03	483.7932	26738.48	2757.14
2	2.00665	79.45994	1.20381	834	13737	38.31581	6530.01	42.80071	15748.67	352.84
3	6.46163	423.37740	7.20187	5997	32733	435.4108	15046.45	1417.96027	28837.78	5557.45
4	2.87830	178.46610	3.02375	1889	12472	105.4611	7521.85	298.92719	21402.01	1901.07
5	3.01654	135.95350	1.70264	795	13699	35.61736	10062.82	88.86487	14085.74	460.31
6	2.27886	135.38190	1.97549	824	9075	23.25944	12236.53	533.1911	19576.83	1972.01
7	1.27445	37.26124	0.81710	227	7856	35.62869	3387.56	7.35512	11929.78	67.39
8	12.97681	652.98200	3.88757	11355	35773	170.985	39482.56	3589.54893	21574.72	17161.94
9	0.77328	26.41343	1.04000	322	4946	1.78061	3912.68	13.56612	12862.53	293.64
10	0.17583	5.78060	1.24906	84	2726	0.55563	1654.21	13.08632	13750.85	54.75
11	3.88080	134.84460	1.12576	691	17970	17.21118	17235.48	156.88902	14718.25	629.17
12	3.70197	109.17040	1.06286	1142	14553	48.855	8587.00	100.82127	12565.98	311.40
13	4.79963	174.75990	1.22226	1129	21188	26.30461	19480.46	73.45376	14371.56	953.23
14	5.12124	213.44900	1.21147	1478	25268	77.03287	12961.1	99.78796	14367.48	1039.52
15	3.49591	153.49950	1.33984	1752	21042	44.04324	13059.69	54.92034	15084.31	648.75
16	10.67826	701.95290	2.91286	5322	47441	108.2184	34457.3	1991.9919	20551.72	13015.35
17	1.83522	75.89360	0.85780	386	6811	9.78927	7655.18	73.68488	14021.54	755.65
18	2.60875	81.36019	1.28305	719	8987	19.75983	7278.75	31.24935	14006.27	537.66
19	5.43947	232.36870	2.14308	1993	20801	119.7095	15212.49	334.14928	15761.38	1313.84
20	0.30013	7.59379	0.98211	35	1240	8.49672	1081.27	2.51876	12691.85	19.22
21	8.33303	519.59200	1.92425	2865	26941	71.9391	33896.65	794.90706	17811.04	4555.71
22	2.52624	80.85633	1.12742	603	6757	16.20675	7358.31	28.37455	13996.55	196.47
23	4.23465	189.50630	1.13144	1342	26403	69.80741	8169.80	39.88149	14128.76	717.04
24	4.87863	214.45900	0.79759	1596	22568	54.59769	14151.28	141.69447	13839.40	1766.76
25	1.22051	37.23044	0.97287	476	7101	10.24687	6169.75	45.13252	14423.93	147.17
26	5.90844	398.83670	3.74258	4818	25638	56.45805	22990.35	1330.12954	24610.81	2672.09
27	1.56993	47.20277	1.11443	326	9982	1.76618	7759.16	83.7537	15451.48	273.67
28	1.27057	52.07259	0.93756	178	3214	14.76515	9740.25	23.15476	15849.19	236.61
29	0.33954	10.44221	1.01806	52	1365	0.89823	1353.31	7.4293	14024.70	32.89
30	0.82683	21.80426	1.19830	120	5688	1.20777	4277.05	109.34563	12257.52	23.74

As already noticed in Chapter 3, Li et al (2012) employed a parametric technique for the assessment of DMUs with extra inputs in the second stage. They first derive parametrically the stage efficiency scores and then they calculate the overall efficiency as the product of the stage efficiencies. For comparison purposes we give in columns (4-6) of Table 5.6 the results of Li et al (2012). Table 5.6 contains also the independent (ideal) stage efficiency scores that derive from analogous models with models (5.4) and (5.5); these models are given in the Appendix.

Table 5.6: Ideal scores and results from Li et al (2012)

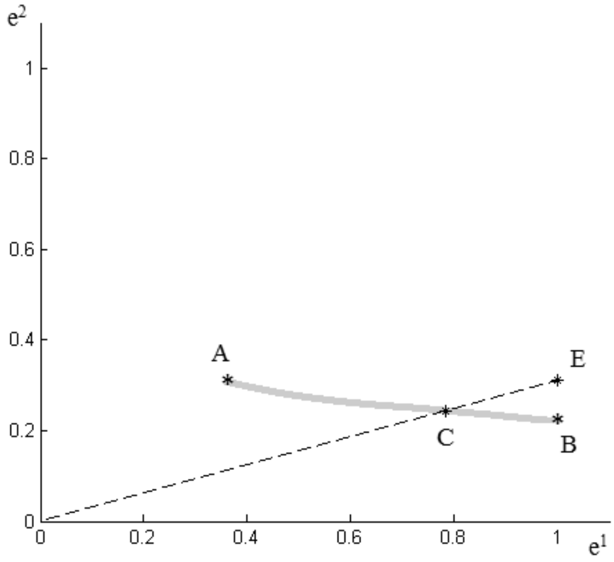
DMU	E^1	E^2	e^1	e^2	$e^o = e^1 \cdot e^2$
1	1	0.1598	1	0.1598	0.1598
2	1	0.2489	1	0.2489	0.2489
3	1	0.5728	0.8950	0.5365	0.4802
4	0.7426	0.5704	0.6774	0.5704	0.3864
5	0.6697	0.3895	0.6697	0.3895	0.2608
6	0.5668	1	0.5668	1	0.5668
7	1	0.3121	1	0.2207	0.2207
8	1	1	1	1	1
9	0.9398	1	0.9398	1	0.9398
10	1	1	1	1	1
11	0.8885	0.8351	0.8885	0.8351	0.7420
12	0.9328	0.2703	0.9328	0.2648	0.2470
13	0.8504	0.7373	0.8493	0.7373	0.6262
14	0.9060	0.3360	0.9060	0.2816	0.2551
15	1	0.3780	1	0.3685	0.3685
16	0.9225	1	0.9225	1	0.9225
17	0.5647	1	0.5644	0.9914	0.5595
18	0.7158	0.5184	0.7152	0.4947	0.3538
19	0.6969	0.3742	0.6671	0.3668	0.2447
20	0.4573	1	0.4573	1	0.4573
21	0.7101	0.8498	0.7101	0.8176	0.5806
22	0.5864	0.5709	0.5708	0.5156	0.2943
23	1	0.2509	1	0.1941	0.1941
24	1	0.4817	1	0.4566	0.4566
25	1	0.6159	1	0.5846	0.5846
26	0.9111	0.9541	0.7293	0.9171	0.6688
27	1	1	1	1	1
28	0.3599	1	0.3599	1	0.3599
29	0.4300	1	0.4300	1	0.4300
30	1	1	1	1	1

In Table 5.7 we present the results obtained from phase I and II, we remark that the stage efficiencies are not improved by phase II, in other words, models (5.16), (5.17) and (5.18) produce the same stage efficiency scores. The adjusted efficiency scores as well as the overall system efficiency are given in columns 5-7. The advantages caused by the extra input L to the second stage are transferred to the system efficiency for 6 of the 30 units. The observed level of extra input L for these units is below the average level of input L for the current technology. For example, for DMUs 6 and 9 the improvement on their adjusted efficiency scores (e^l) led also to improvement on their overall efficiency. Theirs levels of input L are L_6=23.25944 and L_9=1.78061 respectively whereas the average level of input L is 95.3524.

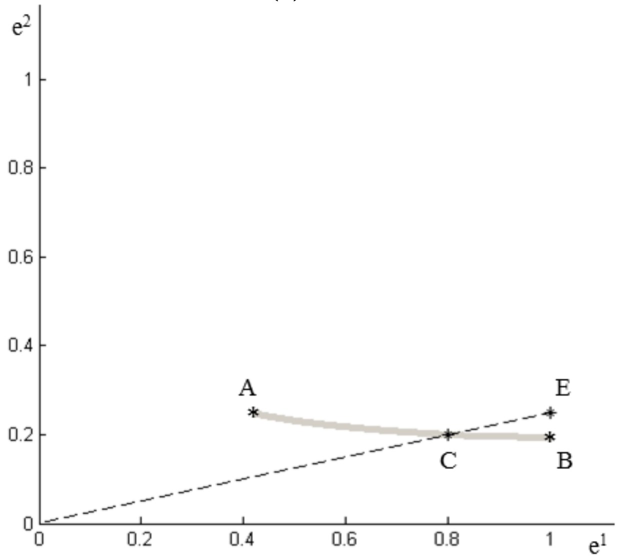
Table 5.7: Results obtained from phases I and II

DMU	θ	$\hat{e}^1$	$\hat{e}^2$	e^I	$e^{II} = \hat{e}^2$	e^o
1	1	1	0.1598	1	0.1598	0.1598
2	1	1	0.2489	1	0.2489	0.2489
3	0.9111	0.9111	0.5219	0.9111	0.5219	0.5219
4	0.9402	0.6982	0.5363	0.6982	0.5363	0.5363
5	1	0.6697	0.3895	0.8497	0.3895	0.3895
6	1	0.5668	1	0.6137	1	0.6137
7	0.7830	0.7830	0.2444	0.7919	0.2444	0.2444
8	1	1	1	1	1	1
9	1	0.9398	1	0.9534	1	0.9534
10	1	1	1	1	1	1
11	1	0.8885	0.8351	0.9288	0.8351	0.8351
12	0.9848	0.9186	0.2662	0.9640	0.2662	0.2662
13	0.9989	0.8495	0.7365	0.9108	0.7365	0.7365
14	0.8614	0.7804	0.2894	0.7804	0.2894	0.2894
15	0.9839	0.9839	0.3719	0.9929	0.3719	0.3719
16	1	0.9225	1	0.9225	1	0.9225
17	0.9920	0.5602	0.9920	0.7049	0.9920	0.7049
18	0.9699	0.6943	0.5028	0.8285	0.5028	0.5028
19	0.9679	0.6746	0.3622	0.6746	0.3622	0.3622
20	1	0.4573	1	0.4573	1	0.4573
21	0.9701	0.6888	0.8244	0.8370	0.8244	0.8244
22	0.9362	0.5490	0.5345	0.7444	0.5345	0.5345
23	0.7984	0.7984	0.2003	0.7999	0.2003	0.2003
24	0.9515	0.9515	0.4583	0.9742	0.4583	0.4583
25	0.9652	0.9652	0.5944	0.9818	0.5944	0.5944
26	0.8659	0.7889	0.8261	0.8804	0.8261	0.8261
27	1	1	1	1	1	1
28	1	0.3599	1	0.3599	1	0.3599
29	1	0.4300	1	0.4308	1	0.4308
30	1	1	1	1	1	1

As shown in Tables 5.6 and 5.7 although the two approaches deem efficient the same units in the second stage and in overall, they yield different stage efficiency scores for 16 of the 30 units. In particular, for the first stage, Li et al (2012) estimate eleven DMUs as efficient whereas only six of them are deemed efficient from our approach. In Li et al (2012) the main goal is the detection of the pair of stage efficiencies that provides the maximal squared geometric average, i.e. the maximal overall efficiency, which is in compliance with the optimality criterion in the multiplicative approach of Kao and Hwang (2008). As a result, when there is a large discrepancy between the independent (ideal) efficiency scores, then the pair of stage efficiencies that maximizes the overall efficiency would tend to lie on the extreme points of the Pareto front. For instance for DMUs 7 and 23, the corresponding solutions obtained from Li et al (2012) are located on the extreme points B of the Pareto fronts as portrayed by Figures 5.8a and 5.8b respectively. As noted, this phenomenon can be associated with the large discrepancy between the independent efficiency scores, i.e. for unit 7 the independent scores are $E^1 =1$, $E^2 =0.3121$ and for unit 23 the corresponding scores are $E^1 =1$, $E^2 =0.2509$. On the contrary, DMUs 7 and 23 are not deemed efficient in stage-1 by our approach, the corresponding Pareto optimal points obtained by our two-phase procedure are depicted by points C on Figures 5.8a and 5.8b. The independent (ideal) stage efficiencies are represented by points E while the extreme Pareto optimal solutions by points A and B. Notice that the points C are formed by the intersection of the ray from the origin to the point E with the Pareto front.



(a) Unit 7



(a) Unit 23

Fig. 5.8: Pareto fronts and Pareto optimal solutions of units 7 and 23

The Fig. 5.9 depicts the Pareto fronts (curves AB) of four indicative units (namely, units 3, 4, 14 and 26) shown in Table 5.7. The independent (ideal) stage efficiency scores are portrayed by points E. The stage efficiencies derived by our two-phase procedure are represented by points C, while the extreme points A and B on the Pareto fronts are obtained by solving analogous LPs with programs (5.12) and (5.13), see the Appendix for details.

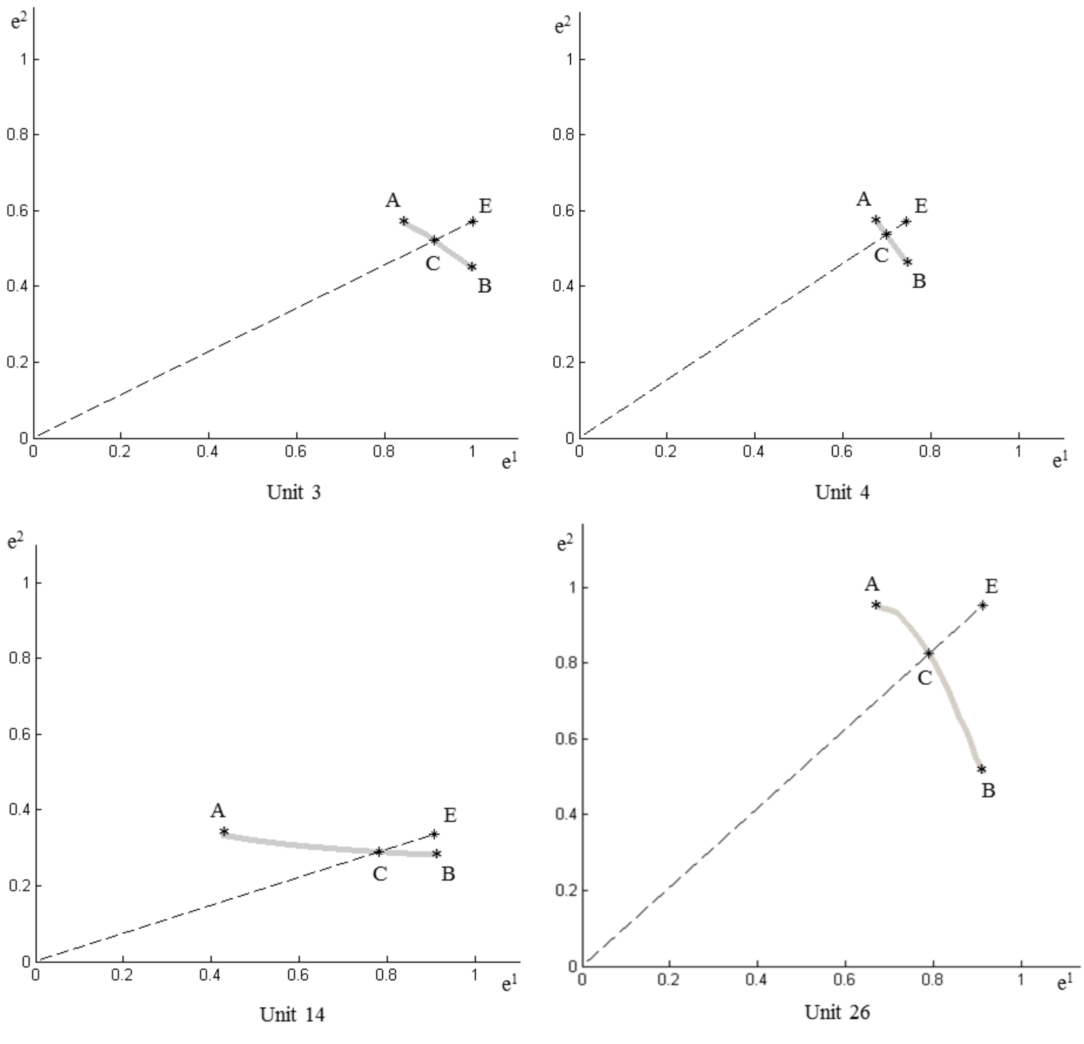


Fig. 5.9: Pareto fronts and Pareto optimal solutions for 4 indicative units

5.1.3 Two-stage process with extra outputs from stage-1 (Type III)

In the structure of Type III, the first stage produces some final outputs *K* that exit the system, beyond the intermediate measures *Z* as depicted in Figure 5.10.

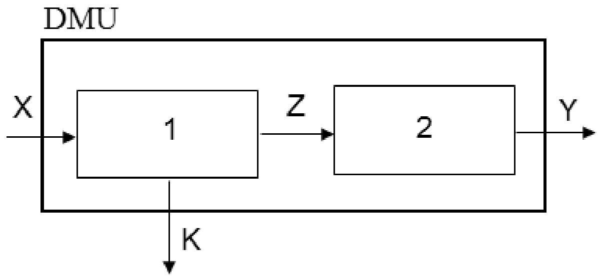


Fig. 5.10: The two-stage process of Type III

An alternative portrayal of the conventional illustration of Fig. 5.10 is given below in Fig. 5.11.

