Theoretical and empirical essays on Value Efficiency Analysis

Panagiotis Ravanos

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

Doctor of Philosophy

University of Macedonia

School of Economic and Regional Studies

Department of Economics

Thessaloniki, Greece

Supervisor: Prof. Giannis Karagiannis

JEL Classification codes: C02; C61; D24; 015

April 2023

To my parents

Supervisor

Professor Giannis Karagiannis, Department of Economics, University of Macedonia, Greece

Thesis advisors

Professor Andreas Georgiou, Department of Business Administration, University of Macedonia, Greece

Professor Nikolaos Tzeremes, Department of Economics, University of Thessaly, Greece

Thesis examiners

Professor Vasileios Aletras, Department of Business Administration, University of Macedonia, Greece

Senior Lecturer Stavros Kourtzidis, Dundee Business School, University of Dundee, UK

Emeritus Professor Emmanuel Thanasoulis, College of Business and Social Sciences, Aston University, UK.

Professor Konstantinos Veletzas, Department of Economics, University of Macedonia, Greece

(this page intentionally left blank)

Summary (in English)

The aim of this thesis is twofold: first, to examine the features of Value Efficiency Analysis (VEA) and second, to provide some innovative applications of it. VEA is a performance evaluation method that incorporates managerial preferences in DEA models through the Most Preferred Solution (MPS), namely a real or artificial efficient DMU that has the most preferred input/output bundle or structure.

The thesis' first part comprises of three empirical essays. In the first of them, we use VEA for the assessing effectiveness, namely the extent that DMUs are "doing the right things" such as following organizational objectives or abiding by agreements. This is demonstrated by using VEA to assess the effectiveness of countries in utilizing their income to develop their citizens' social prosperity or human capabilities.

In the second empirical essay, we introduce VEA to pure inputs DEA models by developing the VEA-Benefit-of-the-Doubt (BoD) model. This model is an alternative to incorporating DM preferences in the construction of composite indicators of socio-economic phenomena, and is used to re-estimate the United Nations Human Development Index (HDI).

In the third empirical essay, we review the list of suggestions for choosing the MPS in VEA and enlarge it with four additional ones. These are based on the relative position of efficient DMUs on the DEA frontier, the notions of the Most Productive Scale size (MPSS) and the Average Production Unit (APU), and common weights. We also conduct comparative empirical analysis of the effect of alternative MPSs on the VEA efficiency scores, the results of which provide useful information regarding the MPS choices that are more likely to offer additional insights to management compared to those of the DEA model, as well as those choices which are frequently similar to each other in practice.

The second part of the thesis comprises of three theoretical essays. In the first of them, we relate VEA to cross efficiency and show that the VEA model is equivalent to the Targetted Benevolence (TB) cross efficiency model. This allows obtaining, for the first time, the cross efficiency scores from the envelopment form of the VEA model, while it provides shortcuts in the estimation of the TB cross efficiency matrix.

In the second theoretical essay, we show that the VEA model can be viewed as a particular class of DEA models with production trade-offs or their dual weight restrictions. The VEA efficiency scores can thus be interpreted as incorporating a particular form of trade-off relations in a DEA model, while the efficiency scores of the DEA models with these particular trade-offs can be viewed as reflecting the judgements of a DM regarding the most preferred input/output structure.

In the third theoretical essay, we show that the VEA efficiency scores are equivalent to those obtained from Cone-Ratio (CR) DEA models that incorporate preferences about efficient DMUs that are considered as examples to follow (*model* DMUs) for the remaining DMUs, or provide upper/lower bound approximations of them. These relations allow obtaining or approximating the efficiency scores of CR-DEA models by means of VEA models, which are less computationally intensive than CR-DEA models.

Summary (in Greek)

Το αντικείμενο της παρούσας διδακτορικής διατριβής είναι η θεωρητική και εμπειρική διερεύνηση της μεθόδου της Αξιακής Ανάλυσης Αποτελεσματικότητας (ΑΑΑ). Η ΑΑΑ είναι μια μέθοδος εκτίμησης της σχετικής αποτελεσματικότητας των Μονάδων Λήψης Απόφασης (ΜΛΑ), η οποία περιλαμβάνει την ενσωμάτωση επιπλέον αξιακών περιορισμών στο υπόδειγμα της Περιβάλλουσας Ανάλυσης Δεδομένων (ΠΑΔ), μέσω της επιλογής της Πλέον Προτιμώμενης Λύσης (ΠΠΛ). Η ΠΠΛ είναι μια αποτελεσματική ΜΛΑ ή ένας αποτελεσματικός γραμμικός συνδυασμός ΜΛΑ, η οποία έχει το επιθυμητό μείγμα εισροών/εκροών.

Το πρώτο μέρος της διατριβής αποτελείται από τρία εμπειρικά άρθρα. Στο πρώτο από αυτά, παρουσιάζεται η χρήση της ΑΑΑ ως εργαλείου για την εκτίμηση της αποδοτικότητας, δηλαδή του κατά πόσο η λειτουργία των αξιολογούμενων ΜΛΑ συμβαδίζει με δεδομένους στόχους (όπως οργανωτικοί στόχοι, συμφωνίες, κ.α.). Στην συνέχεια η ΑΑΑ χρησιμοποιείται για να εκτιμηθεί η αποδοτικότητα των χωρών του κόσμου στην χρησιμοποίηση του εθνικού του εισοδήματος για την επίτευξη της μέγιστης δυνατής κοινωνικής ευημερίας των πολιτών τους.

Στο δεύτερο εμπειρικό άρθρο, η ΑΑΑ χρησιμοποιείται ως εναλλακτική μέθοδος για την ενσωμάτωση αξιακών περιορισμών στην κατασκευή συνθετικών δεικτών. Συγκεκριμένα, κατασκευάζεται ένα υπόδειγμα ΑΑΑ το οποίο διαθέτει μόνο εκροές και χρησιμοποιείται για την επανεκτίμηση του Δείκτη Ανθρώπινης Ανάπτυξης των Ηνωμένων Εθνών.

Στο τρίτο εμπειρικό άρθρο, πραγματοποιείται μια ανασκόπηση των προτάσεων που περιλαμβάνονται στην βιβλιογραφία της ΑΑΑ για την επιλογή της ΠΠΛ και προτείνονται τέσσερις νέες, οι οποίες βασίζονται αντίστοιχα στα χαρακτηριστικά της θέσης των αποτελεσματικών ΜΛΑ στην εν δυνάμει συνάρτηση μετασχηματισμού, στις έννοιες της Αποτελεσματικής Κλίμακας Παραγωγής και της Μέσης Παραγωγικής Μονάδας, και στην εκτίμηση αποτελεσματικότητας μέσω κοινών μεταξύ των ΜΛΑ σχετικών σταθμίσεων για τις εισροές και τις εκροές. Στην συνέχεια αναλύεται χρησιμοποιώντας εμπειρικά δεδομένα η επίδραση της επιλογής διαφορετικών ΠΠΛ στην εκτιμηθείσα σχετική αποτελεσματικότητα μέσω της ΑΑΑ. Από τα αποτελέσματα της εμπειρικής ανάλυσης εξάγονται χρήσιμα συμπεράσματα για τις ΠΠΛ η χρήση των οποίων οδηγεί σε αποτελέσματα τα οποία παρέχουν επιπλέον πληροφόρηση σε σχέση με αυτά της ΠΑΔ, η οποία μπορεί να χρησιμοποιηθεί από τις διοικητικές υπηρεσίες για την λήψη επιχειρηματικών αποφάσεων, αλλά και για τις διαφορετικές ΠΠΛ οι οποίες οδηγούν πρακτικά σε παρόμοια αποτελέσματα ΑΑΑ.

Το δεύτερο μέρος της διατριβής περιλαμβάνει τρία θεωρητικά άρθρα. Στο πρώτο από αυτά, διερευνάται η σχέση μεταξύ της ΑΑΑ και της μεθόδου της σταυροειδούς αποτελεσματικότητας και αποδεικνύεται ότι το υπόδειγμα της ΑΑΑ είναι ισοδύναμο με το υπόδειγμα στοχευμένου αλτρουισμού της σταυροειδούς αποτελεσματικότητας. Αυτή η ισοδυναμία επιτρέπει για πρώτη φορά στην βιβλιογραφία την εκτίμηση σταυροειδών αποτελεσματικοτήτων μέσω του δυικού-αντί του πρωταρχικού-υποδείγματος της ΑΑΑ και οδηγεί σε πιο σύντομη διαδικασία εκτίμησης του πίνακα σταυροειδών αποτελεσματικοτήτων του υποδείγματος στοχευμένου αλτρουισμού.

Στο δεύτερο θεωρητικό άρθρο, αποδεικνύεται ότι το υπόδειγμα της ΑΑΑ αποτελεί μια συγκεκριμένη κατηγορία υποδειγμάτων ΠΑΔ που περιλαμβάνουν περιορισμούς στις σχετικές σταθμίσεις των εισροών και των εκροών. Κατά συνέπεια, μπορεί να εξαχθεί μια επιπλέον ερμηνεία των αποτελεσμάτων της ΑΑΑ, ως

αποτελέσματα ΠΑΔ που ενσωματώνουν μια συγκεκριμένη μορφή περιορισμών στις σχετικές σταθμίσεις των εισροών και των εκροών, ενώ οι σχετικές αποτελεσματικότητες από τα σχετικά υποδείγματα ΠΑΔ ενσωματώνουν τις κρατούσες αντιλήψεις σχετικά με το βέλτιστο μείγμα εισροών/εκροών.

Στο τελευταίο θεωρητικό άρθρο, αποδεικνύεται ότι τα αποτελέσματα της ΑΑΑ (σχετικές αποτελεσματικότητες για τις ΜΛΑ) είναι είτε ισοδύναμα με τα αποτελέσματα υποδειγμάτων ΠΑΔ τα οποία ενσωματώνουν πληροφορίες σχετικά με τις αποτελεσματικές ΜΛΑ που θεωρούνται πρότυπα για τις υπόλοιπες (επιτρέποντας στα διανύσματα των σχετικών σταθμίσεων των εισροών και εκροών να παίρνουν τιμές μέσα σε ένα προκαθορισμένο εύρος / κώνο διανυσμάτων), είτε αποτελούν το ανώτερο και το κατώτερο όριο για αυτά. Οι αποδειχθείσες μαθηματικές σχέσεις επιτρέπουν την εκτίμηση ή την προσέγγιση των αποτελεσμάτων των υποδειγμάτων ΠΑΔ χρησιμοποιώντας τα υποδείγματα ΑΑΑ, τα οποία ενέχουν σχετικά λιγότερη υπολογιστική πολυπλοκότητα.

Keywords

Value Efficiency Analysis; Data Envelopment Analysis; efficiency; Most Preferred Solution; weight restrictions; production trade-offs; model DMUs; Targeted Benevolence cross efficiency; effectiveness; Benefit-of-the-Doubt; Most Productive Scale Size; Average Production Unit; common weights; interior-exterior DMUs, terminal DMUs; Human Development Index; human capabilities; social efficiency; social prosperity

JEL Classification codes

C02; C61; D24; O15

Acknowledgements

If I was Väinämöinen, the legendary wizard-singer in the Finnish epic poem of Kalevala, it would take me about three days and nights of singing and this PhD thesis would be finished. As I obviously do not possess such magical powers, this thesis had to be completed within more than five years through the usual, hard-working, and educative way. Fortunately, there are some people and institutions whose moral, scientific, and pecuniary support made this marathon not only less difficult, but, in the end, worth every minute. This section is my attempt to thank each and every one of them. I sincerely apologize for any omissions.

I am deeply indebted to the institutions which provided me with financial support throughout these five years. I am deeply indebted to the Greek General Secretariat for Research and Technology (GSRT) and the Hellenic Foundation for Research and Innovation (HFRI) for providing me with a PhD Fellowship grant (GA. no. 698) under the program "1st Call for H.F.R.I. Scholarships for PhD Candidates". Also, I graciously thank the Department of Economics of the University of Macedonia for funding my participation in scientific conferences where I was able to present my work and receive valuable feedback from the scientific community. I wholeheartedly acknowledge the support of Mr. Antonios Lazoglou who has provided me with seasonal occupation for almost a decade, and has been a wise advisor whose work ethics have been a source of inspiration.