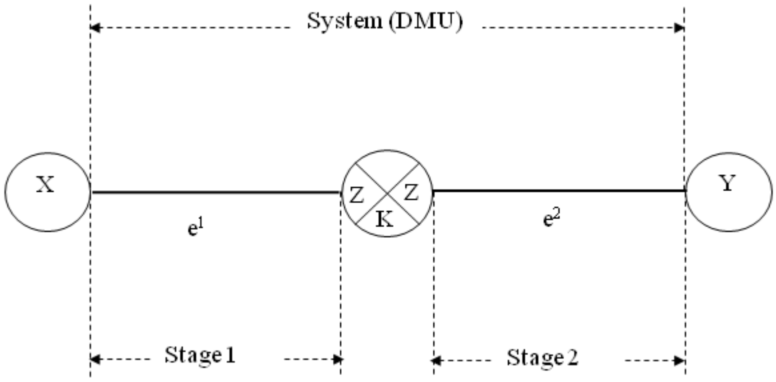


Fig. 5.11: An alternative representation of the two-stage process of Figure 5.10

In this case, the efficiency of the first and the second stage of DMU *j* are typically defined as follows:

$$e_j^1 = \frac{\varphi Z_j + hK_j}{\eta X_j} \quad , \quad e_j^2 = \frac{\omega Y_j}{\varphi Z_j}$$

Working like in the previous sections, we provide below the bi-objective program and the two-phase procedure for estimating the stage efficiencies and the overall efficiency of the two stage process of Figure 5.10.

$$
\begin{aligned}
& max\ e^1 = \frac{\varphi Z_{j_0} + hK_{j_0}}{\eta X_{j_0}} \\
& max\ e^2 = \frac{\omega Y_{j_0}}{\varphi Z_{j_0}} \\
& s.t. \\
& \varphi Z_j + hK_j - \eta X_j \leq 0, \ \ j = 1, \dots, n \\
& \omega Y_j - \varphi Z_j \leq 0, \ \ j = 1, \dots, n \\
& \eta \geq \varepsilon, \varphi \geq \varepsilon, \omega \geq \varepsilon, h \geq \varepsilon
\end{aligned} \tag{5.21}
$$

Applying the C-C transformation with respect to the first objective function we derive the following model:

$$
\begin{aligned}
& max\ e^1 = wZ_{j_0} + \mu K_{j_0} \\
& max\ e^2 = \frac{uY_{j_0}}{wZ_{j_0}} \\
& s.t. \\
& vX_{j_0} = 1 \\
& wZ_j + \mu K_j - vX_j \leq 0, \ \ j = 1, \dots, n \\
& uY_j - wZ_j \leq 0, \ \ j = 1, \dots, n \\
& v \geq \varepsilon, w \geq \varepsilon, u \geq \varepsilon, \mu \geq \varepsilon
\end{aligned} \tag{5.22}
$$

Phase I:

$$
\begin{aligned}
& max\ \ \theta \\
& s.t. \\
& wZ_{j_0} + \mu K_{j_0} \geq \theta E^1_{j_0} \\
& \frac{uY_{j_0}}{wZ_{j_0}} \geq \theta E^2_{j_0} \\
& vX_{j_0} = 1 \\
& wZ_j + \mu K_j - vX_j \leq 0, \ \ j = 1, \dots, n \\
& uY_j - wZ_j \leq 0, \ \ j = 1, \dots, n \\
& v \geq \varepsilon, w \geq \varepsilon, u \geq \varepsilon, \mu \geq \varepsilon, \theta \geq 0
\end{aligned} \tag{5.23}
$$

Let (θ^*, v^*, w^*, u^*, μ^*) be an optimal solution of model (5.23) and the stage efficiencies as follows:

$$
e^{1*}_{j_o} = \frac{w^* Z_{j_o} + \mu^* K_{j_o}}{v^* X_{j_o}} = w^* Z_{j_0} + \mu^* K_{j_0}\ , \quad e^{2*}_{j_o} = \frac{u^* Y_{j_o}}{w^* Z_{j_o}}
$$

The phase I solution is weakly Pareto optimal to the bi-objective program (5.22):

<u>Phase II:</u>

Solve the following linear program:

$$\begin{aligned}
&\max\ s_1\\
&s.t.\\
&wZ_{j_0} + \mu K_{j_0} - s_1 = e_{j_0}^{1*}\\
&uY_{j_0} - e_{j_0}^{2*} wZ_{j_0} \geq 0\\
&vX_{j_0} = 1\\
&wZ_j + \mu K_j - vX_j \leq 0,\ \ j = 1, \dots, n\\
&uY_j - wZ_j \leq 0,\ \ j = 1, \dots, n\\
&v \geq \varepsilon, w \geq \varepsilon, u \geq \varepsilon, \mu \geq \varepsilon\\
&0 \leq s_1 \leq E_{j_0}^1 - e_{j_0}^{1*}
\end{aligned} \tag{5.24}$$

Let $(\hat{s}_1, \hat{v}, \hat{w}, \hat{u}, \hat{\mu})$ be the optimal solution of (5.24). If $\hat{s}_1 > 0$, then the solution is Pareto optimal and the stage efficiency scores are:

$$\hat{e}_{j_o}^1 = \frac{\hat{w}Z_{j_o} + \hat{\mu}K_{j_o}}{\hat{v}X_{j_o}},\ \hat{e}_{j_o}^2 = \frac{\hat{u}Y_{j_o}}{\hat{w}Z_{j_o}}$$

Otherwise (when $\hat{s}_1 = 0$), solve the following program to get a Pareto optimal solution (s'_2, v', w', u', μ') and the stage efficiency scores:

$$e'^1_{j_o} = \frac{w'Z_{j_o} + \mu' K_{j_o}}{v'X_{j_o}},\ e'^2_{j_o} = \frac{u'Y_{j_o}}{w'Z_{j_o}}$$

$$\begin{aligned}
&\max\ s_2\\
&s.t.\\
&wZ_{j_0} + \mu K_{j_0} \geq e_{j_0}^{1*}\\
&\frac{uY_{j_0}}{wZ_{j_0}} - s_2 = e_{j_0}^{2*}\\
&vX_{j_0} = 1\\
&wZ_j + \mu K_j - vX_j \leq 0,\ \ j = 1, \dots, n\\
&uY_j - wZ_j \leq 0,\ \ j = 1, \dots, n\\
&v \geq \varepsilon, w \geq \varepsilon, u \geq \varepsilon, \mu \geq \varepsilon\\
&0 \leq s_2 \leq E_{j_0}^2 - e_{j_0}^{2*}
\end{aligned} \tag{5.25}$$

As noticed in the previous section, using the minimum of e^1, e^2 as the system overall efficiency might underestimate it because the potential advantages caused by the economical low level of the extra inputs to stage-2 may not be reflected to the overall efficiency. Therefore, we adjusted the stage efficiency scores in order to capture and transfer the potential improvements to the overall efficiency. To be more specific, in the two stage process of Figure 5.10 because the first stage has an extra output K, flowing directly out of the system, the overall efficiency of the system should be contributed by K without being stuck by the "weak link". For example, suppose that e^1 is greater than e^2 and this difference is mainly caused by the prominent high output on K. Obviously, still assuming the minimum of these two scores, i.e. e^2, as the overall efficiency for the system does not suffice and the potential high level of output K should be valued by the system. Given the optimal solution $(\hat{v}, \hat{w}, \hat{u}, \hat{\gamma}, \hat{\mu})$, as derived by the two-phase procedure above, we modify the stage efficiencies as follows in order to absorb the aforementioned effects in favor of the system:

$$e_{j_0}^{I} = \frac{\hat{w}Z_{j_o} + \hat{\mu}K_{j_o}}{\hat{v}X_{j_o}}, \; e_{j_0}^{II} = \frac{\hat{u}Y_{j_o} + \hat{\mu}K_{j_o}}{\hat{w}Z_{j_o} + \hat{\mu}K_{j_o}} \tag{5.26}$$

The adjusted stage efficiencies in (5.26) originate from the network structure in Figure 5.12a which is a recast of the two stage process depicted in Figure 5.10. The valued external output $\hat{\mu}K$ of the first stage is used as an input and, at the same time, as an output from the second stage. The term $\hat{\mu}K$ is added to the numerator and to the denominator of the original efficiency ratio to obtain the modified efficiency score $e_{j_0}^{II}$ of the second stage. As already noted, in Kao (2014a) dummy processes were used in a parallel configuration with the original stages for the representation of the modified network structure. Figure 5.12b depicts the modified two stage process in accordance with Kao (2014a).

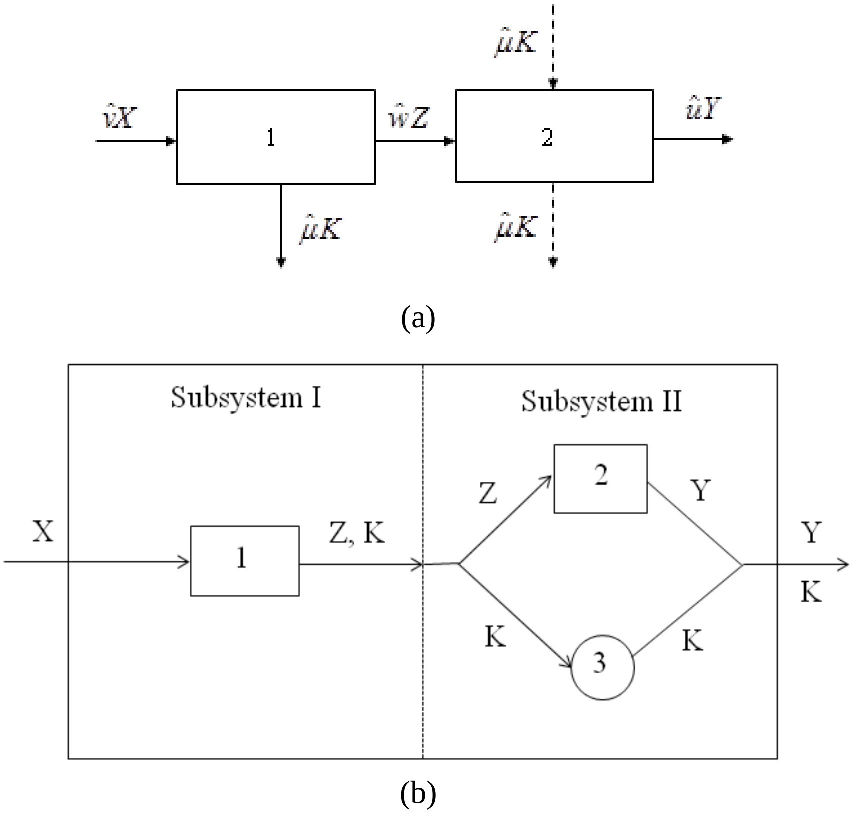


Fig. 5.12: The modified process of Figure 5.10

The adjusted efficiency scores $e_{j_0}^{I}$ and $e_{j_0}^{II}$ are related to their original counterparts $\hat{e}_{j_0}^{1}$ and $\hat{e}_{j_0}^{2}$ as follows:

$$\begin{aligned} e_{j_0}^{I} &= \hat{e}_{j_0}^{1} \\ e_{j_0}^{II} &= \frac{\hat{w}Z_{j_0}}{\hat{w}Z_{j_0}+\hat{\mu}K_{j_0}} \times \hat{e}_{j_0}^{2} + \frac{\hat{\mu}K_{j_0}}{\hat{w}Z_{j_0}+\hat{\mu}K_{j_0}} \times 1 \end{aligned} \quad (5.27)$$

The efficiency of the first stage remains unchanged, whereas for the second stage the adjusted efficiency $e_{j_0}^{II}$ is a weighted average of the original efficiency score $\hat{e}_{j_0}^{2}$ and the efficiency score of the dummy process 3 associated with the second stage, which is 1. Having derived the adjusted efficiency scores, then the overall efficiency of the evaluated unit is $e_{j_0}^{o} = min\{e_{j_0}^{I}, e_{j_0}^{II}\}$.

Illustration

In the following, we apply the "weak-link" approach to a real example taken from Aviles-Sacoto et al (2015) who studied 37 schools of business. Aviles-Sacoto et al (2015) viewed the undergraduate business programmes as two stage processes and they developed a network DEA approach to deal with a peculiar setting where some of the intermediate measures are inputs to the second stage and at the same time external outputs from that stage. At the first stage the assessment is focused on the outcomes which the students achieve before graduation while the second stage captures the accomplishments after graduation. However, for illustration purposes we examine only the scenario where the measures with the controversial role are operate only as external outputs of the first stage. As inputs to stage-1 are used the percentage of applicants rejected (X1), the academic rating (X2) and the percentage of students in top 25% of their classes (X3). The external output of stage-1 is the percentage of students receiving internships (K) while the intermediate measures, outputs of stage-1 and inputs to stage-2, are the percentage of accepted applicants enrolled (Z1) and the percentage of students receiving institutional scholarships (Z2). The external outputs of the second stage are the percentage of students who get jobs (Y1). Table 5.8 exhibits the data set as well as the independent stage efficiencies (last two columns).

Table 5.8: Data from Aviles-Sacoto et al (2015) and independent stage efficiencies

DMU	X1	X2	X3	Z1	Z2	K	Y1	E^1	E^2
1	95	73	90	99	37	79	95	0.8923	0.5603
2	97	68	93	96	20	86.3	78	0.9073	0.7410
3	100	89	97	65	58	90.9	94	0.8861	0.5447
4	100	83	91	75	43	85.2	93	0.8846	0.5898
5	98	81	92	80	42	85.3	93	0.8884	0.5789
6	98	70	92	99	37	86.2	84	0.9238	0.4954
7	99	50	83	82	38	89.7	81	1	0.5233
8	95	70	89	28	30	84.1	83	0.8410	1
9	95	56	74	64	19	84.2	100	1	1
10	92	62	81	36	53	92.3	93	1	0.7216
11	96	68	83	100	48	77.3	97	0.9870	0.5052
12	87	61	94	26	61	77.7	100	0.7826	0.7813
13	74	21	77	20	65	70.8	71	1	0.5625
14	87	60	85	28	46	85.1	64	0.8992	0.5949
15	91	62	89	90	42	78.8	93	0.8915	0.5456
16	71	27	76	46	60	73.8	90	1	0.5889
17	88	54	88	30	50	89	82	0.9129	0.7047
18	79	57	80	33	60	86.2	90	0.9903	0.6633
19	93	64	99	21	50	97	85	0.8958	0.8134
20	98	66	94	95	31	61.7	97	0.7262	0.6303

DMU	X1	X2	X3	Z1	Z2	K	Y1	E^1	E^2
21	89	29	72	45	31	74.9	72	1	0.6947
22	86	48	77	45	26	73.7	95	0.8619	1
23	73	47	76	52	79	74.9	80	1	0.4212
24	13	42	75	74	15	63.6	55	1	0.6967
25	90	57	84	25	90	56.4	62	1	0.3621
26	91	58	72	35	42	82.1	76	1	0.6868
27	87	30	81	10	50	73.5	90	0.9075	1
28	76	41	83	23	68	68.8	82	0.8612	0.6083
29	29	26	75	33	63	62.9	66	1	0.4705
30	91	43	77	99	46	68.3	63	0.9835	0.3367
31	70	27	71	99	26	67.8	92	1	0.6723
32	89	33	71	53	38	58.1	74	0.81781	0.5937
33	73	40	80	38	59	87	68	1	0.4832
34	73	51	84	19	70	87.6	80	1	0.6034
35	89	46	72	99	71	53.2	75	1	0.3221
36	95	41	72	61	22	61.2	78	0.80158	0.7579
37	35	18	84.25	68	23	47.1	69	1	0.6175

Table 5.9 presents the stage efficiency scores obtained by our two-phase procedure as well as the adjusted efficiencies that take into account the potential improvements to the overall efficiency. The results show that the phase II discovers for some units, namely 2, 22, 27, 31 and 34, that their second stage efficiencies could be improved. More specifically, for these DMUs model (5.25) reveals that the optimal solutions derived by model (5.23) are weak Pareto and provides new solutions which are Pareto optimal.

Table 5.9: Results obtained from phases I and II

DMU	θ	e^1	e^2	$\hat{e}^1$	$\hat{e}^2$	$e^I = \hat{e}^1$	e^{II}	e^o
1	0.9584	0.8551	0.5370	0.8551	0.5370	0.8551	0.8745	0.8551
2	0.9149	0.8301	0.6780	0.8301	0.6782	0.8301	1	0.8301
3	0.9931	0.8800	0.5410	0.8800	0.5410	0.8800	0.9007	0.8800
4	0.9862	0.8724	0.5817	0.8724	0.5817	0.8724	0.9004	0.8724
5	0.9785	0.8693	0.5665	0.8693	0.5665	0.8693	0.8948	0.8693
6	0.9639	0.8904	0.4775	0.8904	0.4775	0.8904	0.8680	0.8680
7	0.9918	0.9918	0.5190	0.9918	0.5190	0.9918	0.9021	0.9021
8	1	0.8410	1	0.8410	1	0.8410	1	0.8410
9	1	1	1	1	1	1	1	1
10	1	1	0.7216	1	0.7216	1	1	1
11	0.9608	0.9483	0.4854	0.9483	0.4854	0.9483	0.7638	0.7638
12	0.9758	0.7637	0.7624	0.7637	0.7624	0.7637	0.9535	0.7637
13	1	1	0.5625	1	0.5625	1	0.9833	0.9833
14	1	0.8992	0.5949	0.8992	0.5949	0.8992	1	0.8992
15	0.9678	0.8628	0.5280	0.8628	0.5280	0.8628	0.8719	0.8628
16	1	1	0.5889	1	0.5889	1	0.9198	0.9198
17	1	0.9129	0.7047	0.9129	0.7047	0.9129	1	0.9129
18	0.9892	0.9796	0.6562	0.9796	0.6562	0.9796	0.9356	0.9356
19	1	0.8958	0.8134	0.8958	0.8134	0.8958	1	0.8958
20	0.9337	0.6781	0.5885	0.6781	0.5885	0.6781	0.8060	0.6781
21	1	1	0.6947	1	0.6947	1	1	1
22	0.9989	0.8610	0.9989	0.8610	1	0.8610	1	0.8610
23	1	1	0.4212	1	0.4212	1	0.6578	0.6578
24	1	1	0.6967	1	0.6967	1	0.8913	0.8913
25	0.8154	0.8154	0.2952	0.8154	0.2952	0.8154	0.4122	0.4122
26	1	1	0.6868	1	0.6868	1	1	1
27	0.9974	0.9051	0.9974	0.9051	1	0.9051	1	0.9051
28	0.9470	0.8156	0.5761	0.8156	0.5761	0.8156	0.8643	0.8156
29	1	1	0.4705	1	0.4705	1	0.4705	0.4705
30	0.9736	0.9575	0.3278	0.9575	0.3278	0.9575	0.6595	0.6595
31	0.9522	0.9522	0.6402	0.9522	0.6427	0.9522	1	0.9522
32	0.9999	0.8177	0.5936	0.8177	0.5936	0.8177	0.8753	0.8177
33	1	1	0.4832	1	0.4832	1.0000	0.8932	0.8932
34	0.9715	0.9715	0.5862	0.9715	0.5913	0.9715	1	0.9715
35	1	1	0.3221	1	0.3221	1	0.3221	0.3221
36	0.9761	0.7825	0.7398	0.7825	0.7398	0.7825	0.9469	0.7825
37	0.9956	0.9956	0.6147	0.9956	0.6147	0.9956	0.6147	0.6147

Figure 5.13 depicts the Pareto fronts (curves AB) of four indicative units (namely units 6, 28, 11 and 15). The independent (ideal) stage efficiency scores are portrayed by points E. The stage efficiencies derived by our two-phase procedure are represented by points C, while the extreme points A and B on the Pareto fronts are obtained by solving analogous LPs with programs (5.12) and (5.13), see the Appendix for details.

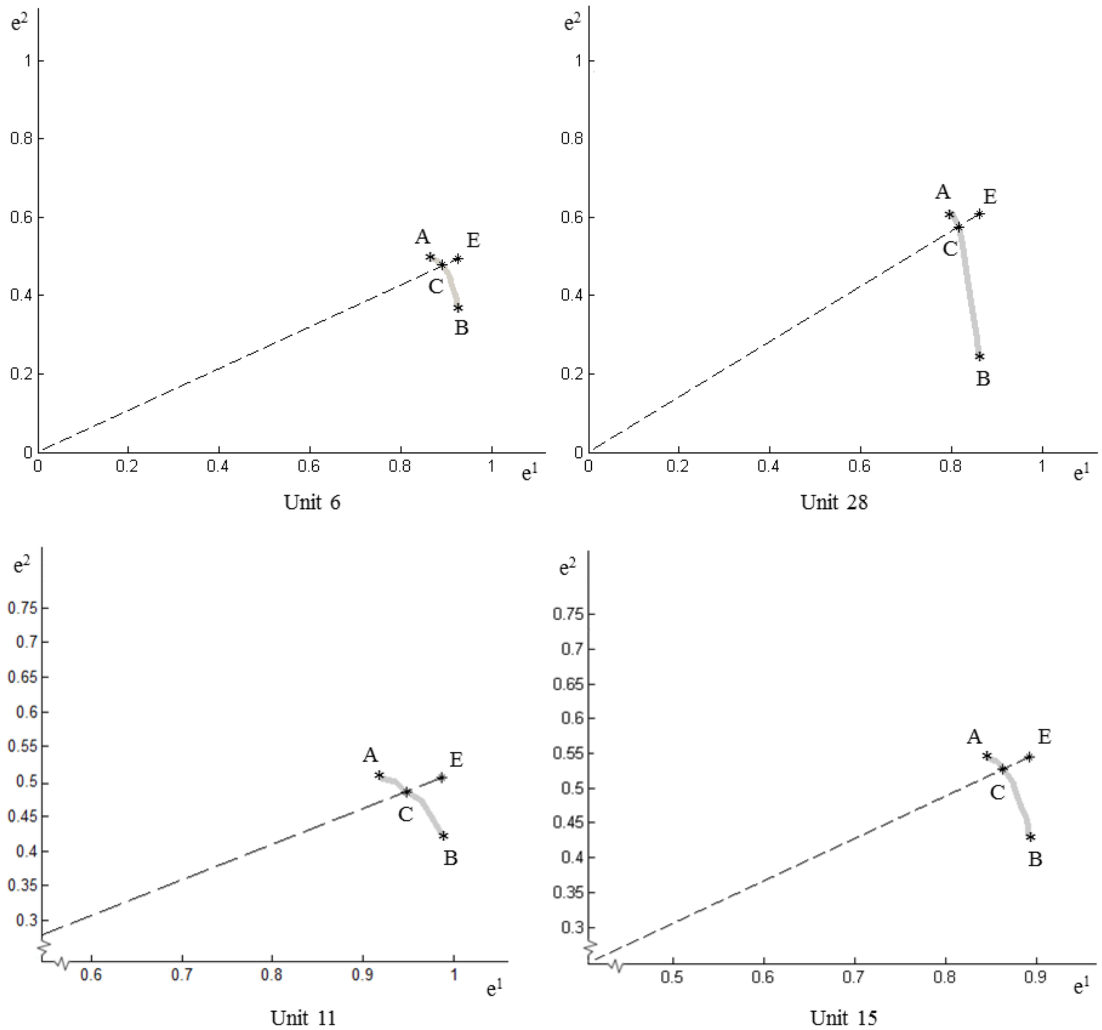


Fig. 5.13: Pareto fronts and Pareto optimal solutions for 4 indicative units

5.1.4 General two-stage process (Type IV)

The general two-stage process depicted in Figure 5.14 differentiates from the other types, in that the first stage produces some final outputs *K* that exit the system, beyond the intermediate measures *Z*, and the second stage uses some extra external inputs *L*.

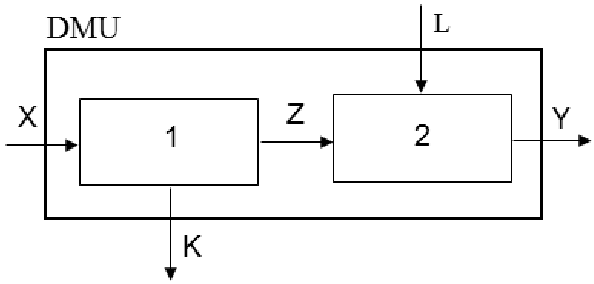


Fig. 5.14: The two-stage process of Type IV

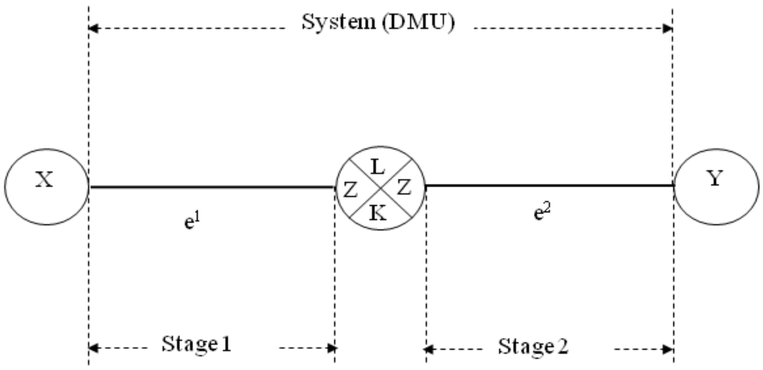


Fig. 5.15: An alternative representation of the two-stage process of Type IV

In this case, the efficiencies of the first and the second stage of DMU j are defined as follows:

$$e_j^1 = \frac{wZ_j+\mu K_j}{vX_j}\ , e_j^2 = \frac{uY_j}{wZ_j+\gamma L_j}$$

Analogously to the other types of processes elaborated in previous sections, the two-phase procedure for estimating the stage efficiencies as well as the overall system efficiency is as follows:

$$\begin{aligned}
&max\ e^1 = \frac{\varphi Z_{j_0} + hK_{j_0}}{\eta X_{j_0}} \\
&max\ e^2 = \frac{\omega Y_{j_0}}{\varphi Z_{j_0} + gL_{j_0}} \\
&s.t. \\
&\varphi Z_j + hK_j - \eta X_j \leq 0,\ \ j = 1, \dots, n \\
&\omega Y_j - \varphi Z_j - gL_j \leq 0,\ \ j = 1, \dots, n \\
&\eta \geq \varepsilon, \varphi \geq \varepsilon, \omega \geq \varepsilon, h \geq \varepsilon, g \geq \varepsilon
\end{aligned} \tag{5.28}$$

By applying the C-C transformation with respect the first objective function we derive the following program:

$$
\begin{aligned}
& max\, e^1 = wZ_{j_0} + \mu K_{j_0} \\
& max\, e^2 = \frac{uY_{j_0}}{wZ_{j_0} + \gamma L_{j_0}} \\
& s.t. \\
& vX_{j_0} = 1 \\
& wZ_j + \mu K_j - vX_j \leq 0, \;\; j = 1, \dots, n \\
& uY_j - wZ_j - \gamma L_j \leq 0, \;\; j = 1, \dots, n \\
& v \geq \varepsilon, w \geq \varepsilon, u \geq \varepsilon, \gamma \geq \varepsilon, \mu \geq \varepsilon
\end{aligned}
\tag{5.29}
$$

<u>Phase I:</u>

$$
\begin{aligned}
& max\;\; \theta \\
& s.t. \\
& wZ_{j_0} + \mu K_{j_0} \geq \theta E^1_{j_0} \\
& \frac{uY_{j_0}}{wZ_{j_0} + \gamma L_{j_0}} \geq \theta E^2_{j_0} \\
& vX_{j_0} = 1 \\
& wZ_j + \mu K_j - vX_j \leq 0, \;\; j = 1, \dots, n \\
& uY_j - wZ_j - \gamma L_j \leq 0, \;\; j = 1, \dots, n \\
& v \geq \varepsilon, w \geq \varepsilon, u \geq \varepsilon, \gamma \geq \varepsilon, \mu \geq \varepsilon, \theta \geq 0
\end{aligned}
\tag{5.30}
$$

Let (θ^*, v^*, w^*, u^*, γ^*, μ^*) be an optimal solution of model (5.30), then the stage efficiencies are calculated as follows:

$$
e^{1*}_{j_0} = \frac{w^* Z_{j_0} + \mu^* K_{j_0}}{v^* X_{j_0}} = w^* Z_{j_0} + \mu^* K_{j_0}, \;\; e^{2*}_{j_0} = \frac{u^* Y_{j_0}}{w^* Z_{j_0} + \gamma^* L_{j_0}}
$$

The phase I solution is weakly Pareto optimal to the bi-objective program (5.29):

Phase II:

Solve the following linear program:

$$
\begin{aligned}
& max\ \ s_1 \\
& s.t. \\
& wZ_{j_0} + \mu K_{j_0} - s_1 = e_{j_0}^{1*} \\
& uY_{j_0} - e_{j_0}^{2*}\left(wZ_{j_0} + \gamma L_{j_0}\right) \geq 0 \\
& vX_{j_0} = 1 \\
& wZ_j + \mu K_j - vX_j \leq 0,\ \ j = 1, \dots, n \\
& uY_j - wZ_j - \gamma L_j \leq 0,\ \ j = 1, \dots, n \\
& v \geq \varepsilon, w \geq \varepsilon, u \geq \varepsilon, \gamma \geq \varepsilon, \mu \geq \varepsilon \\
& 0 \leq s_1 \leq E_{j_0}^1 - e_{j_0}^{1*}
\end{aligned}
\tag{5.31}
$$

Let $(\hat{s}_1, \hat{v}, \widehat{w}, \hat{u}, \hat{\gamma}, \hat{\mu})$ be the optimal solution of (5.31). If $\hat{s}_1 > 0$, then the solution is Pareto optimal and the stage efficiency scores are:

$$\hat{e}_{j_0}^1 = \frac{\widehat{w}Z_{j_o} + \hat{\mu}K_{j_o}}{\hat{v}X_{j_o}}\ ,\ \hat{e}_{j_0}^2 = \frac{\hat{u}Y_{j_o}}{\widehat{w}Z_{j_o} + \hat{\gamma}L_{j_o}}$$

Otherwise, solve the following program to get a Pareto optimal solution $(s'_2, v', w', u', \gamma', \mu')$ and the stage efficiency scores:

$$
\begin{aligned}
& max\ \ s_2 \\
& s.t. \\
& wZ_{j_0} + \mu K_{j_0} \geq e_{j_0}^{1*} \\
& \frac{uY_{j_0}}{wZ_{j_0} + \gamma L_{j_0}} - s_2 = e_{j_0}^{2*} \\
& vX_{j_0} = 1 \\
& wZ_j + \mu K_j - vX_j \leq 0,\ \ j = 1, \dots, n \\
& uY_j - wZ_j - \gamma L_j \leq 0,\ \ j = 1, \dots, n \\
& v \geq \varepsilon, w \geq \varepsilon, u \geq \varepsilon, \gamma \geq \varepsilon, \mu \geq \varepsilon \\
& 0 \leq s_2 \leq E_{j_0}^2 - e_{j_0}^{2*}
\end{aligned}
\tag{5.32}
$$

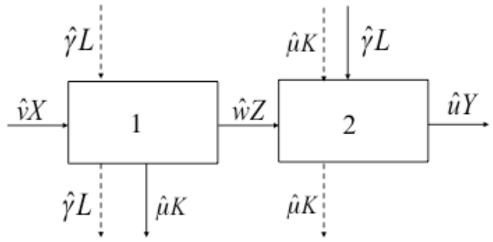
$$e'^1_{j_0} = \frac{w'Z_{j_o} + \mu' K_{j_o}}{v'X_{j_o}}\ ,\ e'^2_{j_0} = \frac{u'Y_{j_o}}{w'Z_{j_o} + \gamma' L_{j_o}}$$

Model (5.32) is non-linear but it can be solved by bisection search in terms of s_2 in the bounded interval $[0, E_{j_0}^2 - e_{j_0}^{2*}]$. Obviously, if the stage efficiency scores, derived from the phase I, are both equal to the corresponding independent efficiency scores, then there is no need to run the phase II since the Pareto front degenerates to the point $\left(E_{j_0}^1, E_{j_0}^2\right)$.

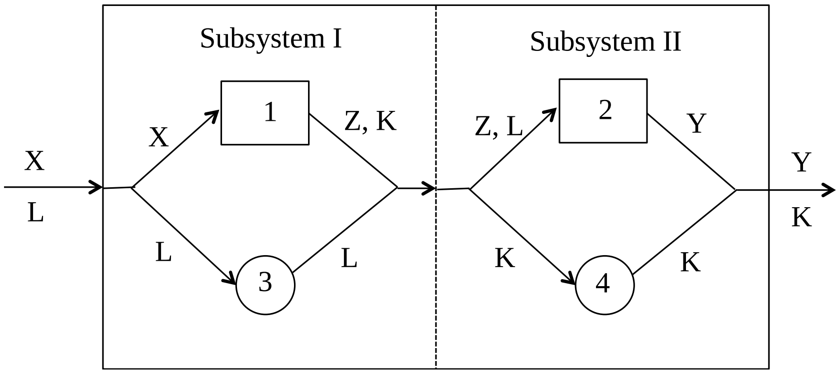
As we already remarked when extra inputs or outputs exist then using the minimum of e^1, e^2 the system efficiency might underestimate it. Indeed, when *K* is missing, as in Fig 5.5, the system should benefit from a potential economical level of *L*. Analogously when the external input L to the second stage is missing, as in Fig 5.10, the overall efficiency of the system should be contributed by a potential prominent high output on K. However these relative advantages of the unit against its peers on L and K may be ignored when taking the overall efficiency as the minimum of the stage efficiencies. In a general two-stage process, these effects must be considered jointly, in a manner that the potential advantages-disadvantages of one stage, caused by the extra inputs *L* and outputs *K*, are transferred to the other stage. Given the optimal solution $(\hat{v}, \hat{w}, \hat{u}, \hat{\gamma}, \hat{\mu})$, as obtained from the two-phase procedure above, the stage efficiencies are adjusted to bring into play these in favor of the system efficiency as follows:

$$e_{j_0}^I = \frac{\hat{w}Z_{j_o} + \hat{\mu}K_{j_o} + \hat{\gamma}L_{j_o}}{\hat{v}X_{j_o} + \hat{\gamma}L_{j_o}},\ e_{j_0}^{II} = \frac{\hat{u}Y_{j_o} + \hat{\mu}K_{j_o}}{\hat{w}Z_{j_o} + \hat{\mu}K_{j_o} + \hat{\gamma}L_{j_o}} \tag{5.33}$$

Figure 5.16a depicts a modification of the network structure of Figure 5.14, the adjusted stage efficiencies that correspond to the modified efficiencies are given in (5.33). The valued input $\hat{\gamma}L$ to the second stage is used as an input and, at the same time, as an output from the first stage. So, the term $\hat{\gamma}L$ is added to the numerator and to the denominator of the original efficiency ratio of stage-1 to obtain the modified efficiency score $e_{j_0}^I$. Analogous is the treatment of the valued external output $\hat{\mu}K$ of the first stage in order to obtain the modified efficiency score $e_{j_0}^{II}$ of the second stage. The modified structure in Figure 5.16b is an alternative, yet equivalent representation of the structure in Figure 5.16a, introduced by Kao (2014a) to deal with general multi-stage systems. In particular, two dummy processes, labelled 3 and 4, were added in a parallel configuration with the original stages 1 and 2 respectively. As it is noticed, each subsystem has a parallel structure composed of one real and one dummy process.



(a)



(b)

Fig. 5.16: The modified process of Figure 5.14

The adjusted efficiency scores $e_{j_0}^{I}$ and $e_{j_0}^{II}$ are related to their original counterparts $\hat{e}_{j_0}^{1}$ and $\hat{e}_{j_0}^{2}$ as follows:

$$e_{j_0}^{I} = \frac{\hat{v}X_{j_0}}{\hat{v}X_{j_0} + \hat{\gamma}L_{j_0}} \times \hat{e}_{j_0}^{1} + \frac{\hat{\gamma}L_{j_0}}{\hat{v}X_{j_0} + \hat{\gamma}L_{j_0}} \times 1$$
$$e_{j_0}^{II} = \frac{\hat{w}Z_{j_0} + \hat{\gamma}L_{j_0}}{\hat{w}Z_{j_0} + \hat{\mu}K_{j_0} + \hat{\gamma}L_{j_0}} \times \hat{e}_{j_0}^{2} + \frac{\hat{\mu}K_{j_0}}{\hat{w}Z_{j_0} + \hat{\mu}K_{j_0} + \hat{\gamma}L_{j_0}} \times 1 \qquad (5.34)$$

That is, $e_{j_0}^{I}$ is a weighted average of the original efficiency score $\hat{e}_{j_0}^{1}$ and the efficiency score of the dummy process 3 associated to the first stage, which is 1. Analogous is the derivation of $e_{j_0}^{II}$ as a function of $\hat{e}_{j_o}^{2}$. Once the adjusted efficiency scores are obtained, the overall efficiency of the evaluated unit is $e_{j_0}^{o} = min\left\{e_{j_o}^{I}, e_{j_o}^{II}\right\}$. As long as the calculation of the adjusted efficiency scores are based on concepts introduced by Kao (2014a), our results are comparable to those obtained by applying his methodology (see illustration below). As already noted though, in essence our methodology differs in the optimality criterion used. The optimality criterion in the "weak-link" approach is to maximize the lowest of the stage efficiencies, following a max-min modelling technique, whilst the optimality criterion in Kao (2014a) is to maximize the overall efficiency. The notions in Kao (2014a) are consistent with the multiplicative decomposition approach (Kao and Hwang, 2008), where the overall efficiency is defined as the ratio of the total virtual external output $(uY + \mu K)$ to the total virtual external input $(vX + \gamma L)$.

Illustration

As an illustrative example we use a synthetic case with 30 DMUs, three inputs to stage-1 (X1–X3), two intermediate measures (Z1, Z2), two final outputs from stage-1 (K1, K2), two extra inputs to stage-2 (L1, L2) and two final outputs from stage-2 (Y1, Y2). Table 5.10 exhibits the data, which are drawn column-wise from a uniform distribution in the intervals given in the last row of Table 5.10.

Table 5.10: Synthetic data for the general two-stage process

DMU	X1	X2	X3	Z1	Z2	K1	K2	L1	L2	Y1	Y2
1	22.3	13.2	54.6	110.1	66.1	21.8	44.6	18	31	13.3	12.5
2	68.3	8.3	15.8	75.4	116.4	19.8	12	19.6	25.8	2.4	18.2
3	52	19.2	31.2	94.3	59.9	47.3	47.4	11.5	22.5	2.3	36
4	31.8	12	40.3	66.4	127.2	10.5	35.8	16.8	37.1	3	19.5
5	95.3	12	29	108.9	52.3	15	22.5	14	27.2	15.8	16.7
6	52.8	6.1	22.6	102.4	78.8	69.6	27	14.8	44.9	12.6	20.4
7	50.5	9.3	48.7	124.6	120.6	52.2	49.8	5.9	38.5	9.5	20.5
8	80.1	17.4	58.4	64.5	131.2	37.7	14.6	10.3	65.6	16.7	39.9
9	53.9	14	36.9	129.8	122.1	60.9	24.1	11.9	49.5	16.8	15
10	20.9	9.5	48.8	66.4	132.5	12.2	68.7	10.1	54.5	10	28.3
11	82.5	7.1	16.8	71.9	138.9	47.7	60.7	5.6	19.1	19.7	33.6
12	27	10.6	25.6	51.9	84.4	47.3	63.3	11	39.6	12.2	43.7
13	49.6	10.7	20.6	125.5	97.3	15.3	32.6	17.7	38.9	18.9	44.7
14	55.7	19.4	46.6	91.5	117.3	79	60.3	11.8	26.4	7.5	38.7
15	55.1	18.2	52.5	90.1	61	12.2	24.9	17	33.5	17.2	43.9
16	66.3	8	34.9	131.1	63.7	57	30.7	10.7	52.5	11.2	15.5
17	93.3	6.3	43.5	53.5	133.9	38.6	32.1	13.4	45	19.7	15.4
18	10.8	11.9	31.5	118.7	89.4	34.9	23.6	11	67.3	8	20.5
19	98.5	6.8	21.3	75.4	133	28	28.9	16.1	26.7	9.5	20.3
20	27.8	17.1	24.9	81	52.2	30.6	14.3	16.9	35.3	17.4	15.4
21	42	7.2	59.7	98.4	147.5	29.2	39.4	14.8	42.3	10.7	44.5
22	98.7	8.5	51	132.8	60.6	27.3	69.3	19.8	61.9	19.9	33.3
23	53.5	15.6	25.7	93.5	121.6	31.3	34.6	19.7	56.5	13	47.6
24	25.1	16.7	56.8	81.6	145.6	62.1	74.8	11.7	17.4	7.6	29.9
25	96.3	15.3	45.1	120.5	133.6	25.7	56.8	19.7	16.9	14.9	38.5
26	97.9	6.8	53.1	103.8	89.8	45.7	49.6	17.7	56.3	4.9	12.5
27	37.4	15	15.5	63.1	128.2	53.1	22	5.5	57.7	5.8	11.8
28	70	12.8	21.5	126.1	97.2	28.3	44.3	11.4	56.7	4.9	47.2
29	24	5.8	33.8	91.2	82.6	73.7	76.2	19.4	42.4	7.8	10.3
30	48.6	18.6	55.9	126	73.8	15.4	57.6	17.1	76.2	13.2	25.6
	[10,100]	[5,20]	[10,60]	[50,150]	[50,150]	[10,80]	[10,80]	[5,20]	[10,80]	[2,20]	[10,50]

Table 5.11 below exhibits the independent stage efficiency scores as well as the results obtained by applying the proposed max-min approach.