The immense moral support of my family and close friends was vital for completing this thesis. I wholeheartedly thank my family for their understanding when work hours got too long or visits could not be paid due to pressing deadlines, for celebrating with me when good news arrived and for being patient with me through the hard times. I owe a lot to my best friends as well, and especially to Lemonia Dervisi for being the best listener of my endless complaints. Special thanks go to Despina Felekidou for our long and heated discussions that always helped me solve problems and move a step forward. Nevertheless, this adventure would be much shorter --and unsuccessful-- if it was not for my partner, Elena. I am immensely indebted to her for being a caring partner who has been next to me since Day One and who had faith in me even when I had lost it.

Through the course of these five years, I had the opportunity to meet, engage and work with several colleagues. Each has supported me in a different manner through this journey. I would like to acknowledge guidance and support provided by the faculty members of the Department of Economics in the University of Macedonia. I am also grateful to the Professors acting as Head of the Department in these five years for acting upon every issue I encountered.

I would like to also express my gratitude for the support provided by my fellow PhD candidates. I owe special thanks to Thanasis Bouzidis for his sharp comments on my ideas and his moral support in the hard times of this adventure and to Stavros Kourtzidis for being an influential research partner providing fresh perspectives on several of my ideas. In addition, I graciously thank Ilias Livanos for giving me the opportunity to work in Cedefop and collaborate with him ever since. Furthermore, I am heavily indebted to the Professors served as my Thesis advisors, Prof. Emeritus Emmanuel Thanassoulis, Prof. Nikolaos Tzeremes and Prof. Andreas Georgiou, for providing valuable comments and insights to further improve this Thesis.

Nevertheless, this thesis would not have been completed if it was not for the support and aid of my supervisor, Prof. Giannis Karagiannis. I am immensely grateful to him for selecting me as a PhD student and providing me with the best possible guidance in order to realize my potential. His support, both moral and scientific, has been massive and I wholeheartedly thank him for working endlessly these five years to forge a researcher out of me. His work ethic and scientific proficiency have inspired and affected me in a multifaceted way and collaborating with him provided me with invaluable experience. I am indebted to him for being a true mentor and I am thrilled for the opportunity to collaborate with and learn from him.

Table of Contents

List of Tables

List of Figures

Prologue

The completion of this thesis was partially funded by the Greek General Secretariat for Research and Technology (GSRT) and the Hellenic Foundation for Research and Innovation (HFRI) under the program “1st Call for H.F.R.I. Scholarships for PhD Candidates” (Grant Agreement no. 698).

Chapter two was presented in the 5th International Conference on Applied Theory, Macro and Empirical Finance (AMEF 2019) held on April 22-23, 2019 in the University of Macedonia, Thessaloniki, Greece under the title “A VEA approach for gauging effectiveness”. It appears in the Discussion Paper Series of the Department of Economics of the University of Macedonia as Discussion Paper No. 3/2020. It is published as “Ravanos, P., & Karagiannis, G. (2021). Using VEA to assess effectiveness in the development of human capabilities. *Economic Change and Restructuring*, 54, 75-99. https://doi.org/10.1007/s10644-020-09276-7 ”.

Chapter three was presented in the 10th North American Productivity Workshop (NAPW) held on June 12-15, 2018 in Miami Business School, Miami, USA under the title “A Value Efficiency Analysis Benefit-of-the-Doubt model”. It is published as “Ravanos, P., & Karagiannis, G. (2021). A VEA Benefit-of-the-Doubt Model for the HDI. *Social Indicators Research*, 155, 27-46. https://doi.org/10.1007/s11205-020-02589-0 “.

Chapter four was presented in the International Conference in Data Envelopment Analysis (DEA40) held on April 14-16, 2018 in Aston Business School, Birmingham, UK under the title “VEA: In search of MPS”. It appears in the Discussion Paper Series of the Department of Economics of the University of Macedonia as Discussion Paper No. 5/2022. It is published as “Ravanos, P., & Karagiannis, G. (2022). In Search for the Most Preferred Solution in Value Efficiency Analysis. *Journal of Productivity Analysis*, *58*, 203-220. https://doi.org/10.1007/s11123-022-00645-0 “.

Chapter five was presented in the 30th European Conference on Operational Research (EURO) held on June 23-26, 2019 in the University College Dublin (UCD), Dublin, Ireland under the title “VEA and Targeted Benevolence cross-efficiency”. It is currently submitted for publication in a peer reviewed journal.

Chapters six and seven were presented in the 16th European Workshop on Efficiency and Productivity Analysis (EWEPA), held on June 10-13, 2019 in the Senate House, London, UK under the auspices of Loughborough University under the joint title “'Exploring the relationship between VEA and types of weight restrictions”. Chapter six appears in the Discussion Paper Series of the Department of Economics of the University of Macedonia as Discussion Paper No. 4/2022 and it is published as “Ravanos, P., & Karagiannis, G. (2022). On VEA, production trade-offs and weight restrictions. *Journal of the Operational Research Society*, (in press). https://doi.org/10.1080/01605682.2022.2128909.” Chapter seven is currently submitted for publication in a peer reviewed journal.

CHAPTER 1

Introduction

1.1. Setting the stage

There are several occasions in the assessment of Decision Making Units' (DMUs) performance, in which it is desired or necessary to incorporate external views and preferences in the evaluation process. This need frequently arises in the public sector, where a Decision Maker (DM), namely a social planner, regulator or supervising agency coordinating the DMUs' operation, aims to monitor their performance and redirect them towards a desired or mandated trajectory. Examples of such public and centrally coordinated groups of DMUs include, but are not limited to, hospitals, education institutions, research centers, as well as large infrastructure industries benefiting from a natural monopoly such as water, electricity and gas networks. In addition, the performance of privately owned entities such as networks of bank branches or retail stores is often assessed by central management with regards to stated organizational goals. These assessments aim to limit the occurrence of dysfunctional incentives and strategic conflict, where the behavior of DMUs is inconsistent with one that best supports overall organizational goals (Epstein and Henderson, 1989).

In the aforementioned cases, it is desired to augment the assessment models with information on holding views over the types of the relatively best-performing entities. When the DMUs have limited control over their resources, the results of such assessments can be used for the redistribution of personnel or intangible inputs, while in the case of more autonomous DMUs the evaluation can be a means to incentivize them towards meeting certain goals.

Popular notions of performance evaluation in which preferences and views are frequently incorporated include, but are not limited to, the assessment of technical efficiency and effectiveness, cost or revenue efficiency and cross efficiency. A DMU's technical efficiency reflects the extent to which it produces the maximum possible bundle of outputs given a bundle of resources and the available technology or, alternatively, the extent to which a particular bundle of outputs is produced using the least possible resources. In essence, technical efficiency measures the extent to which the assessed DMUs are "doing things the right way" (Cooper *et al.,* 2007a, p. 66). Effectiveness is a notion related to that of technical efficiency, in which the ability of DMUs to achieve desired goals or "do the right things" is assessed. These goals may reflect several kinds of non-monetary objectives, such as directions or legislation set out by management or supervising agencies, but also economic objectives such as cost minimization and revenue maximization. In cross efficiency evaluation, the performance of each DMU is assessed relative to that of its "peers" or the "reference" DMUs. In particular, each DMU evaluates the remaining DMUs based on its own "value system", i.e., individual preferences on what constitutes good performance. This is both desired and necessary in cases of group decision making in which transparency matters considerably for stakeholders. These include, among others, budget allocation in multinational companies and international organizations and the assessment of public institutions such as schools or hospitals.

1.2. Data Envelopment Analysis, its extensions, and Value Efficiency Analysis

Data Envelopment Analysis (DEA) is one of the estimation methods used in applied performance assessment. In DEA, the performance of each DMU is expressed as a ratio of the weighted sum of its outputs to the weighted sum of its inputs, in which the vector of input/output weights are selected by the DMU so as to present itself in the best-possible light. The DMUs are assessed with regards to an envelope formed by those DMUs for which the ratio of the weighted sum of outputs to the weighted sum of inputs is the maximum possible across the sample.

Conventional DEA models measure radial technical efficiency (Charnes *et al.,* 1978; Banker *et al.,* 1984). Extensions of these models have been developed for several other performance evaluation cases, a fair share of which augment DEA models with additional information regarding market prices and managerial preferences. In particular, revenue, cost or profit efficiency is assessed by DEA models when price data are available (see, e.g., Färe *et al.,* 1985). Also, additional restrictions on the input/output weighs have been appended in conventional DEA models to accommodate partial information on prices, stated preferences over the relative importance of inputs and/or outputs, and DM perceptions about "good" and "bad" performing DMUs. These may take several forms of linear inequality restrictions on the input/output weights (see., e.g., Allen *et al.,* 1997) and their dual production trade-offs (Podinovski, 2004) which are special cases of the more general cone-ratio DEA (CR-DEA, see Charnes *et al.*, 1989), where the feasible input/output weight vectors for the evaluated DMUs are restricted within suitably defined cones containing selected sets of input/output weight vectors. The addition of such weight restrictions results in efficiency scores which are lower than or equal compared to those of the conventional DEA model. As such, the extent to which the performance of a particular DMU is aligned with the preferences reflected in the additional restrictions can be assessed by examining the differences between the efficiency scores of the conventional and the weight restricted DEA model.

Effectiveness is also assessed through DEA by means of several models, namely by incorporating additional restrictions in conventional DEA models (see, e.g., Asmild *et al.*, 2007), two-stage processes involving the conversion of inputs to outputs and that of outputs to outcomes, which reflect higher goals selected by the DM such as peace and sovereignty or some form of behavioral objectives (Førsund, 2017), and pure output DEA models (Prieto and Zofio, 2001). Pure output DEA models, along with pure input models (see Karagiannis, 2021 for a review) are special cases of DEA models, in which only inputs or outputs are considered. They have been extensively used, among others, for the construction of composite indicators (see Cherchye *et al.*, 2007a).

Furthermore, cross efficiency (see Sexton *et al.*, 1986) is a methodology developed for peer appraisal assessment by means of DEA. In that, each evaluated DMU is the "reference" DMU in turn, namely assesses all other DMUs as well as itself by means of its own optimal vector of input/output weights. This results into a multitude of efficiency scores for each DMU that form the cross efficiency matrix. Based on that, DMs can identify all-around good-performers which should serve as role models for the rest of DMUs and obtain a --frequently complete-- ranking of the DMUs by aggregating the cross efficiency scores for each of them. Extensions of cross efficiency models use secondary objectives to account for the existence of multiple optimal vectors of weights for the efficient or inefficient DMUs (see, e.g., Doyle and Green, 1994).

The focus of this thesis is Value Efficiency Analysis (VEA) (Halme *et al.*, 1999), which is an alternative method for incorporating preferences in DEA. In VEA the views of a social planner, regulator, or manager are expressed through a pseudoconcave value function (i.e., an indifference curve), which is assumed to be strictly increasing in the outputs and strictly decreasing in the inputs. This function may be related to some organizational objective, such as a cost minimization or profit maximization, but it might also reveal other preferences than those related with prices (Thanassoulis *et al.*, 2008, p. 73). Thus, DM preferences in VEA are elicited by a means commonly used in Multi Criteria Decision Analysis (MCDA), namely by incorporating DM information on the desirable quantities for the inputs and outputs of the assessed DMUs. More specifically, it is assumed that the DM's value function is tangent (i.e., is maximized) to the DEA efficient frontier at a point which reflects the most desirable input/output structure by DM's view and is referred to as the Most Preferred Solution (MPS). The MPS is a non-dominated (i.e., strongly DEA-efficient) DMU or a combination of DMUs located on the strongly efficient DEA frontier.[1] Depending on the evaluation setting, the MPS may represent the structure according to which the DM wishes to reorganize a portion of a country's public sector (e.g., the wastewater management network) or it might be viewed as a mentor, i.e., an example-to-follow, for the other DMUs within a private organization. DMs reflect their preferences in VEA by choosing a strongly DEA-efficient DMU or a combination of such DMUs to be the MPS.