Table 5.11: Results obtained from phase II (same as from phase I)

DMU	E^1	E^2	θ	$\hat{e}^1$	$\hat{e}^2$	e^I	e^{II}	e^o
1	0.7338	0.7327	0.9199	0.6750	0.6740	0.7326	0.7607	0.7326
2	1	0.4751	0.9792	0.9792	0.4652	0.9895	0.4732	0.4732
3	0.7608	1	0.9608	0.7310	0.9608	0.7987	0.9735	0.7987
4	0.8938	0.4268	0.8715	0.7789	0.3719	0.8790	0.3719	0.3719
5	0.6914	1	0.8272	0.5719	0.8272	0.6309	0.8272	0.6309
6	1	0.6708	1	1	0.6708	1	0.9507	0.9507
7	0.8812	0.6316	0.9609	0.8467	0.6069	0.9207	0.6117	0.6117
8	0.5496	0.9796	0.9120	0.5012	0.8934	0.5780	0.8934	0.5780
9	0.8656	0.7555	0.9554	0.8270	0.7218	0.8700	0.7253	0.7253
10	1	0.6316	1	1	0.6316	1	0.9089	0.9089
11	1	1	1	1	1	1	1	1
12	1	1	1	1	1	1	1	1
13	1	0.8690	0.9677	0.9677	0.8409	0.9888	0.8409	0.8409
14	0.6687	0.9337	0.9999	0.6686	0.9336	0.6690	0.9997	0.6690
15	0.4215	1	0.9804	0.4132	0.9804	0.6281	0.9804	0.6281
16	0.9813	0.7660	0.7669	0.7525	0.5874	0.7897	0.6318	0.6318
17	1	1	0.9344	0.9344	0.9344	0.9496	0.9383	0.9383
18	1	0.5189	1	1	0.5189	1	0.5538	0.5538
19	1	0.5096	0.9757	0.9757	0.4972	0.9890	0.4972	0.4972
20	0.7676	1	0.9261	0.7109	0.9261	0.7549	0.9317	0.7549
21	1	0.7871	0.9969	0.9969	0.7846	0.9987	0.7846	0.7846
22	0.9591	1	0.8374	0.8031	0.8374	0.8264	0.8437	0.8264
23	0.8255	0.7260	0.9899	0.8171	0.7186	0.9100	0.7367	0.7367
24	1	0.9035	0.9597	0.9597	0.8671	0.9668	0.9571	0.9571
25	0.6613	1	1	0.6613	1	0.7661	1	0.7661
26	0.9447	0.2427	0.9727	0.9189	0.2361	0.9299	0.2688	0.2688
27	1	0.3791	1	1	0.3791	1	0.7900	0.7900
28	1	1	0.9030	0.9030	0.9030	0.9561	0.9143	0.9143
29	1	0.3762	1	1	0.3762	1	0.9807	0.9807
30	0.6246	0.6813	0.8876	0.5544	0.6047	0.6022	0.7161	0.6022

The results obtained by applying the methodology proposed in Kao (2014a) are presented in Table 5.12. The two approaches deem the same DMUs as overall efficient, namely DMUs 11 and 12. As regards the individual stage efficiencies our approach reckons as efficient seven units in stage-1 (i.e. 6, 10, 11, 12, 18, 27 and 29) and three in stage-2 (i.e. 11, 12, and 25), while the approach of Kao (2014a) deems efficient nine units in stage-1 (i.e. 6, 10, 11, 12, 13, 18, 21, 27 and 29) and five in stage-2 (i.e. 5, 11, 12, 20, 25).

Table 5.12: Results obtained from applying the methodology of Kao (2014a)

DMU	e^1	e^2	E_I	E_{II}	e^0
1	0.6883	0.6161	0.6971	0.8960	0.6246
2	0.9565	0.4751	0.9783	0.4902	0.4796
3	0.6375	0.8472	0.9998	0.8472	0.8471
4	0.3856	0.3971	0.3858	0.9990	0.3855
5	0.3873	1	0.5770	1	0.5770
6	1	0.5642	1	0.9999	0.9999
7	0.8020	0.6315	0.9007	0.6455	0.5814
8	0.5419	0.8522	0.7791	0.8527	0.6643
9	0.7223	0.7498	0.7701	0.8361	0.6439
10	1	0.4816	1	0.9999	0.9999
11	1	1	1	1	1
12	1	1	1	1	1
13	1	0.8183	1	0.8425	0.8425
14	0.6137	0.7925	0.9994	0.7925	0.7921
15	0.3624	0.8614	0.8470	0.8614	0.7296
16	0.6033	0.4926	0.6033	0.9998	0.6032
17	0.9395	0.9300	0.9538	0.9342	0.8910
18	1	0.3284	1	0.9998	0.9998
19	0.8792	0.5096	0.9350	0.5378	0.5028
20	0.6679	1	0.7371	1	0.7371
21	1	0.7842	1	0.7842	0.7842
22	0.6347	0.9516	0.7349	0.9612	0.7064
23	0.8094	0.7190	0.9148	0.7527	0.6886
24	0.9331	0.8448	0.9331	0.9999	0.9330
25	0.5644	1	0.9997	1	0.9997
26	0.5552	0.2146	0.5553	0.9996	0.5551
27	1	0.2806	1	0.9998	0.9998
28	0.7393	0.9714	0.9075	0.9714	0.8815
29	1	0.3473	1	0.9998	0.9998
30	0.4752	0.6728	0.5892	0.7602	0.4479

Figure 5.17 exhibits the Pareto fronts (curves AB) of four indicative units (namely, units 1, 5, 16 and 30) shown in Table 5.11. The points E represent the ideal points whose coordinates are the independent efficiency scores of the two stages, whereas the points C correspond to the assessed stage efficiency scores. The points A and B, which represent the extreme points on the Pareto fronts, are obtained by solving analogous LPs to models (5.12) and (5.13), these models are given in the

Appendix. Notice that the points C are formed by the intersection of the ray from the origin to the point E with the Pareto front.

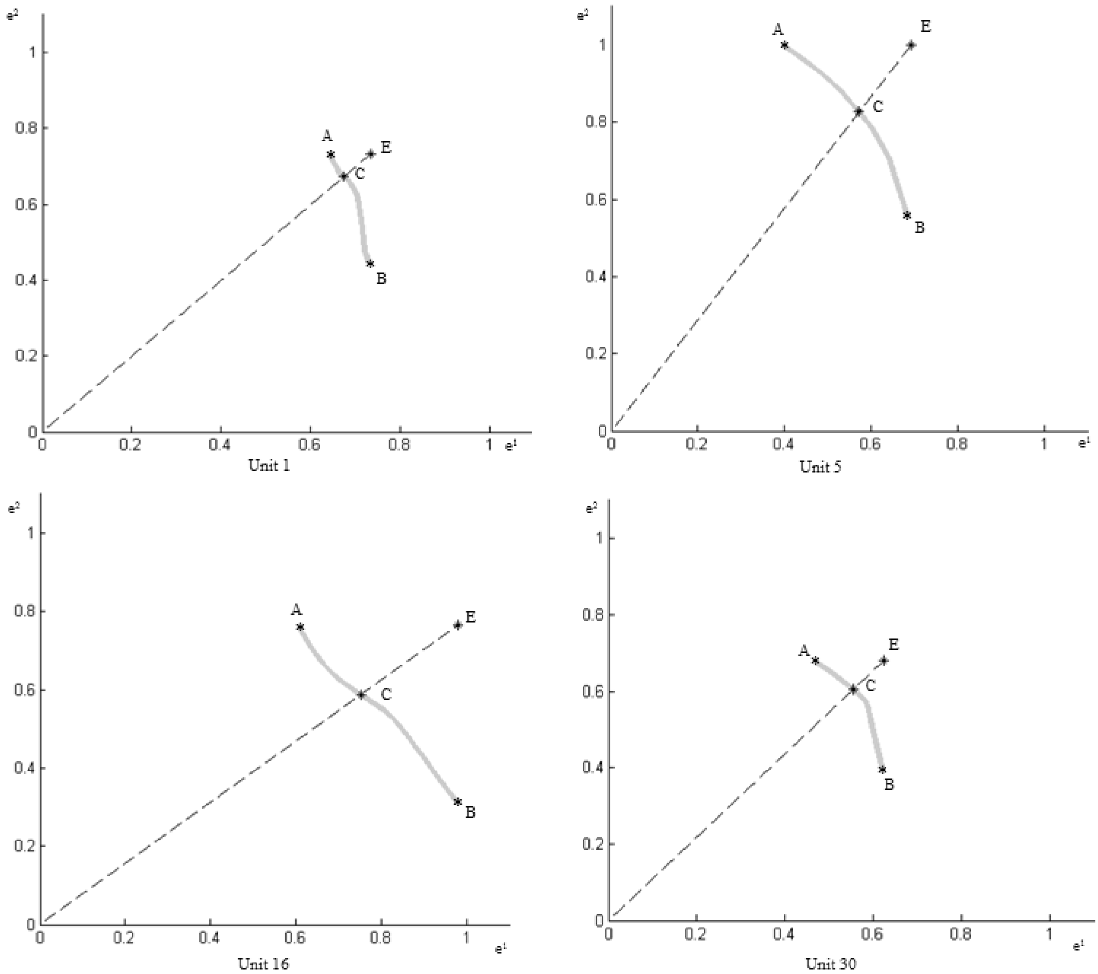


Fig. 5.17: Pareto fronts and Pareto optimal solutions for four indicative units

Before concluding remarks are drawn, it is worth having a more detailed portrayal about the usefulness of the phase II of our approach. For this purpose we create a hypothetical scenario by employing the unit 24. The solution obtained by model (5.30) for the unit 24 is depicted, in Figure 5.18, on the point B(0.9597, 0.8671) which lies on the Pareto front. Therefore, the phase II did not alter this solution. However, in case different weights than the ideal efficiency scores will be used in the max-min model (5.30) then a different solution will derive. For instance, the two extreme points A and F are obtained by using the following weights:

For the point $A(N^1, E^2)$: $q_1^A = \frac{E^2}{N^1+E^2}$, $q_2^A = \frac{N^1}{N^1+E^2}$

For the point $F(E^1, N^2)$: $q_1^F = \frac{N^2}{N^2+E^1}$, $q_2^F = \frac{E^1}{N^2+E^1}$

Bringing into play any weighting scheme between (q_1, q_2) where $q_1 \in [q_1^A, q_1^F]$ and $q_2 \in [q_2^A, q_2^F]$, will yield a solution that is located on the boundary AF. Notice, that the weak efficient solutions lie along the boundary line segment [C→ F), where the parenthesis signifies that F is an open endpoint. The point F depicts a Pareto optimal solution. Assuming for unit 24, the weights $q_1 = 1$ and $q_2 = 0.8997$ instead of $q_1 = E_{24}^1$ and $q_2 = E_{24}^2$ in model (5.30), then the phase I yields the stage efficiencies $e_{24}^1 = 0.9623$ and $e_{24}^2 = 0.8658$, which in Figure 5.18 are represented by point D(0.9623, 0.8658). Point D lies on the segment CD of the boundary which is parallel to the horizontal axis, thus point D is weak Pareto. This issue is treated by the phase II of our method. Indeed, by employing model (5.31) of phase II we derive the Pareto optimal point F(1, 0.8658), which indicates that the efficiency of the first stage is improved. This example demonstrates the potential issues that our phase II can deal with.

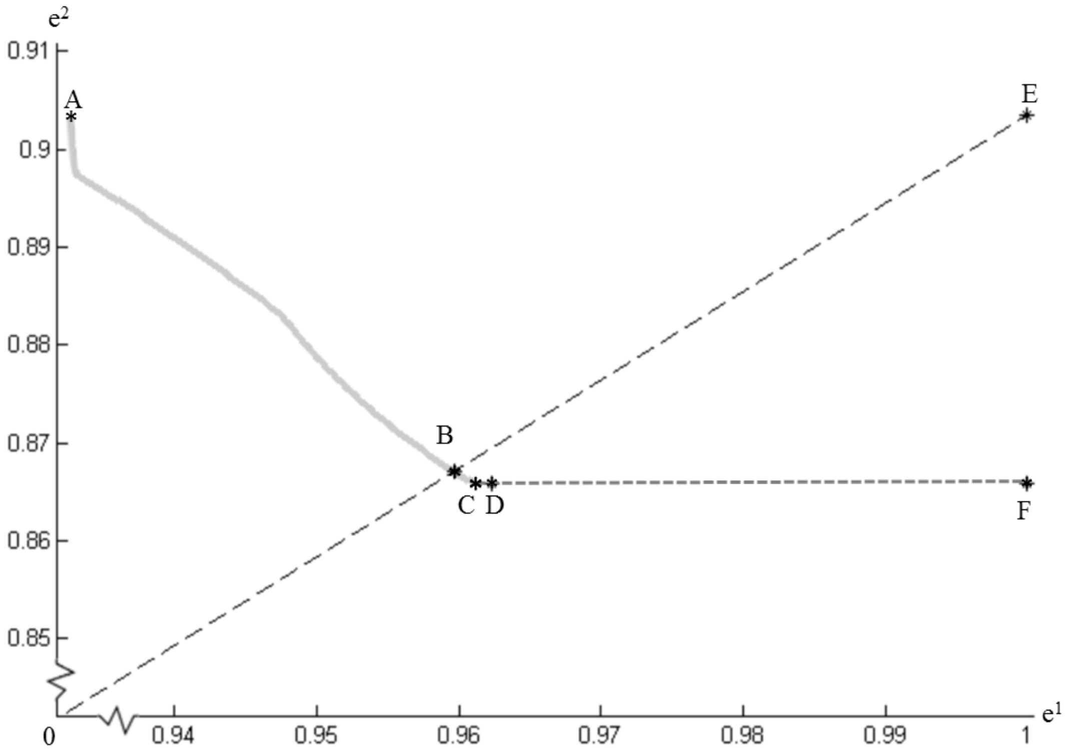


Fig. 5.18: Pareto front and optimal solution for unit 24

Conclusion

We introduced in this chapter a novel approach to two-stage network DEA based on our composition approach presented in Chapter 4. We used a multi-objective formulation and a max-min programming technique to assess the individual stage efficiencies and the overall system efficiency, by maximizing the lowest of the stage efficiencies (weak link). The two-phase procedure that we proposed provides Pareto optimal solutions and a unique point on the Pareto front in the objective functions space, i.e. unique efficiency scores for the two-stages. The search direction is driven by the assumption that the efficiency scores of the two stages are proportional to their independent counterparts. Thus, a point is located on the Pareto front along this direction by maximizing the minimum of the stage efficiency scores. Then the system efficiency is given as the minimum of the stage efficiencies. Although the above assumption is rational, it is not restrictive in our models. External priorities for the two-stages might be assumed to locate a different point on the Pareto front, i.e. different stage and overall efficiencies. A systematic investigation of the sensitivity of the weak link was also provided in order to identify adequately the source of inefficiency. An issue that needs further investigation is the derivation of efficient projections for the inefficient units directly from the proposed models. Our approach is developed to deal with the four types of two-stage processes (as those categorized in Chapter 3). A subject for future research is the extension of the weak-link approach to general network structures involving series and parallel processes.

Chapter 6

Two-stage Network DEA when intermediate measures can be treated as outputs from the second stage

In this chapter we revisit the work of Aviles-Sacoto et al (2015) to provide an alternative modeling approach to the assessment of the efficiencies of undergraduate business programs, in the context of network DEA, in a peculiar situation where one of the intermediate measures must be considered as input to the second stage and, at the same time, as output of the second stage. The motivating situation in Aviles-Sacoto et al (2015) refers to the assessment of the efficiency of undergraduate business programs viewed as two-stage processes, as depicted schematically in Figure 6.1. At the first stage the assessment is focused on the outcomes which the students achieve before graduation while the second stage captures the accomplishments after graduation.

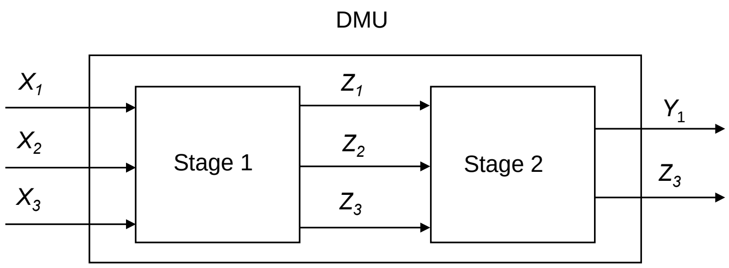


Fig. 6.1: The output Z3 from stage-1 is considered simultaneously as input and output of stage-2

Aviles-Sacoto et al (2015) viewed the undergraduate business programmes as two stage processes and they studied 37 business schools. Table 6.1 exhibits the variables that are taken into account with correspondence to Figure 6.1. A special case of their setup is already examined in section 5.1.3 of previous chapter, where we applied the "*weak-link*" approach and the complete data set of studied 37 schools of business is also given (Table 5.8).

Table 6.1: Variables of the undergraduate business programs

External Inputs (X)	**Intermediate Measures (Z)**	**External Outputs (Y)**
Percentage of applicants rejected (X_1)	Percentage of accepted applicants enrolled (Z_1)	Percentage of students who get jobs (Y_1)
Academic rating (X_2)	Percentage of students receiving institutional scholarships (Z_2)	Percentage of students receiving internships (Z_3)
Percentage of students in top 25% of their classes (X_3)	Percentage of students receiving internships (Z_3)	

Aviles-Sacoto et al (2015) describe convincingly that the output Z_3 of the first stage should be considered both as an input to the second stage and as an output from the second stage (external output). Their modeling approach is based on the additive decomposition method of Chen et al (2009b), according to which the overall efficiency of the unit is defined as a weighted arithmetic average of the stage efficiencies. However, asserting that the conventional additive decomposition methodology is not applicable to such a peculiar situation (we will comment on this issue in the next section), they define the efficiencies of the two stages in an output-oriented VRS setting, in compliance of the modified network structure exhibited in Figure 6.2, as follows:

$$e_1 = \frac{\sum_{i=1}^{3} v_i x_{io} + u^1}{\sum_{d=1}^{2} \eta_d z_{do} + \eta_3 z_{3o}}, e_2 = \frac{\sum_{d=1}^{2} \eta_d z_{do} + u^2}{u_1 y_{1o} + g z_{3o} - h z_{3o}} \tag{6.1}$$

and the overall efficiency of the unit as a weighted arithmetic average of the stage efficiencies:

$$e_0 = w_1 e_1 + w_2 e_2 = \frac{\sum_{i=1}^{3} v_i x_{io} + u^1 + \sum_{d=1}^{2} \eta_d z_{do} + u^2}{\sum_{d=1}^{2} \eta_d z_{do} + \eta_3 z_{3o} + u_1 y_{1o} + g z_{3o} - h z_{3o}} \tag{6.2}$$

with appropriate weights:

$$\begin{aligned} w_1 &= \frac{\sum_{d=1}^{2} \eta_d z_{do} + \eta_3 z_{3o}}{\sum_{d=1}^{2} \eta_d z_{do} + \eta_3 z_{3o} + u_1 y_{1o} + g z_{3o} - h z_{3o}} \\ w_2 &= \frac{u_1 y_{1o} + g z_{3o} - h z_{3o}}{\sum_{d=1}^{2} \eta_d z_{do} + \eta_3 z_{3o} + u_1 y_{1o} + g z_{3o} - h z_{3o}} \end{aligned} \tag{6.3}$$

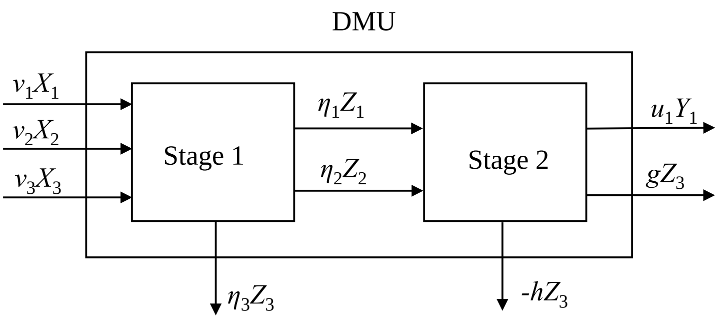


Fig. 6.2: The modified network structure

On the basis of the above definitions, the following mixed-integer linear program with binary variables is proposed to assess the overall and the stage efficiencies:

$$e_o = min \sum_{i=1}^{3} v_i x_{io} + \mu^1 + \sum_{d=1}^{2} \pi_d z_{do} + \mu^2$$

$$s.t.$$

$$\sum_{d=1}^{2} \pi_d z_{do} + \delta z_{3o} + \mu y_{1o} + \gamma z_{3o} - \beta z_{3o} = 1$$

$$\sum_{i=1}^{3} v_i x_{ij} + \mu^1 - \sum_{d=1}^{2} \pi_d z_{dj} - \delta z_{3j} \geq 0, \quad \forall j$$

$$\sum_{d=1}^{2} \pi_d z_{dj} + \mu^2 - \mu y_{1j} - \gamma z_{3j} + \beta z_{3j} \geq 0, \quad \forall j \qquad (6.4)$$

$$\delta - \gamma - \beta = 0$$

$$\gamma - Me \leq 0$$

$$\beta - Mf \leq 0$$

$$e + f = 1$$

$$v_i, \pi_d, \mu, \delta, \gamma, \beta \geq 0; e, f \ binary$$

$$\mu^1, \mu^2 \ unrestricted \ in \ sign$$

Model (6.4) is based on the output oriented VRS variant of the additive decomposition model of Chen et al (2009b). The variables $v_i, \eta_d, \eta_3, u_1, g, h, u^1, u^2$ in the above ratio forms (6.1) and (6.2) of the individual and the overall stage efficiencies are in correspondence with the variables $v_i, \pi_d, \delta, \mu, \gamma, \beta, \mu^1, \mu^2$ in the linear model (6.4), obtained after applying the transformation of Charnes and Cooper (1962).

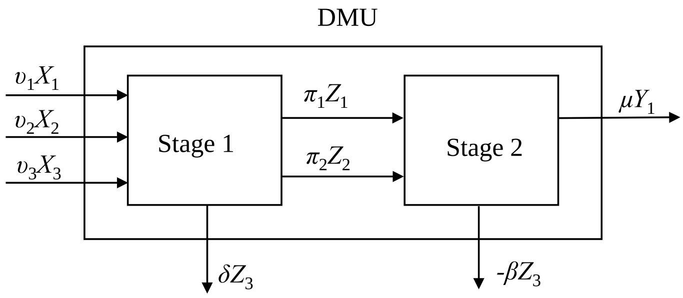
6.1 Comments

While the definition of the efficiency of stage-1 in (6.1) is obvious, in the second stage the intermediate measure Z_3 is moved from the input to the output side with a negative sign ($-hZ_3$). So, in the denominator of the efficiency ratio of stage-2, which in fact represents the total virtual output of the second stage, the Z_3 is counted twice, once with a positive and once with a negative sign. This was deemed necessary by the authors after their observation that putting Z_3 in both the numerator and the denominator of the efficiency ratio of the second stage, as in the following model (6.5), leads to erroneous results, in the sense that the second stage will be always efficient. This argument is supported by assuming a feasible solution where all the variables but g and h are set to zero and g=h. Although such a solution is optimal when the second stage is assessed independently, in a joint assessment, as that imposed by the additive decomposition method, this is not necessarily true. This is validated in the next section. Moreover, forcing the free variable u^2 to take a zero value does not allow the unit to freely exhibit increasing or decreasing returns-to-scale.

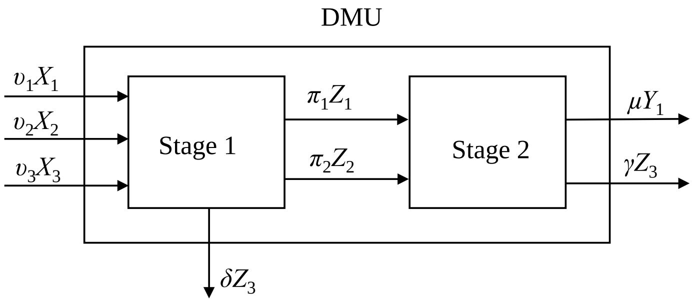
$$
\begin{aligned}
& min\ e_2 = \frac{\sum_{d=1}^{2} \eta_d z_{do} + h z_{3o} + u^2}{u_1 y_{1o} + g z_{3o}} \\
& s.t. \\
& \frac{\sum_{d=1}^{2} \eta_d z_{dj} + h z_{3j} + u^2}{u_1 y_{1j} + g z_{3j}} \geq 1 \ \ \forall j \\
& u_1, \eta_d, g, h \geq 0, u^2\ unrestricted\ in\ sign
\end{aligned}
\qquad (6.5)
$$

The last four constraints in (6.4) designate the role of intermediate measure Z_3. In particular, when in the optimal solution the binary variables take the values $e = 0, f = 1$, then $\gamma = 0, \beta = \delta$ and the intermediate measure Z_3 is considered as input to stage-2, as depicted in Figure 6.3(a). Notice again, that Z_3 is moved to the output side but with a negative sign ($-\beta Z_3$). In the case that the binary variables take the values $e = 1, f = 0$, then $\beta = 0, \gamma =$

δ and the intermediate measure Z_3 is considered as output from stage-2, as depicted in Figure 6.3(b). In every case, the weights assigned to Z_3 are the same, no matter if it is considered as input or as output.



(a) Z_3 is input to stage-2



(b) Z_3 is output from stage-2

Fig. 6.3: The network structures designated by the role of Z3

Our main concern is that the assessments made by model (6.4) differentiate and diverge from the initial argument that the intermediate measure Z_3 should be an input to the second stage *and at the same time* an output from that stage. Actually, the modelling approach followed leads to a different situation where Z_3 is *either* an input to *or* an output from the second stage. In the next section, we present the appropriate model that complies with the original situation set in Aviles-Scoto et al (2015).

6.2 Modelling the situation where some outputs of the first stage are inputs to the second stage and outputs from that stage

Assuming an output orientation in both stages, we define the efficiencies of stage-1 and stage-2 as follows:

$$e_1 = \frac{\sum_{i=1}^{3} v_i x_{io} + u^1}{\sum_{d=1}^{3} \eta_d z_{do}}, e_2 = \frac{\sum_{d=1}^{3} \eta_d z_{do} + u^2}{u_1 y_{1o} + \eta_3 z_{3o}} \tag{6.6}$$

Then the overall efficiency, as per the additive model, is

$$e_0 = w_1 e_1 + w_2 e_2 = \frac{\sum_{i=1}^{3} v_i x_{io} + u^1 + \sum_{d=1}^{3} \eta_d z_{do} + u^2}{\sum_{d=1}^{3} \eta_d z_{do} + u_1 y_{1o} + \eta_3 z_{3o}} \tag{6.7}$$

with weights:

$$\begin{aligned} w_1 &= \frac{\sum_{d=1}^{3} \eta_d z_{do}}{\sum_{d=1}^{3} \eta_d z_{do} + u_1 y_{1o} + \eta_3 z_{3o}} \\ w_2 &= \frac{u_1 y_{1o} + \eta_3 z_{3o}}{\sum_{d=1}^{3} \eta_d z_{do} + u_1 y_{1o} + \eta_3 z_{3o}} \end{aligned} \tag{6.8}$$

The following fractional program provides the overall efficiency of the evaluated DMU:

$$\begin{aligned} & e_0 = min \frac{\sum_{i=1}^{3} v_i x_{io} + u^1 + \sum_{d=1}^{3} \eta_d z_{do} + u^2}{\sum_{d=1}^{3} \eta_d z_{do} + u_1 y_{1o} + \eta_3 z_{3o}} \\ & s.t. \\ & \frac{\sum_{i=1}^{3} v_i x_{ij} + u^1}{\sum_{d=1}^{3} \eta_d z_{dj}} \geq 1 \ \ \forall j \\ & \frac{\sum_{d=1}^{3} \eta_d z_{dj} + u^2}{u_1 y_{1j} + \eta_3 z_{3j}} \geq 1 \ \ \forall j \\ & u_1, \eta_d, v_i \geq 0, u^1, u^2 \ unrestricted \ in \ sign \end{aligned} \tag{6.9}$$

The linear model below, as derived by the Charnes-Cooper transformation, estimates the overall efficiency of the DMU in compliance with the original network structure depicted in Figure 6.1.

$$
\begin{aligned}
& e_0 = min \sum_{i=1}^{3} \upsilon_i x_{io} + \mu^1 + \sum_{d=1}^{3} \pi_d z_{do} + \mu^2 \\
& s.t. \\
& \sum_{d=1}^{3} \pi_d z_{do} + \mu y_{1o} + \pi_3 z_{3o} = 1 \\
& \sum_{i=1}^{3} \upsilon_i x_{ij} + \mu^1 - \sum_{d=1}^{3} \pi_d z_{dj} \geq 0, \quad \forall j \\
& \sum_{d=1}^{3} \pi_d z_{dj} + \mu^2 - \mu y_{1j} - \pi_3 z_{3j} \geq 0, \quad \forall j \\
& \upsilon_i, \pi_d, \mu \geq 0;\ \mu^1, \mu^2\ unrestricted\ in\ sign
\end{aligned}
\tag{6.10}
$$

Table 6.2 exhibits the results obtained from model (6.10). Columns 2-4 show the overall efficiency and the weights w_1 and w_2 that were assumed, whereas columns 5-6 show the efficiencies of the two stages. Notice here that applying the leader/follower notion as introduced in Chen et al (2009b), as a means to estimate extreme values for the stage efficiencies, we have got identical efficiency scores for the two stages, which means that the decomposition of the overall efficiency to the stage efficiencies is unique. Notice also that, as advised in the original paper of Aviles-Sacoto et al (2015), we have carried out the analysis by assuming that each one of the weights w_1 and w_2, which are used to aggregate the stage efficiencies, will take at least a value of 0.1.

Table 6.2: Results obtained from model (6.10)

DMU	e_0	w_1	w_2	e_1	e_2
1	105.89	0.10	0.90	108.04	105.65
2	105.29	0.50	0.50	110.58	100.00
3	102.68	0.47	0.53	100.00	105.03
4	105.29	0.47	0.53	106.38	104.31
5	105.31	0.47	0.53	106.04	104.65
6	105.20	0.50	0.50	110.41	100.00
7	100.17	0.50	0.50	100.34	100.00
8	106.12	0.50	0.50	112.23	100.00
9	100.00	0.10	0.90	100.00	100.00
10	100.00	0.50	0.50	100.00	100.00
11	104.24	0.10	0.90	108.37	103.78
12	101.66	0.10	0.90	116.62	100.00
13	100.00	0.50	0.50	100.00	100.00
14	104.03	0.50	0.50	108.06	100.00
15	107.67	0.10	0.90	107.94	107.64
16	100.00	0.50	0.50	100.00	100.00
17	101.75	0.50	0.50	103.49	100.00
18	101.23	0.50	0.50	102.46	100.00
19	100.00	0.50	0.50	100.00	100.00
20	106.59	0.10	0.90	129.38	104.06
21	100.00	0.50	0.50	100.00	100.00
22	105.79	0.10	0.90	117.28	104.51
23	104.60	0.50	0.50	109.20	100.00
24	100.00	0.50	0.50	100.00	100.00
25	130.22	0.62	0.38	126.30	136.57
26	100.00	0.50	0.50	100.00	100.00
27	103.38	0.50	0.50	106.77	100.00
28	113.53	0.50	0.50	127.06	100.00
29	100.00	0.50	0.50	100.00	100.00
30	111.72	0.50	0.50	123.44	100.00
31	100.00	0.50	0.50	100.00	100.00
32	108.35	0.50	0.50	116.70	100.00
33	100.00	0.50	0.50	100.00	100.00
34	100.00	0.50	0.50	100.00	100.00
35	120.67	0.63	0.37	100.00	155.50
36	113.63	0.50	0.50	127.25	100.00
37	100.00	0.50	0.50	100.00	100.00

In the next section we provide a model to assess the efficiency of the units under the assumption that the intermediate measure Z3 is either input to the second stage or output from that stage.

6.3 Modelling the situation where some outputs of the first stage are either inputs to the second stage or outputs from that stage

In a manner analogous to that introduced in Aviles-Sacoto et al (2015), we define the stage and the overall efficiencies, on the basis of the modified network structure depicted in Figure 6.4, as follows:

$$e_1 = \frac{\sum_{i=1}^{3} v_i x_{io} + u^1}{\sum_{d=1}^{2} \eta_d z_{do} + \eta_3 z_{3o} + h z_{3o}}, e_2 = \frac{\sum_{d=1}^{2} \eta_d z_{do} + h z_{3o} + u^2}{u_1 y_{1o} + g z_{3o}} \tag{6.11}$$

and the overall efficiency of the unit as a weighted average of the above stage efficiencies:

$$e_0 = w_1 e_1 + w_2 e_2 = \frac{\sum_{i=1}^{3} v_i x_{io} + u^1 + \sum_{d=1}^{2} \eta_d z_{do} + h z_{3o} + u^2}{\sum_{d=1}^{2} \eta_d z_{do} + \eta_3 z_{3o} + h z_{3o} + u_1 y_{1o} + g z_{3o}} \tag{6.12}$$

where the weights are selected appropriately as follows (cf. Chen et al, 2009b):

$$w_1 = \frac{\sum_{d=1}^{2} \eta_d z_{do} + \eta_3 z_{3o} + h z_{3o}}{\sum_{d=1}^{2} \eta_d z_{do} + \eta_3 z_{3o} + h z_{3o} + u_1 y_{1o} + g z_{3o}}$$

$$w_2 = \frac{u_1 y_{1o} + g z_{3o}}{\sum_{d=1}^{2} \eta_d z_{do} + \eta_3 z_{3o} + h z_{3o} + u_1 y_{1o} + g z_{3o}} \tag{6.13}$$

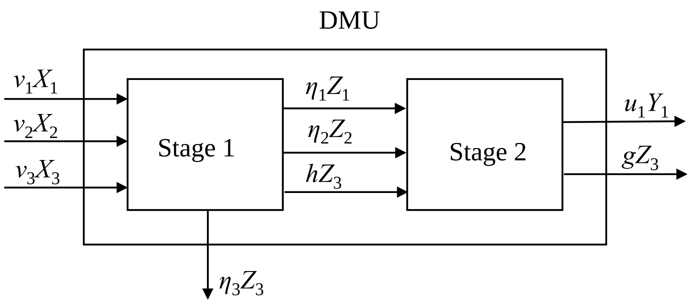


Fig. 6.4: An alternative modified network structure

According to the above definitions and applying the Charnes-Cooper transformation we propose the following linear program with binary variables to assess the overall and the stage efficiencies:

$$e_o = min \sum_{i=1}^{3} v_i x_{io} + \mu^1 + \sum_{d=1}^{2} \pi_d z_{do} + \beta z_{3o} + \mu^2$$

$$s.t.$$

$$\sum_{d=1}^{2} \pi_d z_{do} + \delta z_{3o} + \mu y_{1o} + \gamma z_{3o} + \beta z_{3o} = 1$$

$$\sum_{i=1}^{3} v_i x_{ij} + \mu^1 - \sum_{d=1}^{2} \pi_d z_{dj} - \delta z_{3j} - \beta z_{3o} \geq 0, \quad \forall j$$

$$\sum_{d=1}^{2} \pi_d z_{dj} + \mu^2 - \mu y_{1j} - \gamma z_{3j} + \beta z_{3j} \geq 0, \quad \forall j \tag{6.14}$$

$$\delta - \gamma = 0$$

$$\gamma - Me \leq 0$$

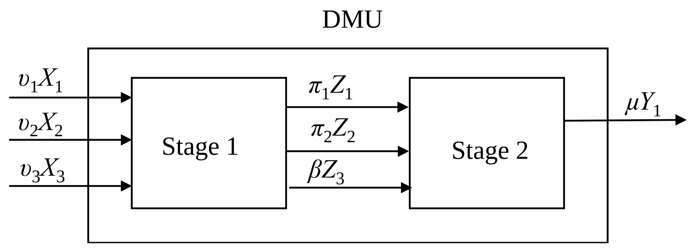
$$\beta - Mf \leq 0$$

$$e + f = 1$$

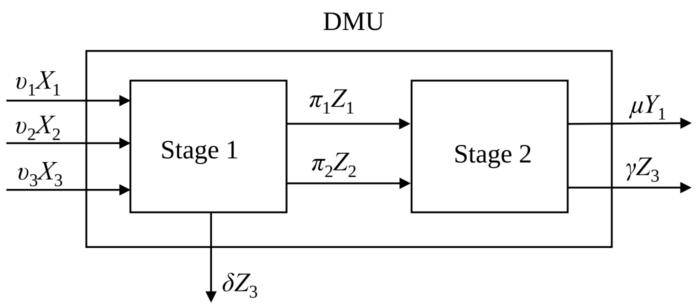
$$v_i, \pi_d, \mu, \delta, \gamma, \beta \geq 0; e, f \; binary$$

$$\mu^1, \mu^2 \; unrestricted \; in \; sign$$

Analogously to (6.4), the last four constraints in (6.14) designate the role of intermediate measure Z_3. In particular, when in the optimal solution the binary variables take the values $e = 0, f = 1$, then $\gamma = \delta = 0$ and the intermediate measure Z_3 is considered as input to stage-2, as depicted in Figure 6.5(a). When the binary variables take the values $e = 1, f = 0$, then $\beta = 0, \gamma = \delta$ and the intermediate measure Z_3 is considered as output from stage-2, as depicted in Figure 6.5(b). Notice that the latter network structure is identical to that assumed in Aviles-Sacoto et al (2015). The differentiation is in the structure of Figure 6.5(a), where Z_3 is conventionally treated, similarly to the other intermediate measures.