The DM preferences are then incorporated in the DEA model by restricting the choice of the optimal vector of input/output weights for each evaluated DMU only among those input/output vectors that are optimal in the DEA model for the DMU or the DMUs that constitute the MPS. In essence, this means that the marginal rates of substitution of inputs or transformation of outputs imposed on the evaluated DMUs are those observed on the DEA frontier in the neighborhood of the MPS. This essentially defines a range of desirable input/output bundles, namely those that have at least one optimal vector of input/output weights in common with the MPS. This range may be viewed as the DM's "margin of error", in the sense that the DMUs contained in it are

[1] The strongly efficient DEA frontier is formed by DEA-efficient DMUs which are not associated with nonzero slacks, i.e., input excesses or output shortfalls, and their linear combinations.

those for which the input/output bundle or mix diverges from the most preferred one (i.e., that of the MPS) to an extent that is considered tolerable by the DM. Such DMUs receive a VEA efficiency score that is equal to their respective DEA score. For the remaining DMUs VEA efficiency scores are lower that their corresponding DEA ones, and the more a DMUs' input/output structure diverges from that of the MPS, the lower is its VEA score. The VEA efficient frontier is the lower envelope of the extended efficient facets of the DEA frontier that intercept at the MPS, and in the VEA scores are optimistic approximations of the scores which would be obtained if an explicit functional form was available for the DM's value function.

VEA provides a useful alternative for incorporating preferences in the performance assessment of DMUs for a number of reasons: *First,* choosing the MPS is a relatively less demanding process for DMs compared to including additional restrictions on the weights. DMs are generally more keen on choosing desirable values for inputs and outputs rather than weight bounds (Korhonen *et al.,* 2002), and this choice can be done without requiring familiarity with the DEA method. This can limit considerably the possibility of DMs expressing their preferences incorrectly and potentially giving rise to misleading evaluation results. *Second,* the range of desirable input/output bundles defined by the MPS choice is less strict compared to e.g., setting explicit targets for DMUs on the efficient frontier, as it allows for tolerated divergences from the most preferred bundle. This is useful in cases where management aims to limit the ability of DMUs in setting their own priorities and identify those falling considerably short of achieving organizational goals or stated norms. Policies can then be designed based on these findings to redirect the DMUs' operation towards the desired trajectory. *Third,* incorporating the MPS in DEA requires only slight modifications in the multiplier and envelopment form of the DEA model.

1.3. Motivation

The aim of this thesis is a more detailed theoretical and empirical examination of VEA. This is motivated by the fact that, despite recent theoretical advancements in the relevant literature, VEA has not been studied to the same extent as DEA (see Table 1.1). Also, the use of VEA in empirical applications is relatively scarce, despite the abundance of performance evaluation cases and operational research problems in which the incorporation of managerial preferences is desired or necessary.

Table 1.1: Uses of DEA and VEA models for performance evaluation

	Data Envelopment Analysis (DEA)	Value Efficiency Analysis (VEA)
introduction	Charnes *et al.* (1978), Banker *et al.* (1984)	Halme *et al.* (1999)
	performance evaluation instances	
incorporating DM preferences when prices are not known	weight restrictions (Thanassoulis and Dyson, 1988)	MPS (Korhonen *et al.*, 2002)
	production trade-offs (Podinovski, 2004)	*Chapter 6 (this thesis)*
	cone-ratio DEA (Charnes *et al.*, 1989)	*Chapter 7 (this thesis)*
	reviews: Allen *et al.* (1997), Thanassoulis *et al.* (2004; 2008)	*Chapter 4 (this thesis)*
cross efficiency	(Sexton *et al.*, 1986; Doyle and Green, 1994)	*Chapter 5 (this thesis)*
effectiveness assessment	Asmild *et al.* (2007), Prieto and Zofio (2001), Førsund (2017)	*Chapter 2 (this thesis)*
composite indicators construction	pure input and pure output DEA models (Van Puyenbroeck, 2018, Karagiannis, 2021) Benefit-of-the Doubt (Cherchye *et al.* (2007a)	*Chapter 3 (this thesis)*

More specifically, studies following the introduction of VEA in Halme *et al.* (1999) examined the potential of obtaining VEA scores which are better approximations of the scores that could be obtained if an explicit functional form was available for the DM's value function, and their interpretation in terms of value differences between the assessed DMU and the MPS (see Joro *et al.*, 2003; Korhonen and Syrjanen, 2005). Furthermore, VEA has been applied for the evaluation of hospital departments (Halme and Korhonen, 2000), higher education institutions (Korhonen *et al.*, 2001), local governments (Marshall and Shortle, 2005) and banks (Eskelinen *et al.*, 2014).

On the other hand, the use of VEA models for several performance evaluation instances for which it seems particularly suitable, such as cross efficiency and effectiveness assessment, has to the best of our knowledge not been examined. Furthermore, the relation of the VEA model with DEA models that introduce additional restrictions on the input/output weights has not been extensively examined, despite the relevant literature recognizes that such relations might exist (see e.g., Sarrico and Dyson, 2004; Kao and Hung, 2005). Korhonen *et al.* (2001) incorporated additional weight restrictions in VEA models, but did not engage in relating VEA and weight restricted DEA models. Moreover, extensive research has been carried out regarding different forms of restrictions on the input/output weights (see, e.g., Angulo-Meza and

Estellita-Lins, 2002 and Thanassoulis *et al.*, 2004;2008 for reviews), the preferences these may represent, their economic interpretation (if any) and the effects of using various forms of weight restrictions on efficiency scores (see Allen *et al.*, 1997). On the other hand, to the best of our knowledge no study has examined in detail the variety of existing suggestions for choosing the MPS in VEA, their economic interpretation, as well as the effects (if any) of using alternative MPSs on VEA efficiency scores.[2] Also, pure input or output DEA models have been studied in detail (see e.g., Van Puyenbroeck, 2018) and have had several empirical uses (see Karagiannis, 2021), the most popular of which being the construction of composite indicators. For the latter purpose, a DEA model known as the Benefit-Of-The-Doubt (BoD) model (Cherchye *et al.*, 2007a) is most frequently used. Corresponding VEA models to these have, to the best of our knowledge, not yet been developed nor used to construct composite indicators of performance.

1.4. Contribution

This thesis is divided in two parts and contributes to the VEA literature by (i) examining the features of VEA in more detail, and (ii) presenting innovative empirical applications of VEA models. More specifically, each of the six chapters that follow investigates in detail an issue among those identified in the previous section which was not up to date addressed in the VEA literature (see Table 1.1). The first part, namely chapters two to four, consists of three empirical essays. In the first of them, we use VEA for the assessment of effectiveness. In particular, we encapsulate the DMs views about the DMUs that are "doing the right things" in the choice of the most desirable input/output bundle, i.e., the MPS. Then, the scores obtained from the respective VEA model are estimates of the DMUs' effectiveness. These are decomposed into an efficiency component capturing the extent of 'doing things right' and a mix component capturing the relative distance of the assessed DMUs' input/output bundle from the DM's range of desirable bundles, as the latter is defined by means of the MPS choice. The mix component is residually estimated as the ratio of DEA and VEA efficiency scores. We use this approach to provide an innovative application of VEA, namely assess the

[2] Korhonen *et al.* (2002) suggest some alternatives for choosing the MPS in VEA, but do not investigate any economic rationales related to these.

effectiveness of countries in utilizing their economic prosperity (proxied by their income) to further develop their citizens' social prosperity or human capabilities (proxied by achievements in terms of health and education) using UN data for the year 2015.

In the second empirical essay, we i) extend VEA towards pure input DEA models and ii) use it as a means to incorporate DM preferences in the construction of composite indicators, by combining the BoD model with VEA. The newly proposed VEA-BoD model is then used to re-estimate the United Nations Human Development Index (HDI). In this application, the MPS is selected based on the notion of uniformity. This reflects an objective and normative overall goal, namely the equal prioritization among the considered indicators (i.e., the achievements in terms of income, health, and education), and means that the DM prefers countries with a relatively balanced prioritization among health, education and income more compared to those with unbalanced achievements. The former would be promoted as peers for improving human capabilities in the latter.

In the third empirical essay, we first review various suggestions made for choosing the MPS in VEA and the preferences these might reflect. We then propose four new, which rely respectively on the relative position of frontier DMUs, the Most Productive Scale size (MPSS), the Average Production Unit (APU), and common vectors of input/output weights. These reflect overall organizational goals such as the pursuit of scale economies and the maximization of structural efficiency, or the need to assess DMUs against common standards because of limited control over the resources allocated to them or autonomy in setting their own priorities. Using a dataset of Greek cotton farms, we then provide comparative empirical results that illustrate the implications of using different MPS choices for the VEA efficiency scores.

Part II, namely chapters five to seven, consists of three theoretical essays. In the first of them, we examine the potential relation between VEA and cross efficiency. In particular, we show that the Targetted Benevolence (TB, see Oral *et al.*, 1991) cross efficiency model, is equivalent to the VEA model, provided that the "reference" DMU, i.e., the one whose optimal multipliers are used to evaluate all other DMUs, in the TB cross efficiency model, if it is an efficient one, or its radial projection on the DEA frontier if it is inefficient, is used as the MPS in the VEA model. The TB model is one among those adopting a secondary objective to account for the possibility of multiple optimal weight vectors in DEA for each “reference” DMU. According to this objective, the evaluated DMU is allowed to select the weight vector that maximizes its cross efficiency score, among those that are optimal in DEA for the “reference” DMU. The identified equivalence implies that the TB cross efficiency scores can be obtained by means of the envelopment form of VEA –which frequently involves fewer constraints compared to its dual multiplier form-- and allows for the estimation of the TB cross efficiency matrix using less linear models.

In the second theoretical essay, we explore the relationship between VEA and DEA models with weight restrictions and their dual production trade-offs. In particular, we show that the VEA model is equivalent to a DEA model with production trade-offs as long as the trade-off coefficient vectors are equal to (i) the negative of the input and output quantities of the DMUs constituting the MPS in VEA, under constant returns to scale, and (ii) the deviations of all evaluated DMUs' input and output quantities from those of the DMUs chosen as the MPS, irrespectively of the returns-to-scale assumption. These production trade-offs are in both cases dual to Type II assurance region weight restrictions (see Thompson *et al.,* 1990). In addition, show that a similar equivalence holds between pure output or input VEA models and DEA models with production trade-offs if the above trade-offs are considered only for the inputs or the outputs. These findings allow for an alternative interpretation of the VEA efficiency scores and the scores of DEA models with production trade-offs and their dual weight restrictions.

In the third theoretical essay, we relate VEA with CR- DEA models incorporating preferences regarding efficient DMUs that management views as examples to follow (*model* DMUs) for the remaining DMUs. In particular, we show that as long as the *model* DMUs chosen in CR-DEA are those that constitute the MPS in VEA, the VEA efficiency scores are i) equal to those obtained from a CR-DEA model in which the cone of feasbile weight vectors is specified as the intersection of the sets containing the weight vectors that are optimal in DEA for each *model* DMU, ii) provide a lower bound to the scores obtained from a CR-DEA model in which the cone of feasible weights is given as the union of the sets containing the optimal weight vectors for each *model* DMU, and iii) constitute an upper bound for the efficiency scores of a Fully-Dimensional-Efficient-Facet (FDEF) CR-DEA model in which the cone of feasible weights contains only those weight vectors that are jointly optimal in DEA for all the model DMUs as well as strictly positive. These findings enable the estimation or the approximation of CR-DEA efficiency scores by means of VEA models. These are less computationally demanding as they do not require to *a priori* identify the cone of feasible weight vectors, as is the case in the CR-DEA models.

Part I: Empirical essays

CHAPTER 2

Using VEA to assess effectiveness in the development of human capabilities

2.1. Introduction

Efficiency and effectiveness are two distinct but related notions of performance evaluation. Efficiency measures the extent to which a decision-making unit (DMU) 'does things the right way', namely whether it produces the maximum possible outputs from given inputs or uses the minimum possible inputs to produce a given bundle of outputs. Effectiveness, on the other hand, measures the ability of DMUs to state and achieve desired goals (Cooper *et al.*, 2007a, p. 66) i.e., it examines the question of doing the right things. The goals or ‘’right things’’ reflect behavioral or organizational objectives of DMUs or their supervising agency, which can be either monetary or non-monetary. The former refers to economic objectives, such as cost minimization or revenue maximization, the extent of which can be assessed as long as price data are available while the latter refers to managerial preferences about the production process itself as well as targets to be achieved by the constituent DMUs (see e.g., Asmild *et al.*, 2007).