(a) Z_3 is input to stage-2



(b) Z_3 is output from stage-2

Fig. 6.5: The two network structures designated by the role of Z3

If it is to avoid solving LPs with binary variables, one could equivalently apply the conventional additive decomposition models with respect to the network structures in Figure 6.5 (a) and (b) and then choose the minimum of the two estimated overall efficiencies as the final result. Table 6.3 exhibits the results obtained from model (6.14). Columns 2-4 show the overall efficiency and the weights assumed, column 5 shows the input/output characterization of the intermediate measure Z_3 as decided by the binary variables e and f, whereas columns 6-7 show the efficiencies of the two stages. Applying the leader/follower notion, we have got identical efficiency scores for the two stages, which means that the decomposition of the overall efficiency to the stage efficiencies is unique.

Table 6.3: Results obtained from model (6.14)

DMU	e_0	w_1	w_2	*Output/Input*	e_1	e_2
1	106.14	0.10	0.90	Output	107.31	106.01
2	110.51	0.40	0.60	Output	106.56	113.09
3	101.19	0.29	0.71	Output	100.85	101.32
4	105.16	0.28	0.72	Output	105.67	104.96
5	105.09	0.29	0.71	Output	104.89	105.17
6	108.26	0.39	0.61	Output	103.18	111.57
7	104.24	0.50	0.50	Output	100.34	108.14
8	110.59	0.37	0.63	Output	114.97	107.99
9	100.00	0.26	0.74	Output	100.00	100.00
10	100.00	0.29	0.71	Output	100.00	100.00
11	104.11	0.10	0.90	Input	105.68	103.93
12	101.42	0.10	0.90	Input	114.24	100.00
13	106.98	0.90	0.10	Input	100.00	169.79
14	111.00	0.50	0.50	Output	108.22	113.83
15	107.92	0.10	0.90	Output	107.74	107.93
16	106.84	0.57	0.43	Input	100.00	115.89
17	105.69	0.39	0.61	Output	103.49	107.12
18	104.14	0.28	0.72	Output	103.07	104.55
19	100.00	0.51	0.49	Output	100.00	100.00
20	104.61	0.10	0.90	Input	135.17	101.22
21	108.25	0.56	0.44	Output	100.00	118.76
22	105.51	0.10	0.90	Input	120.50	103.85
23	114.78	0.39	0.61	Output	100.00	124.30
24	100.00	0.82	0.18	Output	100.00	100.00
25	137.46	0.66	0.34	Input	119.34	173.39
26	106.14	0.58	0.42	Output	100.00	114.70
27	108.51	0.39	0.61	Input	117.32	102.81
28	118.55	0.51	0.49	Input	124.95	112.01
29	105.83	0.90	0.10	Input	100.00	158.32
30	128.37	0.68	0.32	Input	104.15	178.87
31	101.44	0.55	0.45	Input	100.00	103.19
32	112.41	0.60	0.40	Input	100.00	131.26
33	105.75	0.50	0.50	Output	100.00	111.49
34	105.36	0.50	0.50	Output	100.00	110.72
35	116.58	0.61	0.39	Input	104.46	135.87
36	116.15	0.56	0.44	Input	123.35	107.08
37	100.00	0.90	0.10	Input	100.00	100.00

Conclusion

We presented in this chapter an alternative two-stage network DEA approach to the assessment of the efficiencies of undergraduate programs in a peculiar situation where one of the intermediate measures must be considered as input to the second stage and, at the same time, as output of the second stage. Our contribution to this issue is motivated by an observation we made that the original modeling approach followed in Aviles-Sacoto (2015) arbitrarily, yet unnecessarily, deviates from that setting and designates a different situation where the specific intermediate measure is viewed either as input to or as output from the second stage of the process.

Chapter 7

Conclusion

The field of study of this thesis is the network Data Envelopment Analysis, which is an important extension of the Data Envelopment Analysis. We conducted a critical survey of the network DEA literature as well as we introduced novel network DEA methods for the performance assessment of the DMUs that consist of several sub-processes.

Initially, we emphasized the advantages of the network DEA over the conventional DEA and the new insights and possibilities that offers the former in the area of performance measurement. When the internal structure of the DMU is known and the interrelations among its sub-processes can be accurately depicted, then it is strongly recommended to avoid the traditional perception of standard DEA that regards the DMU as a “black box”. Instead, it is proposed that network DEA methods should be employed for the performance evaluation. However, cautions should be taken because as we discussed and proved, there are some deficiencies in the recent developments of network DEA. Hence, we carried out a thorough categorization and critical survey of the state-of-the art network DEA methods, we unveiled their relations and differences, we uncovered their defects and we revealed the effects of these shortcomings in the efficiency assessments. We classified a great volume of network DEA studies based on the assessment approach they follow. In particular, we defined two assessment paradigms, the *independent* and the *joint*. In the *independent assessment paradigm* the standard DEA models are employed to assess the performance of the DMUs and the sub-processes separately. On the contrary, in the *joint assessment paradigm* the DMU and its sub-processes are *jointly* evaluated. We specified three approaches as representatives of the *joint assessment paradigm*, namely the efficiency decomposition approach, the slacks-based measure approach and the system-centric approach. The categorization of the approaches was based on the way they conceptualize the relationship between the system (DMU) and the stage efficiencies as well as on the kind of information that they provide for the performance of the individual stages and the system. We revealed the drawbacks of the existing network DEA methods concerning the returns to scale, the inconsistency between their multiplier and envelopment models, the non-unique efficiency scores and the inability to be universally applied on every type of network structure. Also, we proved that the additive efficiency decomposition method provides biased efficiency assessments and we established the properties that the network DEA methods should meet.

Then, we introduced the *composition approach* to two-stage network DEA, as opposed to the efficiency decomposition approach. Our novel approach overwhelms the shortcomings spotted for the additive and the multiplicative decomposition methods, i.e. it provides unique and unbiased efficiency scores. Contrary to the decomposition approach, in composition approach we first estimate the stage efficiencies and then we aggregate them either additively or multiplicatively to obtain the overall efficiency. In the frame of our composition approach, the efficiency assessment of the two-stage process is formulated as a multi-objective mathematical programming problem. In particular, we formulated a bi-objective mathematical program by assuming an output orientation for the first stage and an input orientation for the second stage, where the intermediate measures were used as the basis to link the efficiency assessments of the two stages. We employed two scalarization techniques so as to convert the bi-objective problem to a single objective LP. Firstly, based on the L_1 norm we aggregated the two objective functions of the bi-objective program additively, without giving any priority between them; the application of this scalarizing function yields an extreme (vertex) Pareto-optimal solution. Then, we employed a min-max scalarization technique, i.e. the Tchebycheff norm (L_∞), which provides a point on the Pareto front not necessarily extreme. Also, we developed two methods to derive the efficient frontier in two-stage DEA and provide efficient projections. The first naturally stems from our composition approach, while the second seeks to provide efficient projections by causing the less change on the original levels of the intermediate measures.

Next, we built upon the composition approach and we introduced the "*weak-link*" approach to two-stage network DEA, which inherits the nice properties of the former, i.e. provides unique and unbiased efficiency scores. Also, the "weak-link" approach can be readily applied to various types of two-stage network structures. In this approach, we introduced a novel definition about the overall efficiency of the DMU, inspired by the "weak link" notion in supply chains and the maximum-flow/minimum-cut problem in networks. We incorporated this notion into the assessment by assuming that given the stage efficiencies, the system efficiency can be viewed as the maximum flow through the network and can be estimated as the min-cut of the network, i.e. the system efficiency derives as the lowest of the stage efficiencies. We mathematically represented this concept by employing a two-phase max-min optimization method in a multi-objective programming framework, which seeks to maximize the minimum weighted achievement from zero-level efficiency, i.e. maximizing the lowest of the stage efficiencies (weak link). The proposed two-phase procedure estimates the stage efficiencies and the overall efficiency simultaneously by providing a unique Pareto optimal solution. The search direction towards the Pareto front is driven by the assumption that the stage efficiencies are proportional to their independent counterparts. External priorities can be also introduced explicitly to our methodology so as to obtain alternative Pareto optimal solutions. We conducted a systematic investigation of the sensitivity of the weak link so as to identify the source of inefficiency in the two-stage processes. A thorough comparison with the multiplicative decomposition method illustrates the advantages of the "weak-link" approach.

Finally, we revisited the work of Aviles-Sacoto et al (2015) who evaluated a peculiar situation of 37 undergraduate business programs in U.S. as two-stage processes, where some of the intermediate measures are inputs to the second stage and at the same time external outputs from that stage. We revealed that their modelling approach departs from the described setting and adapts a different situation, where the specific intermediate measure is viewed either as input to or as output from the second stage of the process. We alternatively proposed a different modelling approach for the performance assessment of the specific two-stage process under examination.

Closing this thesis, we remark that out methods can be straightforwardly applied to real word problems. For instance, the natural representation of the supply chain operations as a multi-stage process is indicative of the synergy of supply chain management with network DEA, as they benefit mutually from the development of methodological tools for performance measurement. A subject for future research is the extension of the *composition* and the "*weak-link*" approaches to general network structures involving series and parallel processes. A universal network DEA method that could be applied to every type of network structure would be advantageous. Also, future studies could be focused on the revision of the system-centric methods, so as to yield the stage efficiency scores except from the overall efficiency. Moreover, a topic that should be revisited is the returns to scale in network DEA models. Another issue that is worth investigating in network DEA is the perfect mapping between the multiplier and the envelopment models of each approach, in order to provide efficient projections. Finally, it is anticipated that the conclusions drawn from this thesis will assist the analyst to be accustomed with the network DEA, will be inspirational for exploring new ideas and will serve to advance and disseminate both the theoretical and the problem-driven research. Undoubtedly, further development of the network DEA methods is needed so as to widen the application field, to aid the decision makers to address the increasing complexity of the organizations and improve their performance and to extend the frontiers of research in DEA.

Implementation comment: All models presented in this thesis are developed and tested in MATLAB (MATLAB 8.0, The MathWorks Inc., Natick, MA, US, 2012) in combination with an open source mixed integer linear programming (MILP) solver (lp_solve 5.5.2.5).

Appendix

The case of a two-stage process with a single intermediate measure

As noticed in Chapter 4 (see Theorems 4.5 and 4.6), in a two-stage production process of Type I (Fig. 4.1) with a single intermediate measure, the stage efficiency scores derived by our *composition* approach are identical to the independent efficiency scores. We provide an illustrative example by applying our approach, under CRS and VRS assumption, to the data set used by Chen and Zhu (2004) for measuring information technology's indirect impact on firm performance in the banking industry in the years 1987-1989. The original data set studied in Wang et al (1997) consists of 36 observations with negative profits, however, as long as the current study have different goals from Wang et al (1997) we follow Chen and Zhu (2004) who chose to remove the observations with negative profits. Twenty seven firms in the banking industry use, in the first stage, as inputs the fixed assets (X1), the number of employees (X2) and the IT investment (X3) to generate the single intermediate measure Deposits (Z1). In the second stage the single intermediate measure Deposits (Z1) is converted into the profit (Y1) and the fraction of loans recovered (Y2). The data set is presented in Table A.1.

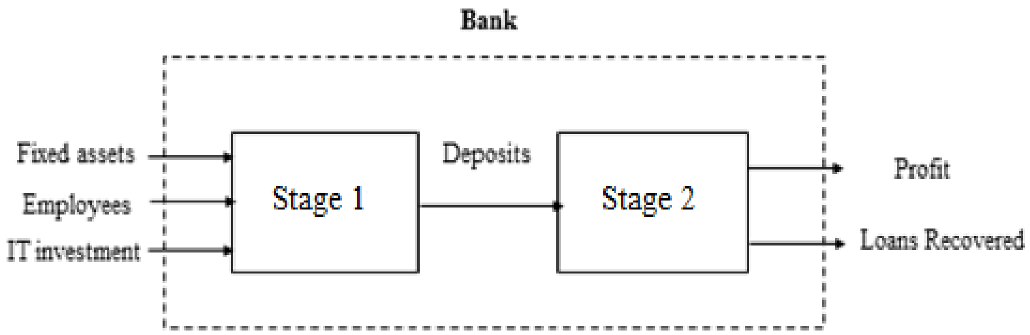


Fig. A.1: Bank operations as a two-stage process

Table A.1: IT data set (source: Chen and Zhu, 2004)

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

Table A.2 exhibits the CRS and VRS efficiency scores that derived from our composition approach presented in Chapter 4. Table A.2 exhibits the CRS and VRS efficiency scores that derived from our composition approach presented in Chapter 4. As it was expected, the efficiency scores are identical to their independent counterparts under both assumptions. In addition, we remark that under CRS assumption, the stage efficiency scores obtained from the additive and the multiplicative approaches are equal to the ones derived from our approach; also the same observation holds for the additive approach under VRS situation.

Table A.2: CRS and VRS efficiencies from models (4.13), (4.16), (4.25) and (4.26)

DMU	$\hat{e}^1_{CRS}$	$\hat{e}^2_{CRS}$	$\hat{e}^o=(\hat{e}^1+\hat{e}^2)/2$	$\hat{e}^1_{VRS}$	$\hat{e}^2_{VRS}$	$\hat{e}^o=(\hat{e}^1+\hat{e}^2)/2$
1	0.6388	0.7459	0.6923	0.6776	0.8458	0.7617
2	0.6507	0.7819	0.7163	0.6678	0.9960	0.8319
3	0.5179	0.7730	0.6454	0.5357	1	0.7679
4	0.5986	0.7142	0.6564	1	0.7144	0.8572
5	0.5556	0.7236	0.6396	0.7079	0.7546	0.7313
6	0.7599	0.5758	0.6679	0.9680	0.6779	0.8229
7	1	0.5758	0.7879	1	1	1
8	0.5352	0.8250	0.6801	0.5400	0.8304	0.6852
9	0.6249	0.6347	0.6298	1	1	1
10	0.4961	0.7188	0.6075	0.5029	0.7680	0.6354
11	0.4945	0.7188	0.6067	0.5012	0.7680	0.6346
12	0.6685	0.5949	0.6317	0.7329	0.5950	0.6639
13	0.9487	0.8582	0.9034	1	0.8589	0.9295
14	0.5880	0.5782	0.5831	0.6967	0.7739	0.7353
15	0.6582	0.6034	0.6308	0.7782	0.8740	0.8261
16	0.6646	0.6434	0.6540	0.6681	1	0.8340
17	0.7177	0.7877	0.7527	1	0.7933	0.8967
18	1	1	1	1	1	1
19	0.8144	0.5926	0.7035	0.8188	0.6544	0.7366
20	0.6933	1	0.8467	1	1	1
21	0.7067	0.9935	0.8501	1	0.9935	0.9968
22	0.7942	0.6408	0.7175	1	0.6410	0.8205
23	0.7802	0.6993	0.7397	0.9600	0.7328	0.8464
24	0.9300	0.7135	0.8218	0.9629	0.9915	0.9772

Estimation of the extreme Pareto points

In the following, we provide the models that yield the independent (ideal) stage efficiencies as well as the ones that produce the extreme boundary points of the feasible set in the objective functions space, for the two-stage processes of Type II, III and IV.

Two-stage process with extra inputs in the stage-2 (Type II)

Concerning the two-stage process with extra inputs to the second stage of Type II (Fig. 5.5), the input-oriented CRS-DEA models that independently estimate the stage-1 and the stage-2 efficiencies for the evaluated unit j_0 are as follows:

$$E^1_{j_0} = max \frac{\varphi Z_{j_0}}{\eta X_{j_0}}$$
$$s.t.$$
$$\varphi Z_j - \eta X_j \leq 0,\ \ j = 1, \dots, n \qquad \text{(A.1)}$$
$$\eta \geq \varepsilon, \varphi \geq \varepsilon$$

$$E^2_{j_0} = max \frac{\omega Y_{j_0}}{\varphi Z_{j_0} + gL_{j_0}}$$
$$s.t.$$
$$\omega Y_j - \varphi Z_j - gL_j \leq 0,\ \ j = 1, \dots, n \qquad \text{(A.2)}$$
$$\varphi \geq \varepsilon, \omega \geq \varepsilon, g \geq \varepsilon$$

The following augmented models derive from models (A.1) and (A.2) by appending the constraints of model (A.1) to model (A.2) and vice versa without affecting their optimal efficiency scores, see Theorems 4.1 and 4.2.

$$E^1_{j_0} = max \frac{\varphi Z_{j_0}}{\eta X_{j_0}}$$
$$s.t.$$
$$\varphi Z_j - \eta X_j \leq 0,\ \ j = 1, \dots, n \qquad \text{(A.3)}$$
$$\omega Y_j - \varphi Z_j - gL_j \leq 0,\ \ j = 1, \dots, n$$
$$\eta \geq \varepsilon, \varphi \geq \varepsilon, \omega \geq \varepsilon, g \geq \varepsilon$$

$$E^2_{j_0} = max \frac{\omega Y_{j_0}}{\varphi Z_{j_0} + gL_{j_0}}$$
$$s.t.$$
$$\varphi Z_j - \eta X_j \leq 0,\ \ j = 1, \dots, n \qquad \text{(A.4)}$$
$$\omega Y_j - \varphi Z_j - gL_j \leq 0,\ \ j = 1, \dots, n$$
$$\eta \geq \varepsilon, \varphi \geq \varepsilon, \omega \geq \varepsilon, g \geq \varepsilon$$

Given the ideal point $\left(E^1_{j_0}, E^2_{j_0}\right)$ defined by the independent stage efficiency scores of the evaluated unit j_0, we derive the extreme points $A\left(N^1_{j_0}, E^2_{j_0}\right)$ and $B\left(E^1_{j_0}, N^2_{j_0}\right)$ on the upper-right boundary of the feasible set in the objective functions space of (5.15) as follows:

For the point $A\left(N^1_{j_0}, E^2_{j_0}\right)$, get $N^1_{j_0}$ as the optimal value of the objective function in the following linear program:

$$N^1_{j_0} = max\ \ wZ_{j_0}$$
$$s.t.$$
$$vX_{j_0} = 1$$
$$\frac{uY_{j_0}}{wZ_{j_0} + \gamma L_{j_0}} \geq E^2_{j_0} \qquad \text{(A.5)}$$
$$wZ_j - vX_j \leq 0,\ \ j = 1, \dots, n$$
$$uY_j - wZ_j - \gamma L_j \leq 0,\ \ j = 1, \dots, n$$
$$v \geq \varepsilon, w \geq \varepsilon, u \geq \varepsilon, \gamma \geq \varepsilon$$