There are four different approaches in the literature to assess effectiveness. The first of them uses a two-stage process (see Førsund (2017) and the references therein) where at the first stage the efficiency of DMUs is assessed by focusing on the process of converting inputs to outputs. Effectiveness, assessed at the second stage, reflects the ability of DMUs to convert outputs to outcomes.[3] Conventional DEA models are used

[3] For example, in assessing effectiveness in transport industry, inputs usually refer to number of vehicles, fuels and labor, outputs refer to the produced transport capacity (e.g., seat-miles) while outcomes refer

in both stages and the behavioral objectives are expressed through the selection of outcomes. Recently, Førsund (2017) and Hanson (2018) provide innovative refinements of this approach, especially suitable for application related to public sector. In the second, additional constraints reflecting behavioral objectives (see Asmild *et al.*, 2007) are introduced into conventional DEA model. If these are related to economic objectives, such as cost minimization or revenue maximization, then effectiveness coincides with the notion of overall efficiency. If the behavioral objectives reflect managerial goals, then we need restrictions on the input and/or output multipliers to incorporate them into the conventional DEA model.[4] The resulting model "evaluate(s) both the technical inefficiency that arises from not fully exploiting production possibilities and the inefficiency due either to lack of fulfillment of managerial goals or to the departure from the specified value system of the inputs and outputs'' (Cooper *et al.*, 2011, p. 101). In the third (see Prieto and Zofio, 2001), effectiveness is estimated by means of pure output DEA models. Here the goals of DMUs are considered as given and we concentrate in estimating the extent to which they are achieved regardless of the amount of resources that might be needed to provide them. This follows the idea of the Koopmans' 'helmsman' that attempts to steer all the outputs towards their maximum levels without considering the inputs used (see Lovell *et al.*, 1995). In the fourth approach, effectiveness is related to the distance of DMUs from target points on the existing DEA efficiency frontier (see Golany *et al.*, 1993). Such targets may minimize the distance of DMUs from the DEA frontier, or maximize the outputs of a DMU under a fixed resource allocation.

In this chapter we propose an alternative way to incorporate behavioral objectives into conventional DEA in order to assess effectiveness. This is based on Value Efficiency Analysis (VEA), where the behavior objectives reflect the preferences of a Decision Maker or supervising agency, which provides the necessary information regarding the right things to do by simply choosing a "model" DMU, instead of having

---

to the extent that produced capacity is consumed by customers (e.g. passenger-kilometer and ton-kilometer) (Yu and Lin, 2008). Another example provided by Hanson (2018) is the assessment of military forces effectiveness, where inputs refer to resources such as personnel and equipment, outputs to countable services or goods such as the number of military units and the quality of their training, and outcomes to country-wide valued states and public goods such as peace, sovereignty or freedom.

[4] Different types of weight restrictions may be used, such as absolute or relative bounds on the multiplier weights, resulting in a set of equal, common across DMUs, or DMU-specific input and output multipliers.

to choose weight restrictions by means of absolute or relative bounds. According to Korhonen *et al.* (2001), this is an easier method to reflect preferences for the Decision Maker, who is more keen on picking a "model" DMU rather than engaging to the task of choosing weight restriction bounds, which is a more technical issue. The "model" DMU reflects the most-preferred solution (MPS) from the Decision Maker's point of view and is then used as a global benchmark that determines a range of preferred input and output bundles which comply with her view of 'doing the right things' and provide the base for estimating effectiveness. Efficiency is estimated by means of the conventional DEA model and the two are related by a mix component. The latter serves as a measure of the closeness of the actual input/output bundle of DMUs to the most-preferred input/output bundle and can aid analysts and Decision Makers identify DMUs with effective operating bundles (which can serve as models) and DMUs which need a restructuring in order to comply with managerial preferences, social norms or supervising agency directives.

We use the proposed approach to provide estimates of countries' ability to efficiently and effectively utilize their economic prosperity to enrich the lives of their citizens using 2015 UNDP data. We rely on Sen's capability approach that views humans as the ultimate ends of the process of economic prosperity and development itself as an expansion of their capabilities, in contrast with the Human Capital approach which views humans as the primary means of economic development. Our empirical models operationalize the differential treatment of income on the capability approach as a means to a number of important ends, rather than an end in itself (Anand and Sen, 2000; Klugman *et al.*, 2011). More specifically, we follow the DEA social efficiency model (see Despotis, 2005a,b; Mariano and Rebelatto, 2014) and use income as an input reflecting economic prosperity, with life expectancy, mean and expected years of schooling as the outputs reflecting social prosperity. The empirical results help classifying countries into groups displaying high and balanced social prosperity provision (Leaders), countries with a balanced bundle but relatively lower achievements, which could use their economic prosperity more efficiently (mix efficient), countries with high but unbalanced provision of health and education (Efficient), which could benefit moving towards a more balanced social prosperity bundle, and finally Laggard countries with both low and unbalanced achievements. Such results can prove useful to both national policy-makers to reshape national policies as well as to international organizations to better allocate development or international aid funds.

The rest of the chapter is organized as follows: In the next section, we introduce VEA and explain how it can be used to estimate effectiveness. The empirical application is presented in the third section, while concluding remarks follow in the last section.

2.2. Effectiveness assessment with VEA

VEA, developed by Halme *et al*. (1999), is a performance evaluation method that takes into account Decision Maker preferences about managerial goals by means of a linear value function (i.e., an indifference curve) that become tangent to the DEA efficient frontier at the point of the most-preferred solution (MPS).[5] This point reflects the Decision Maker's choice of a virtual or real *non-dominated* (i.e., DEA-efficient) DMU as a model DMU. Then, the VEA frontier is constructed by extending towards the axes the hyperplanes of the DEA efficient facets intercepting at the MPS. As DEA facets are generated by extreme-efficient DMUs, the MPS will in essence be either a single extreme-efficient DMU or a combination of extreme-efficient DMUs that are *jointly* efficient, in the sense that they generate at least one common facet.[6] In Figure 2.1(a), by choosing for example DMU B as the MPS, the two efficient facets AB and BC are extended towards the axes, creating the VEA frontier (the blue kinked line). The range of preferred input/output bundles is given between rays OA and OC. All DMUs producing within the preferred range receive a VEA score that is equal to their respective DEA score whereas DMUs producing outside of the preferred range are penalized by receiving VEA scores less than their corresponding DEA scores.

---

[5] A detailed presentation of VEA can be found in Joro and Korhonen (2015).

[6] In Charnes *et al.* (1991a) the DMUs with a DEA efficiency score of one are classified into three categories: (a) extreme-efficient DMUs (E) that reside at a point of the convex DEA frontier where more than one facets intercept, (b) non-extreme-efficient DMUs (E'), namely DMUs located on the interior of a facet, and (c) weakly-efficient DMUs (F) that have at least one positive optimal value for an input or output slack. If the DM chooses a DEA-inefficient or weakly-efficient (i.e., a dominated) DMU, or a non-extreme-efficient DMU as the MPS, then the combination of the extreme-efficient DMUs that are identified as its peers in DEA can be used as the MPS instead (see e.g., Halme *et al.*, 1999). The use of the peers of the DEA-inefficient DMU rather than its radial projection in the DEA frontier is advocated, as the latter might be associated with input and/or output slacks and thus may not be a non-dominated DMU. The same is the case for weakly efficient DMUs.

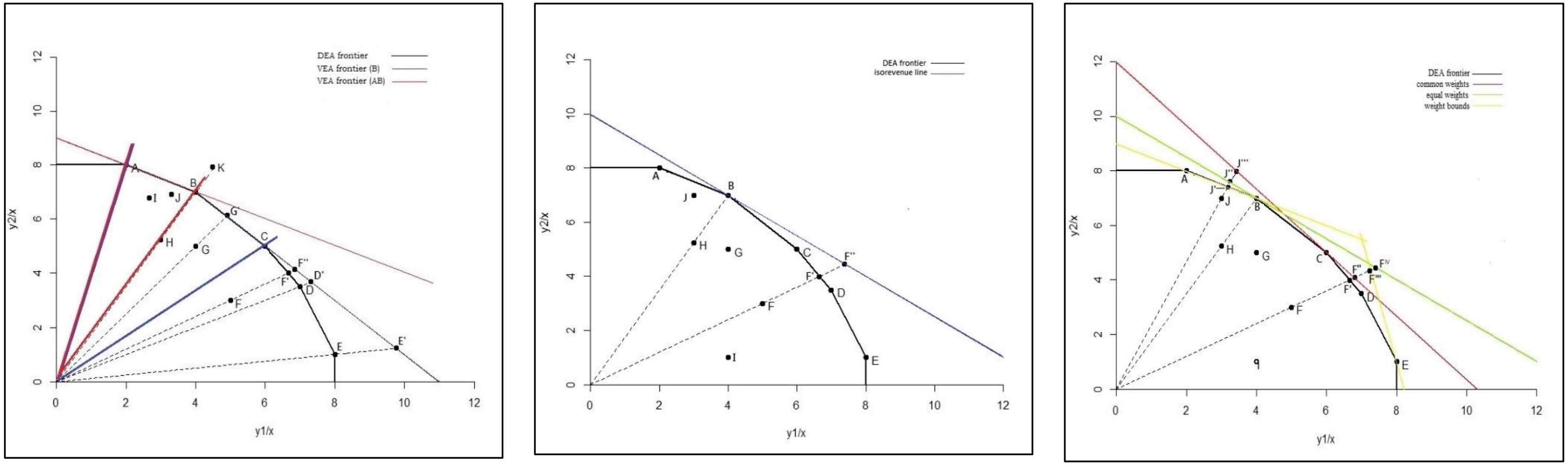
Choosing the model DMU is a crucial step in VEA, as the chosen MPS affects the preferred range of input and output bundles and consequently, the resulting VEA scores.[7] Although Decision Makers are more inclined to simply choosing a DMU from the set of DEA-efficient ones (see Korhonen *et al.*, 2001) such as DMU B in Figure 6.1(a), they also have the freedom to choose an MPS that is DEA-inefficient (Korhonen *et al.*, 2002) or propose instead an artificially constructed MPS that may or may not be efficient. In the latter case, the chosen DMUs are first projected on to the DEA efficient frontier and then their peers are instead used as the MPS. In Figure 2.1(a), consider for example DMUs H and G, which are DEA-inefficient and K which is an artificial DMU. H and K are (for ease of presentation) both projected into point B of the DEA efficient frontier. Then, their use as MPS implies instead the use of DMU B, their peer, as the MPS. In a similar fashion, the use of G as the MPS, which is projected on the efficient facet BC, implies the joint use of DMUs B and C as the MPS. Note that projecting an artificial DMU such as K on the DEA frontier requires solving a superefficiency DEA model (see Andersen and Petersen, 1993).

A variable-returns-to-scale formulation of the VEA model, in its multiplier form is given as (Halme and Korhonen, 2000):

$$
\begin{aligned}
\min_{v_i^o, u_j^o, u^o} \varphi_{VEA}^o = & \sum_{i=1}^{I} v_i^o x_i^o - u^o \\
s.t. \quad & -\sum_{j=1}^{J} u_j^o y_j^k + \sum_{i=1}^{I} v_i^o x_i^k - u^k \geq 0 \quad && k = 1, \dots, K, k \neq r \\
& -\sum_{j=1}^{J} u_j^o y_j^r + \sum_{i=1}^{I} v_i^o x_i^r - u^r = 0 \quad && r = 1, \dots, R \\
& \sum_{j=1}^{J} u_j^o y_j^o = 1 \\
& v_i^o \geq 0 && i = 1, \dots, I \\
& u_j^o \geq 0 && j = 1, \dots, J \\
& u^o \, free
\end{aligned}
\tag{2.1}
$$

[7] See Korhonen *et al*. (2002) for more details regarding the several alternatives underlying the choice of the model DMU.

Figure 2.1:Effectiveness assessment based on different approaches



Panel (a): Reflecting managerial preferences through VEA

Panel (b): Known prices, effectiveness coincides with overall efficiency

Panel (c): Approximating prices with weight restrictions where x and y refer to input and output quantities, φ_{VEA} to the (inverse) of the VEA efficiency score, v and u are parameters to be estimated, k is used to index DMUs ($k = 1, \dots, o, \dots, K$), i is used to index inputs ($i = 1, \dots, I$), j is used to index outputs ($j = 1, \dots, J$) and $r = 1, \dots, R$ refers to the DMUs chosen as the MPS. The above formulation is only slightly different from the conventional DEA model: the restriction corresponding to the MPS is turned from inequality to equality. This affects the optimal values of the input/output multipliers and essentially determines the range of preferred input/output bundles such as those between the rays OA and OB in Figure 2.1(a).

In this chapter we use VEA to estimate effectiveness and compare it to efficiency which is estimated by means of DEA, in the sense that it reflects the behavioral objectives of a Decision Maker or supervising agency, which provides the necessary information regarding the "right things" by means of a "model" DMU, that determines the MPS and the range of preferred input/output bundles. Then, the distance of a DMU from the VEA frontier is used to measure effectiveness while its distance from the DEA frontier is used to measure efficiency. Consider for example DMU F in Figure 2.1(a) where $\frac{OF}{OF''}$ measures the extent to which DMU F "does the right things" while $\frac{OF}{OF'}$ measures the extent to which the DMU F does "things the right way". From that we see that effectiveness and efficiency are related to each other as follows:

$$\underbrace{\frac{1}{\varphi_{VEA}^{F}}}_{effectiveness} = \underbrace{\frac{1}{\varphi_{DEA}^{F}}}_{efficiency} \times \underbrace{\frac{OF'}{OF''}}_{\substack{mix\\component}} \qquad (2.2)$$

The second term in (2.2), i.e., the mix component, reflects the extent to which the DMU operates inside the given range of preferred input/output bundles and it is given by the ratio of the effectiveness to the efficiency score, taking values within the [0,1] range.[8,9] When a DMU operates within the preferred range of bundles, effectiveness and

[8] The mix component is similar (but not the same) to Filippetti and Peyrache (2011) compositional index and to the Li and Zhao (2015) dimension mix index, with the main difference being that their non-DEA frontiers result from a set of common (across DMUs) weights which in terms of Figure 2.1 implies a linear frontier; see Figure 2.1(c).
[9] Effectiveness scores are never higher than efficiency scores, as the VEA frontier envelops the DEA frontier.

efficiency scores coincide (for example DMU G Figure 2.1(a)) and the mix component is equal to one. This in general indicates that the particular DMU operates in a manner that is in line with the behavioral objectives set out by the manager or the supervising agency but has different implications when it occurs for efficient and inefficient DMUs. Inefficient DMUs with a mix component equal to one are on the "right operating path" and their ineffectiveness is caused only by inefficient utilization of inputs to produce outputs (i.e., inefficiency) while efficient DMUs producing within the preferred range of bundles are classified as effective and receive a score of one in all three scores. Such DMUs can serve as examples to follow for the rest of the group. On the other hand, when production takes place outside of the preferred range (see DMU F in Figure 2.1(a)) effectiveness is lower than efficiency and the mix component is lower than unity. This indicates that a DMU has diverged from the "right things" norm or mandate and there is a need to change its operating bundle, while if the DMU is also inefficient, additional actions are needed to eliminate technical inefficiencies.

The "right thighs" norm or mandate can include directions set out by the management authorities of a corporation to its branches (e.g. in the case of a bank branch), regulations set out by the government agencies regulating an economic sector (e.g. in the case of financial sector regulations set out by Capital Market Commissions) or the international organizations supporting and monitoring a nation's actions (e.g. in the case of a country being part of the European Monetary System, NATO or the UN). This broad definition highlights the generality of our approach and the fact that it can be applied in several real-world cases.

We can now compare the VEA formulation of effectiveness to those of the second approach referred to in the Introduction, namely that of imposing behavioral (e.g., economic or managerial) objectives. The use of economic objectives is depicted in Figure 2.1(b) and that of managerial objectives by means of weight restrictions in Figure 2.1(c). In Figure 2.1(b) the straight blue line refers to a known output price ratio (i.e., iso-revenue line), which defines a single optimal output bunlde along the ray OB. Effectiveness, which in this case coincides with the notion of overall (revenue) efficiency, of DMU F is given by the ratio $\frac{OF}{OF''}$ while (technical) efficiency is given by the ratio $\frac{OF}{OF'}$. In this case, the mix component, which is given by the ratio $\frac{OF'}{OF''}$, coincides with allocative efficiency. In Figure 2.1(c) we depict different cases of weight restrictions that are used to reflect "the right things to do". The straight green and red lines tangent to the DEA frontier in points B and C correspond respectively to a common (across DMUs) and an equal weight scheme. Both define a single optimal output bundle, although a common-weights scheme reflected into lines AB or BC would define a range of preferred input/output bundles. The broken yellow line corresponds to a form of relative weight bounds, which is similar to VEA frontier in Figure 2.1(a), and both define a preferred range of bundles.[10] Taking DMU F as an example, in Figure 2.1(b) efficiency is defined by the ratio OF/OF^{I}. Effectiveness is defined as the ratio OF/OF^{II} for the common weights scheme, the ratio OF/OF^{III} for the relative weights bounds scheme and the ratio OF/OF^{IV} for the equal weights scheme. This indicates that effectiveness estimates for a DMU may differ when different weighting schemes are used to reflect "the right things to do".[11]

2.3. Estimating effectiveness in the development of human capabilities

2.3.1. Methods and Materials

In this section, using 2015 UNDP data, we employ VEA to estimate the extent to which countries utilize their economic prosperity efficiently and effectively to enhance the development of human capabilities for their citizens, i.e., to increase their nations social prosperity. Social prosperity is considered within the capability approach which focuses on the ability of people to live the lives they have reason to value (Sen, 1999, p. 293) and views development as a process that is "removing restrictions" (Fukunda-Parr, 2003) and "enlarging people's choices" (UNDP, 1990). People themselves are the primary ends of the process of development, in addition to them being the principal

---

[10] VEA can also lead to a common set of weights. If for example both DMUs A and B were chosen as the MPS in Figure 2.1(a), the VEA frontier would extend only facet AB towards the axes, thus creating a common set of weights that nevertheless defines again a range of preferred bundles. The same would occur if the inefficient DMUs I or J were chosen to be the MPS, as for both of them the efficient peers identified by DEA are DMUs A and B.

[11] DMUs may be 'favored' by specific weight restrictions more or less than others, as e.g., DMUs J and F: the former is more (less) favored by the green (red) line of common (equal) weights while the opposite holds for the latter. However, the same holds for effectiveness by means of the VEA model, as some DMUs are favored by the chosen MPS more or less than others: with DMU B as the MPS in Figure 2.1(a), DMU E is ineffective while if DMU D is chosen as the MPS the DMU E would be effective instead.

means of economic production and subsequent economic growth.[12] This differs from the human capital literature that tends to concentrate merely on the role of human beings in augmenting production possibilities, i.e., seeks what people "put into" development. According to Sen (1999, p. 293-295) the latter is a narrower view that tells us nothing about why economic growth is sought in the first place and can fit in the more inclusive perspective of human capabilities, which seeks "what people get from development" (Anand and Sen, 2000a). The two approaches are of course related to each other in a causal way; see Ranis *et al.* (2000) and Suri *et al.* (2011).

Economic prosperity, which is usually reflected through a country's per capita income, is viewed by the capability approach as being merely a means to the ends of human development rather than an end in itself (UNDP, 1990; Anand and Sen, 2000; Fukunda-Parr, 2003; Alkire, 2005; Klugman *et al.*, 2011). Nevertheless, Sen (1993) noted that means of development such as income can indirectly influence the evaluation of human well-being through their effects on variables included in the evaluative space of human well-being (p. 33).[13] This brings forth the question of whether countries are able to efficiently utilize their economic prosperity to enhance the social prosperity of their people. The need to provide an answer to such a question is necessary because, despite the high correlation of income levels with longevity and education outcomes, "this tight relation does not obtain" (Sen, 2003, p. 3). There exist many examples of countries with similar levels of income that achieve very different outcomes in terms of basic capabilities such as being healthy and receiving adequate education (see e.g., Sen, 1983, pp. 753-754 and Sen, 2003, pp. 3-4) and for that reason, Sen (1983, p. 754) noted that "not merely is it the case that economic growth is a means rather that an end, it is also the case that for some important ends it is not a very efficient means either".[14]

This line of reasoning was operationalized within the DEA framework by what is now referred to as DEA social efficiency model (Despotis, 2005a,b; Marianno and

---

[12] The capability approach is the underpinning of the construction of the Human Development Index (HDI), which concentrates in a set of basic and universally valued capabilities-longevity and education as well as gross national income.

[13] "the income of a person can tell us a good deal about her ability to do things that she has reason to value" (Anand and Sen, 2000a, p. 100).

[14] Anand and Sen (2000, p. 101) also referred to outlier countries that are "doing much more to enhance life expectancy than their GNP per capita would suggest". These outlier countries need to be identified and used as benchmarks for other countries.

Rebelatto, 2014) where income is treated as an input and life expectancy and educational attainment as outputs.[15,16]. In the context of the DEA social efficiency model, efficiency does not have a strictly production-oriented meaning, i.e., it does not explicitly refer to producing a given set of outputs with the minimum possible inputs. Instead, a "socially efficient" country is one which manages to provide to its citizens high social prosperity levels given its current economic prosperity levels in a relative sense, i.e., given the achievements in social prosperity of other countries with similar economic prosperity levels. This definition adheres to our earlier definition of efficiency as "doing the right things" and does not include any considerations about the relative composition of health and education indicator levels.

Nevertheless, additional information regarding "the right things to do", i.e., norms about the preferred performance of nations should be considered. Examples include institutional constraints laid down by international bodies, positive or negative externalities pointing towards desirable performance and fairness or social conscience (Golany and Thore, 1997). Such an example may be the intention to simultaneously improve the provision of health and education services. Mishra and Nathan (2018) refer to such a balanced realization of performance as the uniformity axiom and state that it is a desirable property for any index of material well-being and capabilities. Also, from a policy perspective, such a balanced prioritization norm between health and education provision, if followed, would aid the country to exploit possible spillover effects existing between the two.[17] We adopt this equal prioritization norm to define

[15] DEA is a non-parametric methodology for estimating production frontiers and measuring efficiency. Compared to its parametric counterpart, Stochastic Frontier Analysis (SFA), there are advantages and disadvantages. The main advantage of using DEA is that it does not require any information more than input and output quantities, while SFA requires an explicit specification of a functional form for the production function and an explicit distributional assumption for the inefficiency terms. Also, in DEA all deviations from the frontier are readily attributed to inefficiencies, i.e., it does not incorporate stochastic noise in the data as is done by SFA. The latter is a particularly important advantage when additional restrictions are incorporated in the model (as is the case of this chapter), as the extension of the DEA frontier by the extra restrictions (see e.g., Figure 6.1(a)) is not guaranteed to take place in the presence of stochastic noise.

[16] All previous studies using this model assumed variable returns to scale, in order to reflect the diminishing returns as income increases and used an output orientation to gauge efficiency. Output orientation displays a focus towards increasing the current provision of health and education given the resources currently available. It also reflects the views of Ranis *et al.* (2000) and Suri *et al.* (2011) that improving levels of education and health should have priority or at least move together with direct efforts to enhance growth.

[17] Ranis *et al.* (2000, p. 200) offer an example of such a spillover effect, citing studies that provide evidence that "education, especially female, tends to improve infant survival and nutrition".

effectiveness and to choose our “model” country for the VEA model. Thus, a country should not only manage to provide to its citizens high social prosperity levels given its current economic prosperity, but also to equally prioritize between the provision of health and education services. We chose Norway which is the country that ranked 1st in the 2015 version of the UN HDI and it is a good example of balanced prioritization among health and education provision. Norway is a DEA-inefficient country and thus the units comprising its reference set, namely the efficient countries Australia, Switzerland and Hong-Kong, are instead used as MPS in its place.[18]

Our empirical models use the natural logarithm of GNI per capita in 2011 PPP $ as the single input. We consider three outputs, the first of which is life expectancy at birth, a proxy for health provision. To proxy educational attainment, we follow Lozano and Gutiérrez (2008) and Sayed *et al.* (2015) and use the indicators of mean and expected years of schooling as two separate outputs instead of taking their arithmetic average, to better reflect their different focus on the future expectations of education versus the current realizations of it. Such a choice is also grounded on recent statistical results by Canning *et al.* (2013) who found that combining the two variables into a composite causes a substantial loss of information. The data were normalized using the distance-to-the-leader scheme, as suggested by Herrero *et al.* (2012). This normalization scheme retains the unit invariance property for our models, while also leads to normalized values that necessarily lie within the [0,1] range.[19] Descriptive statistics of the model variables are given in Table 2.1.