For the point $B\left(E_{j_0}^1, N_{j_0}^2\right)$, get $N_{j_0}^2$ as the optimal value of the objective function in the following linear program:

$$
\begin{aligned}
& N_{j_0}^2 = max\ uY_{j_0} \\
& s.t. \\
& wZ_{j_0} + \gamma L_{j_0} = 1 \\
& \frac{wZ_{j_0}}{vX_{j_0}} \geq E_{j_0}^1 \\
& wZ_j - vX_j \leq 0,\ \ j = 1, \dots, n \\
& uY_j - wZ_j - \gamma L_j \leq 0,\ \ j = 1, \dots, n \\
& v \geq \varepsilon, w \geq \varepsilon, u \geq \varepsilon, \gamma \geq \varepsilon
\end{aligned}
\tag{A.6}
$$

Two-stage process with extra outputs from stage-1 (Type III)

Concerning the two-stage process with extra outputs flowing out from the first stage of Type III (Fig. 5.10), we obtain the independent (ideal) efficiency scores using the input-oriented CRS-DEA models for the evaluated unit j_0 as follows:

$$
\begin{aligned}
& E_{j_0}^1 = max \frac{\varphi Z_{j_0} + hK_{j_0}}{\eta X_{j_0}} \\
& s.t. \\
& \varphi Z_j + hK_j - \eta X_j \leq 0,\ \ j = 1, \dots, n \\
& \eta \geq \varepsilon, \varphi \geq \varepsilon, h \geq \varepsilon
\end{aligned}
\tag{A.7}
$$

$$
\begin{aligned}
& E_{j_0}^2 = max \frac{\omega Y_{j_0}}{\varphi Z_{j_0}} \\
& s.t. \\
& \omega Y_j - \varphi Z_j \leq 0,\ \ j = 1, \dots, n \\
& \varphi \geq \varepsilon, \omega \geq \varepsilon
\end{aligned}
\tag{A.8}
$$

By appending the constraints of one model to the other, as already described, we derive the following augmented models:

$$
\begin{aligned}
& E_{j_0}^1 = max \frac{\varphi Z_{j_0} + hK_{j_0}}{\eta X_{j_0}} \\
& s.t. \\
& \varphi Z_j + hK_j - \eta X_j \leq 0,\ j = 1, \dots, n \\
& \omega Y_j - \varphi Z_j \leq 0,\ \ j = 1, \dots, n \\
& \eta \geq \varepsilon, \varphi \geq \varepsilon, \omega \geq \varepsilon, h \geq \varepsilon
\end{aligned}
\tag{A.9}
$$

$$
\begin{aligned}
& E_{j_0}^2 = max \frac{\omega Y_{j_0}}{\varphi Z_{j_0}} \\
& s.t. \\
& \varphi Z_j + hK_j - \eta X_j \leq 0,\ j = 1, \dots, n \\
& \omega Y_j - \varphi Z_j \leq 0,\ \ j = 1, \dots, n \\
& \eta \geq \varepsilon, \varphi \geq \varepsilon, \omega \geq \varepsilon, h \geq \varepsilon
\end{aligned}
\tag{A.10}
$$

Given the ideal point $\left(E_{j_0}^1, E_{j_0}^2\right)$ obtained from the above models for the evaluated unit j_0, we obtain the extreme boundary points $A\left(N_{j_0}^1, E_{j_0}^2\right)$ and $B\left(E_{j_0}^1, N_{j_0}^2\right)$ on the upper-right boundary of the feasible set in the objective functions space of (5.22) as follows:

For the point $A\left(N_{j_0}^1, E_{j_0}^2\right)$, get $N_{j_0}^1$ as the optimal value of the objective function in the following linear program:

$$
\begin{aligned}
& N_{j_0}^1 = max \;\; wZ_{j_0} + \mu K_{j_0} \\
& s.t. \\
& vX_{j_0} = 1 \\
& \frac{uY_{j_0}}{wZ_{j_0}} \geq E_{j_0}^2 \\
& wZ_j + \mu K_j - vX_j \leq 0, \;\; j = 1, \dots, n \\
& uY_j - wZ_j \leq 0, \;\; j = 1, \dots, n \\
& v \geq \varepsilon, w \geq \varepsilon, u \geq \varepsilon, \mu \geq \varepsilon
\end{aligned}
\qquad \text{(A.11)}
$$

For the point $B\left(E_{j_0}^1, N_{j_0}^2\right)$, get $N_{j_0}^2$ as the optimal value of the objective function in the following linear program:

$$
\begin{aligned}
& N_{j_0}^2 = max \;\; uY_{j_0} \\
& s.t. \\
& wZ_{j_0} = 1 \\
& \frac{wZ_{j_0} + \mu K_{j_0}}{vX_{j_0}} \geq E_{j_0}^1 \\
& wZ_j + \mu K_j - vX_j \leq 0, \;\; j = 1, \dots, n \\
& uY_j - wZ_j \leq 0, \;\; j = 1, \dots, n \\
& v \geq \varepsilon, w \geq \varepsilon, u \geq \varepsilon, \mu \geq \varepsilon
\end{aligned}
\qquad \text{(A.12)}
$$

General two-stage process (Type IV)

As regards the general two-stage process of Type IV (Fig. 5.14), the input-oriented CRS-DEA models that estimate the stage-1 and the stage-2 efficiencies for the evaluated unit j_0 independently are as follows:

$$
\begin{aligned}
&E_{j_0}^1 = max \frac{\varphi Z_{j_0} + hK_{j_0}}{\eta X_{j_0}} \\
&s.t. \\
&\varphi Z_j + hK_j - \eta X_j \leq 0, j = 1, \dots, n \\
&\eta \geq \varepsilon, \varphi \geq \varepsilon, h \geq \varepsilon
\end{aligned}
\tag{A.13}
$$

$$
\begin{aligned}
&E_{j_0}^2 = max \frac{\omega Y_{j_0}}{\varphi Z_{j_0} + gL_{j_0}} \\
&s.t. \\
&\omega Y_j - \varphi Z_j - gL_j \leq 0, j = 1, \dots, n \\
&\varphi \geq \varepsilon, \omega \geq \varepsilon, g \geq \varepsilon
\end{aligned}
\tag{A.14}
$$

By appending the constraints of one model to the other we derive the following augmented models:

$$
\begin{aligned}
&E_{j_0}^1 = max \frac{\varphi Z_{j_0} + hK_{j_0}}{\eta X_{j_0}} \\
&s.t. \\
&\varphi Z_j + hK_j - \eta X_j \leq 0,\ j = 1, \dots, n \\
&\omega Y_j - \varphi Z_j - gL_j \leq 0,\ j = 1, \dots, n \\
&\eta \geq \varepsilon, \varphi \geq \varepsilon, \omega \geq \varepsilon, h \geq \varepsilon, g \geq \varepsilon
\end{aligned}
\tag{A.15}
$$

$$
\begin{aligned}
&E_{j_0}^2 = max \frac{\omega Y_{j_0}}{\varphi Z_{j_0} + gL_{j_0}} \\
&s.t. \\
&\varphi Z_j + hK_j - \eta X_j \leq 0,\ j = 1, \dots, n \\
&\omega Y_j - \varphi Z_j - gL_j \leq 0,\ j = 1, \dots, n \\
&\eta \geq \varepsilon, \varphi \geq \varepsilon, \omega \geq \varepsilon, h \geq \varepsilon, g \geq \varepsilon
\end{aligned}
\tag{A.16}
$$

Given the ideal point $\left(E_{j_0}^1, E_{j_0}^2\right)$ obtained from the above models for the evaluated unit j_0, we obtain the extreme boundary points $A\left(N_{j_0}^1, E_{j_0}^2\right)$ and $B\left(E_{j_0}^1, N_{j_0}^2\right)$ on the upper-right boundary of the feasible set in the objective functions space of (5.29) as follows:

For the point $A\left(N_{j_0}^1, E_{j_0}^2\right)$, get $N_{j_0}^1$ as the optimal value of the objective function in the following linear program:

$$
\begin{aligned}
& N_{j_0}^1 = max \;\; wZ_{j_0} + \mu K_{j_0} \\
& s.t. \\
& vX_{j_0} = 1 \\
& \frac{uY_{j_0}}{wZ_{j_0} + \gamma L_{j_0}} \geq E_{j_0}^2 \\
& wZ_j + \mu K_j - vX_j \leq 0, \;\; j = 1, \dots, n \\
& uY_j - wZ_j - \gamma L_j \leq 0, \;\; j = 1, \dots, n \\
& v \geq \varepsilon, w \geq \varepsilon, u \geq \varepsilon, \mu \geq \varepsilon, \gamma \geq \varepsilon
\end{aligned}
\tag{A.17}
$$

For the point $B\left(E_{j_0}^1, N_{j_0}^2\right)$, get $N_{j_0}^2$ as the optimal value of the objective function in the following linear program:

$$
\begin{aligned}
& N_{j_0}^2 = max \;\; uY_{j_0} \\
& s.t. \\
& wZ_{j_0} + \gamma L_{j_0} = 1 \\
& \frac{wZ_{j_0} + \mu K_{j_0}}{vX_{j_0}} \geq E_{j_0}^1 \\
& wZ_j + \mu K_j - vX_j \leq 0, \;\; j = 1, \dots, n \\
& uY_j - wZ_j - \gamma L_j \leq 0, \;\; j = 1, \dots, n \\
& v \geq \varepsilon, w \geq \varepsilon, u \geq \varepsilon, \mu \geq \varepsilon, \gamma \geq \varepsilon
\end{aligned}
\tag{A.18}
$$

References

1. Abad C, Thoreb SA, Laffarga J (2004). Fundamental Analysis of Stocks by Two-stage DEA. *Managerial and Decision Economics*, 25, 231-241.
2. Adler N, Liebert V, Yazhemsky E (2013). Benchmarking airports from a managerial perspective. *Omega*, 41, 442–458.
3. Agrell PJ, Hatami-Marbini A (2013). Frontier-based performance analysis models for supply chain management: State of the art and research directions. *Computers and Industrial Engineering*, 66(3), 567-583.
4. Akther S, Fukuyama H, Weber WL (2013). Estimating two-stage network slacks-based inefficiency: An application to Bangladesh banking. *Omega*, 41, 88-96.
5. Amirteimoori A, Kordrostami S (2005). DEA-like models for multi-component performance measurement. *Applied Mathematics and Computation*, 163, 735–743.
6. Amirteimoori A, Shafiei M (2006). Measuring the efficiency of interdependent decision making sub-units in DEA. *Applied Mathematics and Computation*, 173, 847–855.
7. Amirteimoori A (2013). A DEA two-stage decision processes with shared resources. *Central European Journal of Operations Research*, 21, 141–151.
8. An Q, Yan H, Wu J, Liang L (2016). Internal resource waste and centralization degree in two-stage systems: An efficiency analysis. *Omega*, 61, 89-99.
9. Avilés-Sacoto SV, Cook WD, Imanirad R and Zhu J (2015). Two-stage network DEA: when intermediate measures can be treated as outputs from the second stage. *Journal of the Operational Research Society*, 66(11): 1868–1877.
10. Avkiran N (2009). Opening the black box of efficiency analysis: An illustration with UAE banks. *Omega*, 37, 930–941.
11. Ballestero E, Romero C (1991). A theorem connecting utility function optimization and compromise programming theorem connecting utility function. *Operations Research Letters*, 1, 421-427.
12. Banker RD (1984). Estimating most productive scale size using data envelopment analysis. *European Journal of Operational Research*, 17, 35-44.

13. Banker RD, Charnes A, Cooper WW (1984). Some models for estimating technical and scale inefficiencies in data envelopment analysis. *Management Science*, 30, 1078-1092.
14. Banker RD, Cooper WW, Seiford LM, Thrall RM, Zhu J (2004). Returns to scale in different DEA models. *European Journal of Operational Research*, 154, 345–362.
15. Banker RD, Thrall RM (1992). Estimation of returns to scale using Data Envelopment Analysis. *European Journal of Operational Research*, 62, 74-84.
16. Bazaraa MS, Jarvis JJ, Sherali HD (2011). Linear programming and network flows. Third Edition, John Wiley & Sons, Inc., Hoboken, New Jersey, USA.
17. Beasley JE (1995). Determining teaching and research efficiencies. *The Journal of the Operational Research Society*, 46(4), 441-452.
18. Bowman Jr VJ (1976). On the Relationship of the Tchebycheff Norm and the Efficient Frontier of Multiple-Criteria Objectives. In: Thiriez H, Zionts S (eds) Multiple Criteria Decision Making. Lecture Notes in Economics and Mathematical Systems, 130, Springer-Verlag, Berlin, Heidelberg, 76-85.
19. Buchanan J, Gardiner L (2003). A comparison of two reference point methods in multiple objective mathematical programming. *European Journal of Operational Research*, 149(1), 17-34.
20. Cao X, Yang F (2011). Measuring the performance of Internet companies using a two-stage data envelopment analysis model. *Enterprise Information Systems*, 5(2), 207-217.
21. Castelli L, Pesenti R, Ukovich W (2010). A classification of DEA models when the internal structure of the Decision Making Units is considered. *Annals of Operations Research*, 173(1), 207-235.
22. Caves DW, Christensen LR, Diewert WE (1982). The Economic Theory of Index Numbers and the Measurement of Input, Output, and Productivity. *Econometrica*, 50(6), 1393-1414.
23. Censor Y (1977). Pareto Optimality in Multiobjective Problems. *Applied Mathematics and Optimization*, 4, 41-59.
24. Chang YT, Lee S, Park HK (2017). Efficiency analysis of major cruise lines. *Tourism Management*, 58, 78-88.
25. Chankong V, Haimes YY (1983).Optimization-Based Methods for Multiobjective Decision-Making: An Overview. *Large Scale Systems*, 5(1), 1-33.

26. Charnes A, Cooper WW (1961). Management models and industrial applications of linear programming. Wiley, New York.
27. Charnes A, Cooper WW (1962). Programming with linear fractional functional. *Naval Research Logistics*, 9, 181-185.
28. Charnes A, Cooper WW, Rhodes E (1978). Measuring the efficiency of decision making units. *European Journal of Operational Research*, 2(6), 429-444.
29. Charnes A, Cooper WW, Golany B, Seiford L, Stutz J (1985). Foundations of data envelopment analysis for Pareto–Koopmans efficient empirical production functions. *Journal of Econometrics*, 30(1-2), 91–107.
30. Charnes A, Cooper WW, Golany B, Halek R, Klopp G, Schmitz E, Thomas D (1986). Two phase data envelopment analysis approach to policy evaluation and management of army recruiting activities: tradeoffs between joint services and army advertising. Research Report CCS no. 532, Center for Cybernetic Studies, The University of Texas, Austin, Texas.
31. Charnes A, Cooper WW, Wei QL, Huang ZM (1989). Cone ratio data envelopment analysis and multi-objective programming. *International Journal of Systems Science*, 20(7), 1099-1118.
32. Chen C, Yan H (2011). Network DEA model for supply chain performance evaluation. *European Journal of Operational Research*, 213(1), 147–155.
33. Chen C, Zhu J, Yu JY, Noori H (2012). A new methodology for evaluating sustainable product design performance with two-stage network data envelopment analysis. *European Journal of Operational Research*, 221(2), 348–359.
34. Chen Y and Zhu J (2004). Measuring information technology's indirect impact on firm performance. *Information Technology and Management*, 5, 9-22.
35. Chen Y, Liang L, Yang F, Zhu J (2006). Evaluation of information technology investment: A data envelopment analysis approach. *Computers and Operations Research*, 33, 1368-1379.
36. Chen Y, Liang L, Zhu J (2009a). Equivalence in two-stage DEA approaches. *European Journal of Operational Research*, 193, 600-604.
37. Chen Y, Cook WD, Li N, Zhu J (2009b). Additive efficiency decomposition in two-stage DEA. *European Journal of Operational Research*, 196, 1170-1176.
38. Chen Y, Cook WD, Zhu J (2010a). Deriving the DEA frontier for two-stage DEA processes. *European Journal of Operational Research*, 202, 138-142.

39. Chen Y, Du J, Sherman HD, Zhu J (2010b). DEA model with shared resources and efficiency decomposition. *European Journal of Operational Research*, 207,339–349.
40. Chen Y, Cook WD, Kao C, Zhu J (2013). Network DEA pitfalls: Divisional efficiency and frontier projection under general network structures. *European Journal of Operational Research*, 226, 507-515.
41. Chilingerian JA, Sherman HD (1990). Managing physician efficiency and effectiveness in providing hospital services. *Health Services Management Research*, l3(1), 3-15.
42. Choo EU, Atkins DR (1980). An interactive algorithm for multicriteria programming. *Computers and Operations Research*, 7, 81-87.
43. Cook WD, Hababou M, Tuenter HJH (2000). Multicomponent efficiency measurement and shared inputs in data envelopment analysis: An application to sales and service performance in bank branches. *Journal of Productivity Analysis*, 14, 209–224.
44. Cook WD, Hababou M (2001). Sales performance measurement in bank branches. *Omega*, 29, 299-307.
45. Cook WD, Green RH (2004). Multicomponent efficiency measurement and core business identification in multiplant firms: A DEA model. *European Journal of Operational Research*, 157, 540–551.
46. Cook WD, Seiford LM (2009). Data envelopment analysis (DEA) – Thirty years on. *European Journal of Operational Research*, 192(1), 1-17.
47. Cook WD, Liang L, Zhu J (2010a). Measuring performance of two-stage network structures by DEA: A review and future perspective. *Omega*, 38(6), 423–430.
48. Cook WD, Zhu J, Bi G, Yang F (2010b). Network DEA: Additive efficiency decomposition. *European Journal of Operational Research*, 207(2), 1122-1129.
49. Cook WD, Zhu J (eds) (2014). Data Envelopment Analysis - A Handbook of Modeling Internal Structure and Network, Springer, New York.
50. Cooper WW (2005). Origins, Uses of, and Relations Between Goal Programming and Data Envelopment Analysis. *Journal of Multi-Criteria Decision Analysis*, 13, 3-11.
51. Cooper WW (2014). Origin and Development of Data Envelopment Analysis: Challenges and Opportunities. *Data Envelopment Analysis Journal*, 1, 3–10.
52. Cooper WW, Lovell CAK (2011). History lessons. *Journal of Productivity Analysis*, 36, 193–200.