The decomposition of effectiveness estimates into efficiency and the mix component, as in (2.2), allows the classification of countries into five groups based on their relative ability to provide an increased *as well as* a balanced provision of health and education to their citizens. The “Leaders” group contains those DMUs which score above 99% in both efficiency and the mix component and therefore are considered as effective. The “mix efficient” group contains those DMUs that have a mix component score higher than 99% but an inefficiency score less than 99%. The reverse occurs for “efficient” DMUs which have efficiency scores higher than 99%, but lower mix

---

[18] In terms of Figure 2.1, Norway corresponds to DMU J.

[19] This normalization scheme also avoids the process of truncating normalized values to unity, which is criticized by Lind (2019, p, 410) since it “suggests that human development has an upper limit”.

Table 2.1: Descriptive statistics of model variables

Variable	min	Max	standard deviation	median	average
		raw variables			
Life expectancy at birth	48.943 (Swaziland)	84.163 (Hong Kong)	8.297	73.415	71.353
Expected years of schooling	4.872 (South Sudan)	20.433 (Australia)	2.897	13.140	12.983
Mean years of schooling	1.442 (Burkina Faso)	13.370 (Switzerland)	3.097	8.656	8.372
GNI per capita	587.474 (Central African Republic)	129915.601 (Qatar)	19069.312	10415.970	17313.866

Note: The country in parenthesis indicates where the respective minimum/maximum is found.

component scores. The group of "Laggards" consists of the relatively worst performing countries, which achieve efficiency and mix component scores below a certain threshold, which was set at 80%. Thus, we consider inefficiencies below 80% as significant enough to raise alarms to supervising agencies. The remaining DMUs are relatively inefficient with respect to both measures to some extent, but not as severely as the Laggards, i.e., their efficiency scores and mix component are both below 99% but at least one is above 80%. These were altogether grouped as "inefficient".

2.3.2. Empirical Results

Estimates of effectiveness, efficiency and the mix component by group, income class and geographical region are given in Table 2.2. The arithmetic average and aggregate values of efficiency scores and the mix component for the full sample of 188 countries are 0.927 and 0.906 respectively, indicating that ineffectiveness is caused more by countries' imbalanced prioritization on health and education provision (captured by the mix component) than by having relatively low achievements relative to their economic prosperity levels (inefficiency).[20] This is clearly reflected in the shape of the kernel distributions of efficiency scores (see Figure 2.2) where the mix component distribution

---

[20] According to Färe and Karagiannis (2017), the aggregate values are computed using potential output shares. However, as we have more than one outputs for which there are no market prices, we have to approximate their "market" shares. Here we follow the approximation suggested by Färe and Zelenuyk (2003) that assumes that the value of the total amount of any output is the same as the value of the total amount of any other output. This implies that the aggregation weights are equal to the unweighted average of the shares of the individual countries corresponding to each output, i.e., $\frac{1}{K}\sum_{j=1}^{J}\left(y_j^k / \sum_{k=1}^{K} y_j^k\right)$.

Table 2.2: Estimates of effectiveness, efficiency and the mix component

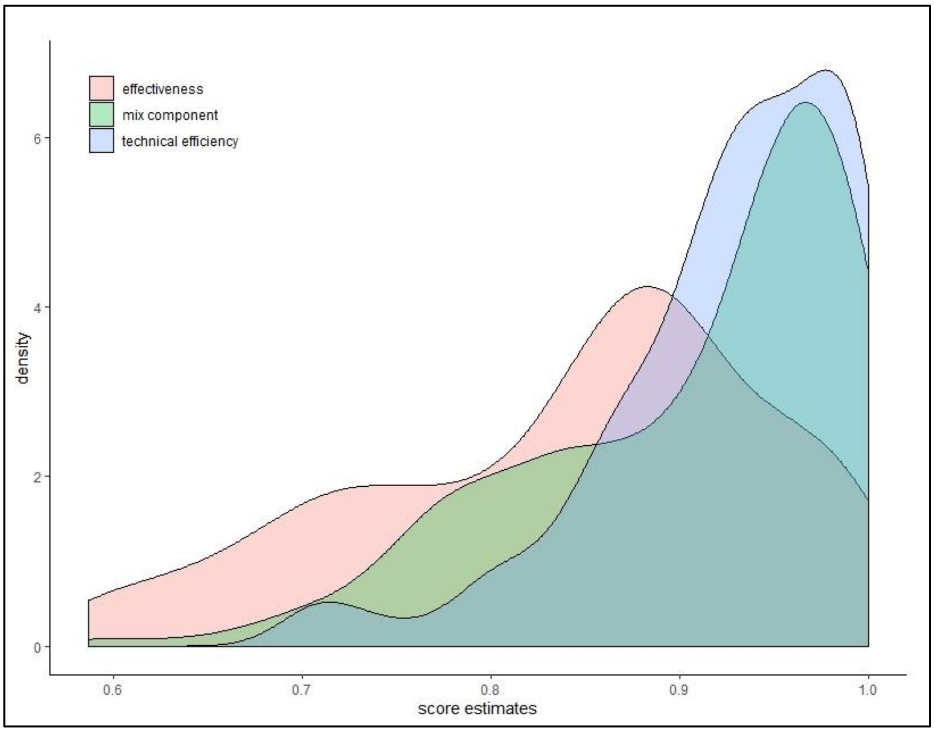
		effectiveness	efficiency	mix component
World (188 countries)	maximum	1.000	1.000	1.000
	minimum	0.587	0.704	0.594
	average	0.841	0.927	0.906
	aggregate	0.849	0.930	0.913
by cluster				
leaders	maximum	1.000	1.000	1.000
	minimum	0.989	0.996	0.991
	average	0.995	0.999	0.996
	aggregate	0.995	0.999	0.996
mix efficient	maximum	0.989	0.989	1.000
	minimum	0.885	0.893	0.991
	average	0.963	0.966	0.996
	aggregate	0.963	0.967	0.996
efficient	maximum	0.988	1.000	0.988
	minimum	0.594	0.991	0.594
	average	0.840	0.998	0.841
	aggregate	0.850	0.998	0.852
inefficient	maximum	0.968	0.988	0.990
	minimum	0.588	0.704	0.706
	average	0.822	0.910	0.903
	aggregate	0.829	0.911	0.909
laggards	maximum	0.595	0.789	0.790
	minimum	0.587	0.743	0.749
	average	0.591	0.765	0.773
	aggregate	0.591	0.766	0.772
by income class				
high income	maximum	1.000	1.000	1.000
	minimum	0.845	0.857	0.913
	average	0.948	0.963	0.984
	aggregate	0.949	0.964	0.984
upper-middle income	maximum	0.951	1.000	0.980
	minimum	0.588	0.704	0.829
	average	0.855	0.908	0.941
	aggregate	0.855	0.909	0.940
lower-middle income	maximum	0.897	1.000	0.961
	minimum	0.595	0.707	0.779
	average	0.794	0.911	0.871
	aggregate	0.796	0.913	0.872
low income	maximum	0.801	1.000	0.826
	minimum	0.587	0.743	0.594
	average	0.691	0.918	0.755
	aggregate	0.692	0.918	0.754

(Table 2.2 continued)

		effectiveness	efficiency	mix component
	by geographical region			
North America and the Caribbean	maximum	0.990	1.000	0.994
	minimum	0.728	0.857	0.791
	average	0.884	0.940	0.940
	aggregate	0.886	0.940	0.942
Europe (all)	maximum	1.000	1.000	1.000
	minimum	0.890	0.916	0.913
	average	0.949	0.968	0.980
	aggregate	0.949	0.968	0.980
Europe (EU)	maximum	0.988	1.000	1.000
	minimum	0.890	0.916	0.913
	average	0.948	0.966	0.981
	aggregate	0.948	0.966	0.981
Europe (non-EU)	maximum	1.000	1.000	1.000
	minimum	0.892	0.934	0.945
	average	0.951	0.972	0.978
	aggregate	0.952	0.973	0.979
North Africa	maximum	0.879	0.948	0.949
	minimum	0.833	0.885	0.909
	average	0.855	0.917	0.933
	aggregate	0.856	0.916	0.934
South, East Asia and Oceania	maximum	1.000	1.000	1.000
	minimum	0.720	0.856	0.784
	average	0.861	0.944	0.912
	aggregate	0.868	0.947	0.916
South America	maximum	0.969	1.000	0.969
	minimum	0.785	0.842	0.879
	average	0.874	0.924	0.945
	aggregate	0.874	0.925	0.945
Sub-Saharan Africa	maximum	0.883	1.000	0.977
	minimum	0.587	0.704	0.594
	average	0.700	0.872	0.807
	aggregate	0.702	0.862	0.814
North, West and Central Asia	maximum	0.991	1.000	0.991
	minimum	0.694	0.812	0.789
	average	0.861	0.935	0.921
	aggregate	0.864	0.936	0.923

has a higher density that the technical efficiency one for lower values of estimates (below 0.85). Also, from the 20 countries that DEA identifies as offering the highest possible social prosperity relative to their economic prosperity (i.e., the efficient countries) only four (namely Australia, Hong Kong, Japan and Switzerland) are denoted

Figure 2.2: Kernel density estimates for the efficiency, effectiveness and mix component, 2015.



by VEA as 100% effective.[21] These countries receive a score of one for all three measures.

Eight countries in total (Australia, Hong Kong, Japan and Switzerland which are efficient and Norway, Denmark, Singapore and Sweden which are inefficient) have a mix component equal to one, i.e., their DEA and VEA scores are equal to each other. The inefficient countries with a mix component equal to one manage to offer to their citizens a high balance in the provision of health and education services, but not the "highest possible" amount relative to their economic prosperity, as there are other countries having slightly higher social achievements with the same levels of economic prosperity. The lowest efficiency score among those four countries is 0.972, by Denmark. Nevertheless, the negative skewness for the three measures (see Figure 2.2) along with the minimum scores indicates the existence of highly inefficient countries, which are in dire need of restructuring actions. Such actions could include increases in

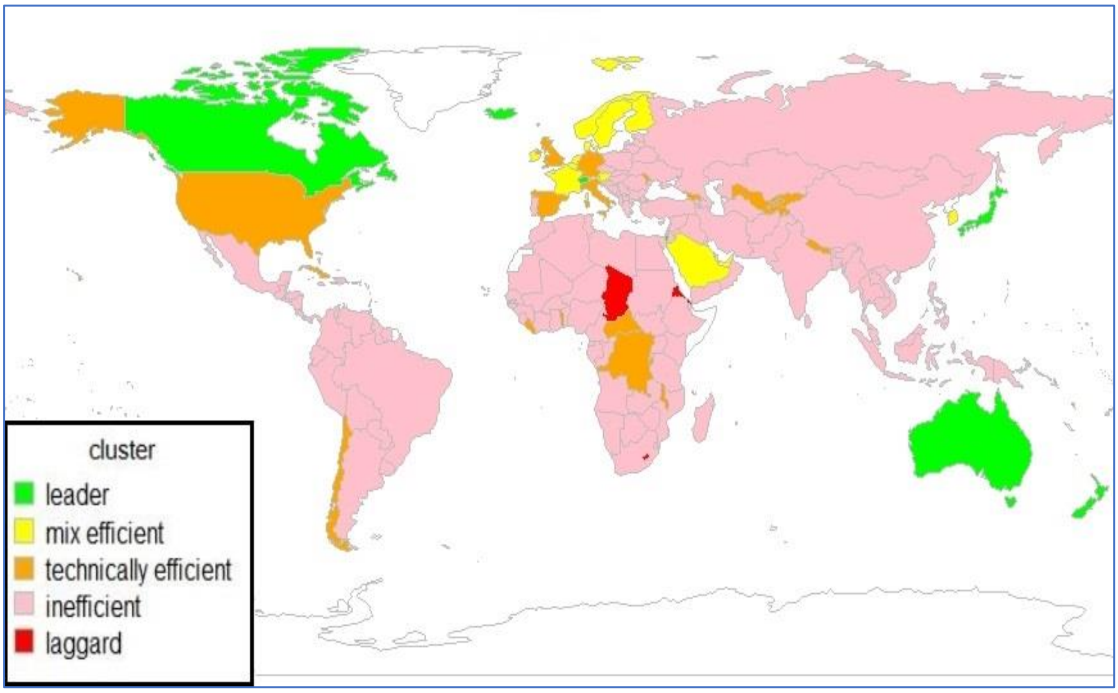
[21] The efficient countries are (in alphabetical order) Australia, Burundi, Central African Republic, Cuba, Georgia, Hong Kong, Iceland, Italy, Israel, Japan, Kyrgyzstan, Liberia, Moldova, Nepal, Republic of Congo, Solomon Islands, Switzerland, Tajikistan, United Kingdom and Uzbekistan.

health and education expenditures and a better management in order to decrease resource waste. For countries with low mix component scores, budget redistribution could also be an action leading to increased levels of future social prosperity.

In the second panel of Table 2.2 we present the results for the clustering of countries according to their efficiency and mix component scores. This clustering is also portrayed depicted in Figure 2.3 and Table 2.3 presents in more detail the countries in each cluster. Note that in Table 2.3 we have split the inefficient group into six sub-groups based on an additional threshold set at 95% for technical efficiency and the mix component. The Leader group of countries outperforms on average all other groups in all three measures. The eight Leader countries (see Table 2.3) include industrialized and well-performing countries in terms of social prosperity such as Canada, Australia and Japan. These are the ones doing 'the right things' and can be considered by the supervising agency (e.g. the UN) as undeniable best performers whose behavior should be copied by other countries in the future. The 17 "mix efficient" countries include 12 European ones, among which we find most of the Nordic countries along with many EU members such as France, Ireland and Belgium. The group is filled with two Asian (Korea, Singapore) and three Arabic countries (Qatar, Saudi Arabia and United Arab Emirates). These countries, which provide to their citizens the highest possible balance in social prosperity outcomes (i.e., the operate within the preferred range of bundles set out by the "model" country), are also providing very high levels of social prosperity relative to their economic prosperity (their average efficiency score is 0.966) and consequently they display high effectiveness (average 0.963).

On the contrary, the 22 countries of the "efficient" group, which display the "highest possible" social prosperity achievements relative to their economic prosperity, are not concentrated on a specific region but are scattered across the world, including countries as diverse as USA, Chile and Uzbekistan. Furthermore, this group of countries is well performing only with respect to efficiency while having mediocre average performances in terms of balance in the provision of social prosperity (the groups' mix component varies from the low 59.4% in Central African Republic to the well-performing 98.8% in Italy) and consequently, in terms of effectiveness. This group should focus disproportionately more in improving balance in their social prosperity outcomes through gradual changes in their mix. The inefficient countries slightly outperform the efficient countries in terms of the mix component (average estimate

Figure 2.3: Country clusters by means of efficiency and the mix component



0.903 compared to 0.841). The three Laggard countries, namely Chad, Lesotho and Sierra-Leone (see Table 2.3) belong to the Sub-Saharan African region. For those countries there seems to be vast room for future improvement, as they are displaying both very low as well as very unbalanced social prosperity achievements relative to their levels of economic prosperity. The average estimates of efficiency (0.765) and mix component (0.773) are the lowest across groups and lead to an average effectiveness score of 59.1%, also the lowest among all groups.

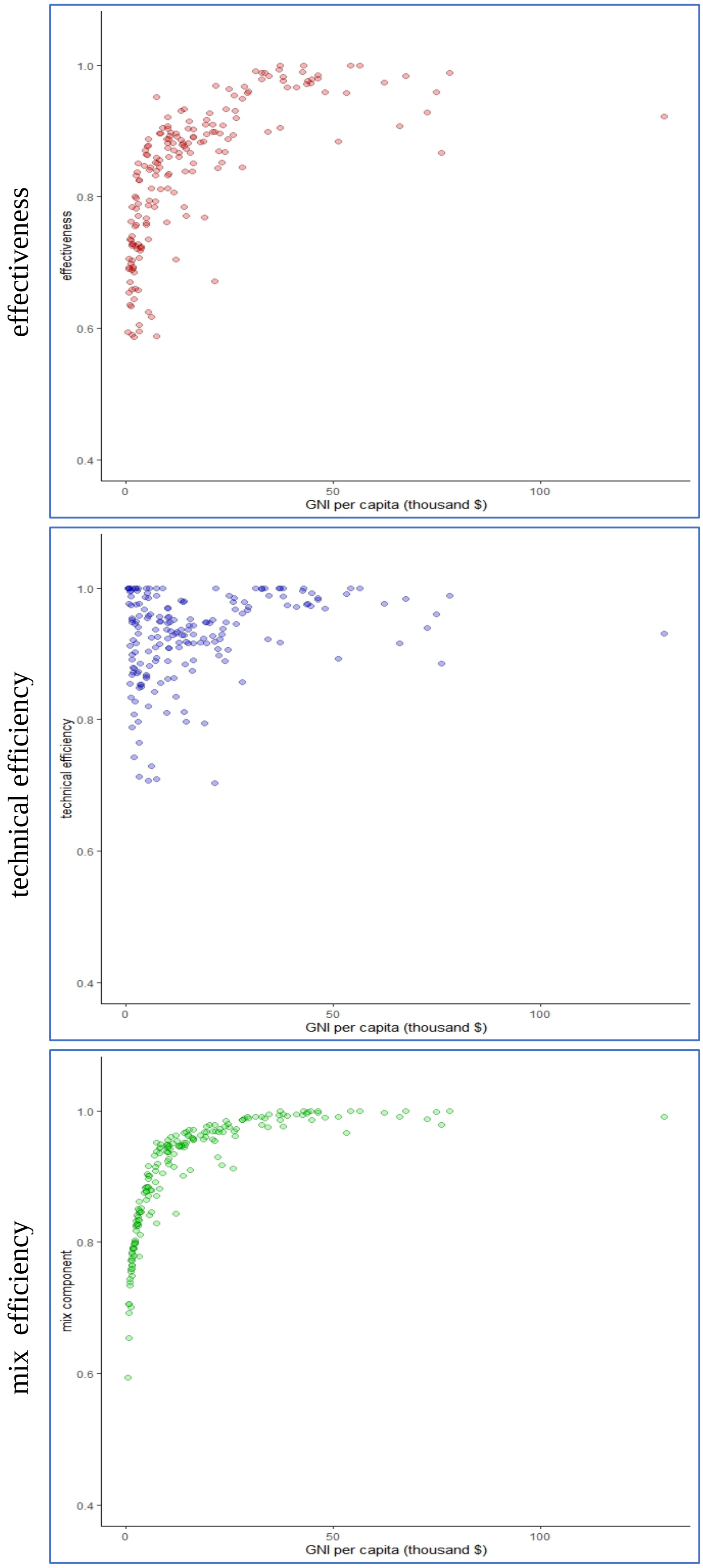
The performance of countries by income class is given in the third panel of Table 2.2 and in Figure 2.4 we plot effectiveness, efficiency and the mix component against GNI per capita.[22] Effectiveness and the mix component seem to follow an S-curve with respect to income, which is more intense in the mix component case. As income increases average effectiveness and the mix component also increase, with the highest shift in average values being between low and lower-middle income classes. This suggests that a small initial ‘’push’’ in a country’s income can spark significant improvements in the provision of health and education services, i.e., that returns to

[22] The respective information for the 2015 country clustering by income class was retrieved from the World Bank.

Table 2.3: Classifying countries by means of efficiency and the mix component

		mix component			
		[1-0.99]	(0.99-0.95]	(0.95-0.8]	(0.8-0]
efficiency	[1-0.99]	Australia, Canada, Hong Kong (China, SAR), Iceland, Israel, Japan, New Zealand, Switzerland	Chile, Cuba, Germany, Italy, Spain, United Kingdom, United States	Georgia, Kyrgyzstan, Moldova (Republic of), Nepal, Tajikistan, Tonga, Uzbekistan, Vanuatu	Burundi, Central African Republic, Congo (Democratic Republic of), Liberia, Malawi, Solomon Islands, Togo
	(0.99-0.95]	Austria, Belgium, Cyprus, Denmark, Finland, France, Ireland, Korea (Republic of), Liechtenstein, Luxembourg, Netherlands, Norway, Singapore, Sweden	Andorra, Argentina, Costa Rica, Czech Republic, Estonia, Greece, Malta, Portugal, Slovenia	Albania, Bangladesh, Belarus, Bosnia and Herzegovina, Cabo Verde, Dominica, Fiji, Grenada, Honduras, Lebanon, Lithuania, Maldives, Micronesia (Federated States of), Nicaragua, Palau, Palestine (State of), Samoa, Syrian Arab Republic, Ukraine, Viet Nam	Comoros, Ethiopia, Madagascar, Niger, Rwanda
	(0.95-0.8]	Qatar, Saudi Arabia, United Arab Emirates	Antigua and Barbuda, Azerbaijan, Bahamas, Bahrain, Barbados, Brunei Darussalam, Bulgaria, Croatia, Hungary, Iran (Islamic Republic of), Jordan, Kuwait, Latvia, Malaysia, Mauritius, Mexico, Montenegro, North Macedonia, Oman, Panama, Peru, Poland, Romania, Saint Kitts and Nevis, Serbia, Seychelles, Slovakia, Sri Lanka, Suriname, Thailand, Trinidad and Tobago, Turkey, Turkmenistan, Uruguay, Venezuela (Bolivarian Republic of)	Algeria, Armenia, Belize, Bhutan, Bolivia, Brazil, Cambodia, China, Colombia, Congo, Djibouti, Dominican Republic, Ecuador, Egypt, El Salvador, Ghana, Guatemala, Guyana, India, Indonesia, Iraq, Jamaica, Kazakhstan, Kenya, Kiribati, Lao People's Democratic Republic, Libya, Mauritania, Mongolia, Morocco, Myanmar, Namibia, Pakistan, Papua New Guinea, Paraguay, Philippines, Russian Federation, Saint Lucia, Saint Vincent and the Grenadines, Sao Tome and Principe ,South Africa, Sudan, Tanzania (United Republic of), Timor-Leste, Tunisia, Yemen, Zambia	Afghanistan, Benin, Burkina Faso, Eritrea, Gambia, Guinea, Guinea-Bissau, Haiti, Mali, Mozambique, Senegal, South Sudan, Uganda, Zimbabwe
	(0.8-0]	-	Botswana, Equatorial Guinea, Gabon	Angola, Cameroon, Côte d'Ivoire, Nigeria, Swaziland	Chad, Lesotho, Sierra Leone

Figure 2.4: Relation between income and performance



income are increasing in lower income levels. After a certain level of GNI per capita (around 12.000 $ which is close to the threshold between upper-middle and high-income class) the curvature changes and the returns to income in social prosperity become decreasing, indicating that further economic prosperity increases improve only slightly a country's achievements in terms of social prosperity. There is no high-income country that departs more than 8.7% from the preferred range of health and education bundles, as the minimum estimate of the mix component for a high-income country is as high as 0.913 (see Table 2.2). Thus, the virtually zero gain in human development and well-being when income surpasses a certain threshold that is found by Kahneman and Deaton (2010) can be partly explained by that fact that most high-income countries are already displaying both very high as well as highly balanced achievements in terms of social prosperity relative to their (also high) levels of economic prosperity. They are prioritizing relatively even between health and education provision and exploiting heavily the spillovers between simultaneous improvements in both of them.

For efficiency however there does not seem to be a particular pattern. There is no significant difference between low- and the two middle-income classes in terms of efficiency while the high-income countries are on average only 6% more efficient that the low-income countries. Thus, a low level of economic prosperity does not appear to prevent a country from exploiting it to the highest possible extent to provide social prosperity outcomes (i.e., ''doing things right'') as efficiency is realized for a wide range of income levels, from the extremely low-income Central African Republic (GNI per capita 587.474 $ in 2015) to high-income Switzerland (PPP GNI per capita 56,363.958 $ in 2015). On the other hand, equal prioritization and efficient resource use seem to be associated with higher income levels. High-income countries seem to have the know-how about ''doing the right things'' in terms of enhancing social prosperity and further developing the capabilities of their citizens.