53. Cooper WW, Seiford LM, Tone K, Zhu J (2007). Some models and measures for evaluating performances with DEA: past accomplishments and future prospects. *Journal of Productivity Analysis*, 28, 151–63.
54. Da Cruz NF, Carvalho P, Marques RC (2013). Disentangling the cost efficiency of jointly provided water and wastewater services. *Utilities Policy*, 24, 70-77.
55. Debreu G (1951). The coefficient of resource utilization. *Econometrica*, 19(3), 273-292.
56. Despotis DK (1996). Fractional minmax goal programming: A unified approach to priority estimation and preference analysis in MCDM. *The Journal of the Operational Research Society*, 47, 989-999.
57. Diez-Ticio A, Mancebon MJ (2002). The efficiency of the Spanish police service: An application of the multiactivity DEA model. *Applied Economics*, 34, 351–362.
58. Dyson RG, Allen R, Camanho AS, Podinovski VV, Sarrico CS, Shale EA (2001). Pitfalls and protocols in DEA. *European Journal of Operational Research*, 132(2), 245–259.
59. Ehrgott M (2000). Multicriteria optimization. Lecture notes in economics and mathematical systems. Springer-Verlag, Berlin, Heidelberg.
60. Ehrgott M (2005). Multicriteria Optimization. Springer-Verlag, Berlin, Heidelberg.
61. Gattoufi S, Oral M, Reisman A (2004). A taxonomy for data envelopment analysis. *Socio-Economics Planning Science*, 38(2-3), 141-158.
62. Fare R, Primont D (1984). Efficiency measures for multiplant firms. *Operations Research Letters*, 3, 257-260.
63. Fare R (1991). Measuring Farrell Efficiency for a Firm with Intermediate Inputs. *Academia Economic Papers*, 19, 829-40.
64. Fare R, Grosskopf S (1992a). Malmquist Productivity Indexes and Fisher Ideal Indexes. *The Economic Journal*, 102(410), 158-160.
65. Fare R, Grosskopf S, Li SK (1992b). Linear Programming Models for Firm and Industry Performance. *The Scandinavian Journal of Economics*, 94(4), 599-608.
66. Fare R, Primont D (1993). Measuring the efficiency of multiunit banking: An activity analysis approach. *Journal of Banking and Finance*, 17, 539-544.
67. Fare R, Whittaker G (1995). An intermediate input model of dairy production using complex survey data. *Journal of Agricultural Economics*, 46(2), 201–23.
68. Fare R, Grosskopf S (1996). Productivity and intermediate products: a frontier approach. *Economic Letters*, 50, 65-70.

69. Fare R, Grabowski R, Grosskopf S, Kraft S (1997). Efficiency of a fixed but allocatable input: A non-parametric approach. *Economics Letters*, 56, 187-193.
70. Fare R, Grosskopf S (2000). Network DEA. *Socio-Economic Planning Sciences*, 34(1), 35-49.
71. Farrell MJ (1957). The measurement of productive efficiency. *Journal of the Royal Statistical Society*. *Series A (General)*, 120(3), 253-281.
72. Farrell MJ and Fieldhouse M (1962). Estimating Efficient Productions Functions under Increasing Returns to Scale. *Journal of the Royal Statistical Society*. *Series A (General)*, 125(2), 252–267.
73. Thanassoulis E, Portela MCS, Despic O (2008). Data Envelopment Analysis: The Mathematical Programming Approach to Efficiency Analysis. In: Fried HO, Lovell CAK, Schmidt SS (eds) The Measurement of Productive Efficiency and Productivity Growth. Oxford University Press.
74. Forsund FR, Sarafoglou N (2002). On the Origins of Data Envelopment Analysis. *Journal of Productivity Analysis*, 17, 23–40.
75. Fukuyama H, Weber WL (2010). A slacks-based inefficiency measure for a two-stage system with bad outputs. *Omega*, 38, 398-409.
76. Fukuyama H, Mirdehghan SM (2012). Identifying the efficiency status in network DEA. *European Journal of Operational Research*, 220, 85-92.
77. Gass S, Saaty T (1955).The computational algorithm for the parametric objective function. *Naval Research Logistics Quarterly*, 2, 39-45.
78. Geoffrion AM (1968). Proper efficiency and theory of vector maximization. *Journal of Mathematical Analysis and Applications*, 22, 618–630.
79. Golany B (1988). An interactive MOLP procedure for the extension of DEA to effectiveness analysis. *The Journal of the Operational Research Society*, 39, 725-734.
80. Guan JC, Chen KH (2010). Measuring the innovation production process: A cross-region empirical study of China’s high-tech innovations. *Technovation*, 30, 348-358.
81. Guan JC, Chen KH (2012). Modeling the relative efficiency of national innovation systems. *Research Policy*, 41, 102-115.
82. Guo C, Shureshjani RA, Foroughi AA, Zhu J (2017). Decomposition weights and overall efficiency in two-stage additive network DEA. *European Journal of Operational Research*, 257(3), 896-906.

83. Halkos GE, Tzeremes NG, Kourtzidis SA (2014). A unified classification of two-stage DEA models. *Surveys in Operations Research and Management Science*, 19, 1–16.
84. Halkos GE, Tzeremes NG, Kourtzidis SA (2015a). Weight assurance region in two-stage additive efficiency decomposition DEA model: an application to school data. *Journal of the Operational Research Society*, 66(4), 696-704.
85. Halkos GE, Tzeremes NG, and Kourtzidis SA (2015b).Regional sustainability efficiency index in Europe: an additive two-stage DEA approach. *Operational Research*, 15(1), 1-23.
86. Halkos GE, Tzeremes NG, Kourtzidis SA (2016). Measuring Sustainability Efficiency Using a Two-Stage Data Envelopment Analysis Approach. *Journal of Industrial Ecology*, 20(5), 1159-1175.
87. Hsieh LF, Lin LH (2010). A performance evaluation model for international tourist hotels in Taiwan: An application of the relational network DEA. *International Journal of Hospitality Management*, 29, 14–24.
88. Hwang C-H, Masud ASM (1979). Multiple objective decision making, methods and applications: a state-of-the-art survey. Springer-Verlag, Berlin, Heidelberg.
89. Jahanshahloo GR, Amirteimoori AR, Kordrostami S (2004). Multi-component performance, progress and regress measurement and shared inputs and outputs in DEA for panel data: an application in commercial bank branches. *Applied Mathematics and Computation*, 151, 1-16.
90. Joro T, Korhonen P, Wallenius J (1998). Structural comparison of data envelopment analysis and multiple objective linear programming. *Management Science*, 44, 962–970.
91. Kaliszewski I (1994). Qualitative Pareto Analysis by Cone Separation Technique, Kluwer Academic Publishers, Boston, Massachusetts.
92. Kaliszewski I (2004). Out of the mist-towards decision-maker-friendly multiple criteria decision making support. *European Journal of Operational Research*, 158, 293-307.
93. Kao C (1998). Measuring the efficiency of forest districts with multiple working circles. *The Journal of Operational Research Society*, 49(6), 583-590.
94. 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.

95. Kao C (2009a). Efficiency decomposition in network data envelopment analysis: A relational model. *European Journal of Operational Research*, 192(3), 949–962.
96. Kao C (2009b). Efficiency measurement for parallel production systems. *European Journal of Operational Research*, 196, 1107–1112.
97. Kao C, Hwang S-N (2010). Efficiency measurement for network systems: IT impact on firm performance. *Decision Support Systems*, 48, 437–446.
98. Kao C, Hwang S-N (2011). Decomposition of technical and scale efficiencies in two-stage production systems. *European Journal of Operational Research*, 211, 515-519.
99. Kao C (2012). Efficiency decomposition for parallel production systems. *Journal of the Operational Research Society*, 63, 64–71.
100. Kao C (2014a). Efficiency decomposition for general multi-stage systems in data envelopment analysis. *European Journal of Operational Research*, 232(1), 117–124.
101. Kao C (2014b). Network data envelopment analysis: A review. *European Journal of Operational Research*, 239(1), 1–16.
102. Kao C (2016). Efficiency decomposition and aggregation in network data envelopment analysis. *European Journal of Operational Researc*h, 255(3), 778-786.
103. Kao C (2017). Network Data Envelopment Analysis - Foundations and Extensions. Springer International Publishing, Switzerland.
104. Keh HT, Chu S (2003). Retail productivity and scale economies at the firm level: a DEA approach. *Omega*, 31(2), 75–82.
105. Keh HT, Chu S, Xu J (2006). Efficiency, effectiveness and productivity of marketing in services. *European Journal of Operational Research*, 170(1), 265–276.
106. Koopmans TC (1951). An analysis of production as an efficient combination of activities. In: Koopmans TC (ed) Activity analysis of production and allocation, cowles commission for research in economics, Monograph 13, John Wiley and Sons, New York, 33-97.
107. Kornbluth JSH (1991). Analysing policy effectiveness using cone restricted data envelopment analysis. *Journal of the Operational Research Society*, 42, 1097–1104.
108. Lee CY, Johnson AL (2011). A decomposition of productivity change in the semiconductor manufacturing industry. *International Journal of Production Research*, 49(16), 4761–4785.
109. Lee CY, Johnson AL (2012). Two-dimensional efficiency decomposition to measure the demand effect in productivity analysis. *European Journal of Operational Research*, 216, 584-593.

110. Lewis HF, Sexton TR (2004). Network DEA: efficiency analysis of organizations with complex internal structure. *Computers and Operations Research*, 31, 1365-1410.
111. Li Y, Chen Y, Liang L, Xie J (2012). DEA models for extended two-stage network structures. *Omega*, 40(5), 611-618.
112. Li Y, Lei X, Dai Q, Liang L (2015) .Performance evaluation of participating nations at the 2012 London Summer Olympics by a two-stage data envelopment analysis. *European Journal of Operational Research*, 243(3), 964–973.
113. Liang L, Yang F, Cook WD, Zhu J (2006). DEA models for supply chain efficiency evaluation. *Annals of Operations Research*, 145(1), 35-49.
114. Liang L, Cook WD, Zhu J (2008). DEA models for two-stage processes: Game approach and efficiency decomposition. *Naval Research Logistics*, 55, 643-653.
115. Liang L, Li ZQ, Cook WD, Zhu J (2011). Data envelopment analysis efficiency in two-stage networks with feedback. *IIE Transactions*, 43, 309–322.
116. Lightner M, Director S (1981). Multiple criterion optimization for the design of electronic circuits. *IEEE Transactions on Circuits and Systems*, 28(3), 169-179.
117. Limaei SM (2013). Efficiency of Iranian forest industry based on DEA models. *Journal of Forestry Research*, 24, 759-765.
118. Lin JG (1976). Three Methods for Determining Pareto-Optimal Solutions of Multiple-Objective Problems. In: Ho YC, Mitter SK (eds) Directions in Large-Scale Systems: Many-Person Optimization and Decentralized Control, Springer US, Boston, 117-138.
119. Lin TY, Chiu SH (2013). Using independent component analysis and network DEA to improve bank performance evaluation. *Economic Modelling*, 32, 608-616.
120. Liu JS, Lu LYY, Lu WM, Lin BJY (2013). Data envelopment analysis 1978-2010: A citation-based literature survey. *Omega*, 41(1), 3-15.
121. Liu ST (2011). Performance measurement of Taiwan financial holding companies: An additive efficiency decomposition approach. *Expert Systems with Applications*, 38, 5674–5679.
122. Liu ST, Wang RT (2009). Efficiency measures of PCB manufacturing firms using relational two-stage data envelopment analysis. *Expert Systems with Applications*, 36, 4935-4939.
123. Lo SF, Lu WM (2009). An integrated performance evaluation of financial holding companies in Taiwan. *European Journal of Operational Research*, 198, 341-350.

124. Lo SF (2010). Performance evaluation for sustainable business: A profitability and marketability framework. *Corporate Social Responsibility and Environmental Management*, 17, 311-319.
125. Lozano S, Gutierrez E, Moreno P (2013). Network DEA approach to airports performance assessment considering undesirable outputs. *Applied Mathematical Modelling*, 37, 1665–1676.
126. Lothgren M, Tambour M (1999). Productivity and customer satisfaction in Swedish pharmacies: A DEA network model. *European Journal of Operational Research*, 115, 449-458.
127. Lu WC (2009). The evolution of R&D efficiency and marketability: Evidence from Taiwan's IC-design Industry. *Asian Journal of Technology Innovation*, 17, 1-26.
128. Lu WM, Wang WK, Hung SW, Lu ET (2012). The effects of corporate governance on airline performance: Production and marketing efficiency perspectives. *Transportation Research Part E: Logistics and Transportation Review*, 48(2), 529-544.
129. Lu WM, Kweh QL, Huang CL (2014). Intellectual capital and national innovation systems performance. *Knowledge-Based Systems*, 71, 201-210.
130. Luo XM (2003). Evaluating the profitability and marketability efficiency of large banks: An application of data envelopment analysis. *Journal of Business Research*, 56, 627–635.
131. Mar Molinero C (1996). On the Joint Determination of Efficiencies in a Data Envelopment Analysis Context. *The Journal of the Operational Research Society*, 47(10), 1273-1279.
132. Matthews K (2013). Risk management and managerial efficiency in Chinese banks: A network DEA framework. *Omega*, 41(2), 207-215.
133. Michalowski W, Szapiro T (1992). A bi-reference procedure for interactive multiple criteria programming. *Operations Research*, 40, 247-258.
134. Miettinen K (1999). Nonlinear Multiobjective Optimization. Springer Science & Business Median, New York.
135. Miettinen K, Makela MM (2002). On scalarizing functions in multiobjective optimization. *OR Spectrum*, 24, 193–213.
136. Mirdehghan SM, Fukuyama H (2016). Pareto–Koopmans efficiency and network DEA. *Omega*, 61, 78-88.

137. Ogryczak W (2001). Comments on properties of the minmax solutions in goal programming. *European Journal of Operational Research*, 132(1), 17-21.
138. Olson DL (1993). Tchebycheff norms in multi-objective linear programming. *Mathematical and Computer Modelling*, 17, 113–124.
139. Premachandra IM, Zhu J, Watson J, Galagedera DUA (2012). Best performing US mutual fund families from 1993 to 2008: Evidence from a novel two-stage DEA model for efficiency decomposition. *Journal of Banking and Finance*, 36, 3302-3317.
140. Prieto AM, Zofío JL (2007). Network DEA efficiency in input-output models: With an application to OECD countries. *European Journal of Operational Research*, 178, 292–304.
141. Reeves G R, MacLeod KR (1999). Some experiments in Tchebycheff-based approaches for interactive multiple objective decision making. *Computers and Operational Research*, 26(13), 1311 – 1321.
142. Rogge N, Jaeger S (2012). Evaluating the efficiency of municipalities in collecting and processing municipal solid waste: A shared input DEA-model. *Waste Management*, 32, 1968-1978.
143. Seiford LM (1996). Data Envelopment Analysis: The Evolution of the State of the Art (1978-1995). *The Journal of Productivity Analysis*, 7, 99-137.
144. Seiford LM, Zhu J (1999). Profitability and marketability of the top 55 US commercial banks. *Management Science*, 45(9), 1270-1288.
145. Sheth C, Triantis K, Teodorovic D (2007). Performance evaluation of bus routes: A provider and passenger perspective. *Transportation Research Part E: Logistics and Transportation Review*, 43(4), 453–478.
146. Sexton TR, Lewis HF (2003). Two-stage DEA: An application to major league baseball. *Journal of Productivity Analysis*, 19(2-3), 227-249.
147. Steuer RE, Choo EU (1983). An interactive weighted Tchebycheff procedure for multiple objective programming. *Mathematical Programming*, 26(3), 326-44.
148. Steuer RE (1986). Multiple criteria optimization: Theory, computation, and applications. Wiley, New York.
149. Stewart TJ (1996). Relationships between data envelopment analysis and multicriteria decision analysis. *Journal of the Operational Research Society*, 47(5), 654–665.
150. Soteriou A, Zenios SA (1999). Operations, quality, and profitability in the provision of banking services. *Management Science*, 45, 1221-1238.

151. Toloo M, Emrouznejad A, Moreno P (2015). A linear relational DEA model to evaluate two-stage processes with shared inputs. *Computational and Applied Mathematics*, 1, 1-17.
152. Tamiz M, Jones D, Romero C (1998). Goal programming for decision making: an overview of the current state-of-the-art. *European Journal of Operational Research*, 111, 569-581.
153. Tone K (2001). A slacks-based measure of efficiency in data envelopment analysis. *European journal of operational research*, 130(3), 498-509.
154. Tone K, Tsutsui M (2009). Network DEA: A slacks-based measure approach. *European Journal of Operational Research*, 197, 243-252.
155. Tsai PF, Mar Molinero C (2002). A variable returns to scale data envelopment analysis model for the joint determination of efficiencies with an example of the UK health service. *European Journal of Operational Research*, 141, 21–38.
156. Tsolas IE (2011). Relative profitability and stock market performance of listed commercial banks on the Athens Exchange: A non-parametric approach. *IMA Journal of Management Mathematics*, 22, 323–342.
157. Tsolas IE (2013). Modeling profitability and stock market performance of listed construction firms on the Athens Exchange: Two-stage DEA approach. *Journal of Construction Engineering and Management*, 139, 111–119.
158. Tsutsui M, Goto M (2009). A multi-division efficiency evaluation of US electric power companies using a weighted slacks-based measure. *Socio-Economic Planning Sciences*, 43(3), 201-208.
159. Vaz CB, Camanho AS, Guimaraes RC (2010). The assessment of retailing efficiency using network data envelopment analysis. *Annals of Operations Research*, 173, 5–24.
160. Wang CH, 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.
161. Wang K, Huang W, Wu J, Liu YN (2014). Efficiency measures of the Chinese commercial banking system using an additive two-stage DEA. *Omega*, 44, 5-20.
162. Wanke PF (2013). Physical infrastructure and shipment consolidation efficiency drivers in Brazilian ports: A two-stage network-DEA approach. *Transport Policy*, 29, 145–153.
163. Wanke P, Barros C (2014). Two-stage DEA: An application to major Brazilian banks. *Expert Systems with Applications*, 41, 2337–2344.

164. Wanke P, Blackburn V, Barros CP (2016). Cost and learning efficiency drivers in Australian schools: a two-stage network DEA approach. *Applied Economics*, 48(38), 3577-3604.
165. Wierzbicki AP (1980). The use of reference objectives in multiobjective optimization. In: G. Fandel G, Gal T (eds) Multiple Criteria Decision Making Theory and Application. Lectures Notes in Economics and Mathematical Systems 177, Springer, New York, 468-486.
166. Wu J, Yin P, Sun J, Chu J, Liang L (2016). Evaluating the environmental efficiency of a two-stage system with undesired outputs by a DEA approach: An interest preference perspective. *European Journal of Operational Research*, 254(3), 1047-1062.
167. Yang F, Wu D, Liang L, Bi G, Wu DD (2011). Supply chain DEA: production possibility set and performance evaluation model. *Annals of Operations Research*, 185(1), 195–211.
168. Yang CH, Lin HY, Chen CP (2014). Measuring the efficiency of NBA teams: additive efficiency decomposition in two-stage DEA. *Annals of Operations Research*, 217(1), 565-589.
169. Yu MM (2008). Measuring the efficiency and return to scale status of multimode bus transit – evidence from Taiwan's bus system. *Applied Economics Letters*, 15, 647–653.
170. Yu MM (2010). Assessment of airport performance using the SBM-NDEA model. *Omega*, 38(6), 440-452.
171. Yu MM, Fan CK (2009). Measuring the performance of multimode bus transit: A mixed structure network DEA model. *Transportation Research Part E: Logistics and Transportation Review*, 45, 501–515.
172. Yu PL (1973). A Class of Solutions for Group Decision Problems. *Management Science*, 19, 936-946.
173. Zeleny M (1973). Compromise Programming. In: Cochrane JL, Zeleny M (eds) Multiple Criteria Decision Making. University of South Carolina Press, Columbia, SC, 262-301.
174. Zha Y, Liang L (2010). Two-stage cooperation model with input freely distributed among the stages. *European Journal of Operational Research*, 205, 332-338.
175. Zhu J (2000). Multi-factor performance measure model with an application to Fortune 500 companies. *European Journal of Operational Research*, 123(1), 105-124.

176. Zhu J (2011). Airlines Performance via Two-Stage Network DEA Approach. *Journal of CENTRUM Cathedra*, 4(2), 260-269.