The lower panel of Table 2.2 we present the results by geographical region. From there we can see: *first*, North-Central America and the Caribbean is a rather diverse region whose good average efficiency and mix component performance is mainly supported by the two North American countries, USA and Canada. Most Central American and Caribbean countries are classified as inefficient. *Second*, South American countries on average appears to provide slightly less social prosperity achievements relative to their economic prosperity, compared to their Northern neighbors (average efficiency in South America is 0.842 compared to 0.857 in North America) but on the same time, offering considerably more balanced provision of health and education (the average mix component in South America is 0.879 while that of North America is 0.791). *Third*, South American countries located north or south of the tropical Amazon rainforest seem, to offer a more balanced mix of health and education outcomes compared to countries in the center of South America (e.g. Brazil, Argentina, and Paraguay, see also Table 2.3). *Fourth*, Europe is the best performing region on average in all three measures, while EU-member countries slightly outperform the non-EU countries in terms of balance in the provision of social prosperity but lag slightly behind in terms of efficiency. *Fifth*, three of the Nordic countries (Norway, Denmark, Sweden) have a mix component equal to one, with the rest of the Nordic group (Iceland and Finland) following suit with scores greater than 0.99. On the opposite, the worst performing European countries in terms of the mix component are three Balkan countries (North Macedonia, Albania, Bosnia and Herzegovina) and two Eastern European countries (Estonia and Lithuania). *Sixth*, the Sub-Saharan African (SSA) group of countries is the worst performing region on average in all three measures. The majority of the SSA countries score below 90% in terms of the mix component while half of them (26) score below the threshold of 80%. This poor performance suggests that the SSA region is in urgent need of restructuring actions, such as shifting the allocation of natural resources revenues from recruitment and administrative to health and education expenditures (Raheem *et al.*, 2018) or using the same revenues in order to boost human capital (Oyinlola *et al.*, 2020).

2.3.3. Robustness checks

We next present a robustness check by (a) considering two alternative MPS choices for the year 2015, namely Australia (the country that ranked second in the 2015 HDI) and an artificial country comprised by the average of the five efficient countries with the highest ranks in the 2015 HDI, namely Australia, Switzerland, Iceland, Hong-Kong and United Kingdom and (b) extending the period under consideration to 2014-2018, using our initial "model" country.

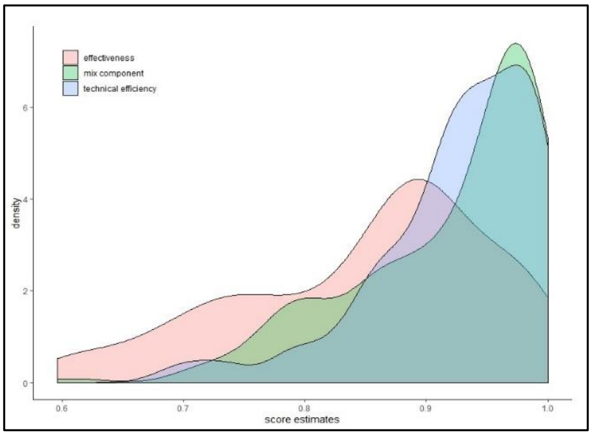
The use of Australia as the "model" country for the year 2015 results into relatively higher effectiveness and mix component scores relative to Norway being the "model" country, mostly because it results to a wider range of preferred bundles. This is to be expected as Australia was a DEA-efficient country while Norway was not. This is also clear from Figure 2.1(a) if we consider DMU J as Norway and DMU B as Australia. Using DMU B as the MPS expands both facets AB and BC (blue line) and results into more countries attaining a score of one for the mix component and consequently, to higher effectiveness scores compared to the use of DMU J as the MPS, which expands only facet AB (the red line). The average effectiveness score with Australia as the "model" country was 0.917 (compared to 0.841 with Norway as the "model" country) while the number of "mix efficient" countries was 126 (compared to 17 in the case of Norway). On the other hand, using as "model" country the artificial "average best performing" country was operationalized by using its efficient peers as the MPS, i.e., Australia, Switzerland, Hong-Kong and Japan. This set of peers is the same as that for Norway with the addition of Japan. Adding Japan in the set of MPS brings virtually no changes to effectiveness and mix component scores with respect to the case of Norway being the "model" country; all average scores as well as the classification of countries remains the same. Referring again to Figure 2.1(a), let Norway and the artificial average country correspond respectively to DMUs J and I, which both are projected in facet AB and therefore, their use as MPS extends the same facet.

Estimates of effectiveness, technical efficiency and mix components scores for years 2014, 2017 and 2018 using Norway as the "model" country are given in Table 2.4 and their kernel density distributions are portrayed in Figure 2.5. Overall, the results across years remain relatively stable. The only notable change occurs in 2017, where the distribution of the mix component became more skewed towards unity compared to other years (see Table 2.4). This is due to changes in the set of peers for Norway. More specifically, Norway's peers were the same in 2014 and 2015 (namely, Australia, Switzerland and Hong-Kong), while in 2017 changed to Australia, Hong-Kong and Japan. The exclusion of Switzerland, a country with a relatively narrower preferred range of bundles, caused the preferred bundle range to widen in 2017 relative to other years. Referring to Figure 2.1(a), this is as the preferred range of bundles temporarily moved from OA-OB to OA-OC. In 2018, Norway's peers are Australia, Switzerland and Japan, highlighting a return to a narrower preferred range of bundles which is similar to those of years 2014 and 2015, as it can be seen from Table 2.4.

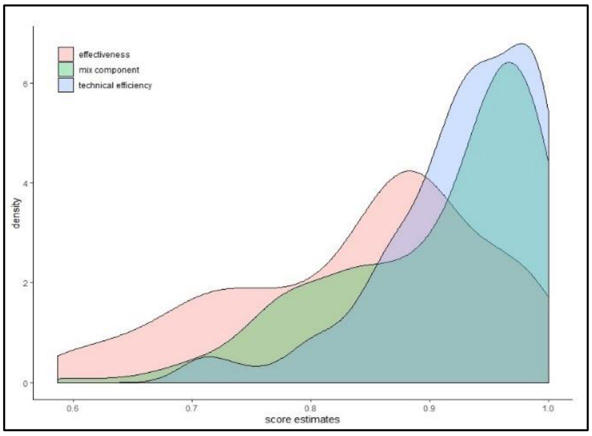
Table 2.4: Distribution of effectiveness, technical efficiency and mix component scores, 2014, 2015, 2017 and 2018 with the same MPS choice (Norway)

	2014	2015	2017	2018	2014	2015	2017	2018	2014	2015	2017	2018
	effectiveness				efficiency				mix component			
(zero)	0	0	0	0	0	0	0	0	0	0	0	0
(0,0.1)	0	0	0	0	0	0	0	0	0	0	0	0
[0.1,0.2)	0	0	0	0	0	0	0	0	0	0	0	0
[0.2,0.3)	0	0	0	0	0	0	0	0	0	0	0	0
[0.3,0.4)	0	0	0	0	0	0	0	0	0	0	0	0
[0.4-0.5)	0	0	0	0	0	0	0	0	0	0	0	0
[0.5,0.6)	2	5	0	3	0	0	0	0	0	1	0	0
[0.6,0.7)	16	19	7	23	1	0	0	0	1	2	1	3
[0.7,0.8)	37	37	33	37	11	11	6	6	20	26	5	26
[0.8,0.9)	65	71	64	69	41	40	43	46	42	40	36	43
[0.9,1)	64	52	82	53	117	117	125	124	119	112	134	108
1	4	4	3	4	18	20	15	13	6	7	13	9
maximum	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
minimum	0.596	0.587	0.644	0.571	0.685	0.704	0.704	0.714	0.610	0.594	0.652	0.658
average	0.850	0.841	0.871	0.838	0.924	0.927	0.925	0.928	0.918	0.906	0.941	0.900
aggregate	0.859	0.849	0.880	0.847	0.928	0.930	0.929	0.932	0.925	0.913	0.947	0.909
standard deviation	0.104	0.107	0.086	0.108	0.067	0.066	0.060	0.058	0.078	0.085	0.064	0.086
median	0.875	0.867	0.891	0.867	0.938	0.939	0.934	0.944	0.951	0.943	0.969	0.920

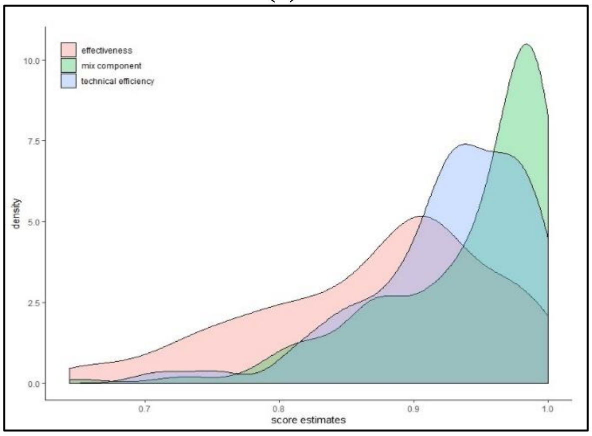
Figure 2.5:Kernel density estimates for the efficiency, effectiveness and mix component, 2014, 2015, 2017 and 2018 with the same MPS choice (Norway)



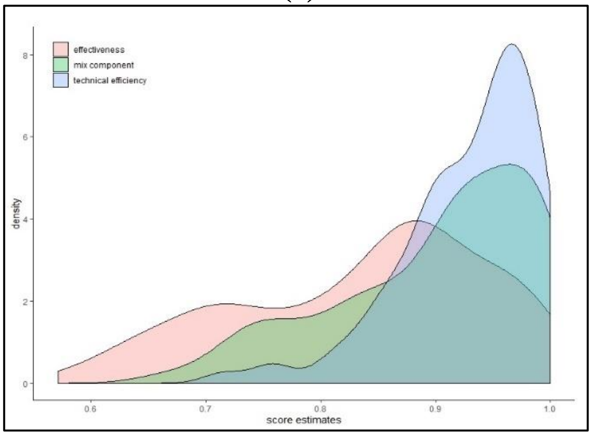
(a) 2014



(b) 2015



(c) 2017



(d) 2018

2.4. Concluding remarks

In this chapter we use VEA to assess effectiveness. VEA captures the extent of ‘’doing the right things’’ through the choice of a “model” DMU which defines a preferred range of input/output bundles. DMUs operating within that range are perceived as effective by managerial authorities, while DMUs operating outside the preferred range should be directed towards mix changes and restructuring. Effectiveness was then decomposed to two measures reflecting the extent of ‘’doing things the right way’’ (efficiency) and producing out of the given range of input and output mixes’ (mix component

The proposed approach could be utilized in many real-world instances where an evaluation of units is sought and managerial preferences need to be taken into account along with efficiency issues. At a micro level, our approach could aid firm owners, CEOs or HR departments to make decisions upon hiring, promoting or allocating personnel based on both their operating efficiency and effectiveness or allocate pay-for-performance funds within a firm. Similarly at a macro level, international organizations such as the International Development Association, Development Banks, the World Bank or the European Union could utilize such a tool with the aim of allocating several kinds of funds towards countries or regions which are either the best (if the funds act as rewards) or the worst performers (if the funds act as aid).[23] A final note regards the sensitivity of the method to the choice of the MPS. As the chosen MPS unit defines the preferred range of bundles, it certainly affects the effectiveness scores and their decomposition into efficiency and the mix component.

In the case of the development of human capabilities considered in the empirical application, the use of a "model" country to assess the effectiveness of converting economic to social prosperity allows us to identify the countries providing inappropriate mixes of health and education services and those providing an inappropriate amount of the suggested mixes of health and education services. The former may use policies to correct their deficiencies such as redistributing government expenditures more evenly across health and education in their future balance sheets or directly targeting "priority areas", i.e. the service provision sector which is the relatively most neglected among health and education, through the creation of infrastructure (schools, hospitals) or the implementation of new regulations (e.g. population immunization policies through mandatory vaccination).[24] The latter can benefit from policies that redistribute government expenditures from other uses (e.g. administrative expenditures) to health and education to further improve their achievements, i.e. increase their HD-allocation ratio (see Ranis *et al.*, 2000), as well as from policies that enhance the efficient use of those expenditures, such as better monitoring mechanisms for government officials that handle the relevant contracts.

---

[23] A case of reward-funds is considered by Golany and Thore (1997): the evaluation by the World Bank or some UN agency of loan requests made by developing countries.

[24] Ranis *et al.* (2000) refer to the proportion of government expenditures for sectors related to human development that is attributed to such priority areas as HD priority ratio and argue that the latter is affected positively by the extent of government decentralization.

CHAPTER 3

A VEA Benefit-of-the-Doubt model for the HDI

3.1. Introduction

The Benefit-of-the-Doubt (BoD) as an input-oriented Data Envelopment Analysis (DEA) model with a single constant input, mainly used for constructing composite indicators (CI) (see OECD, 2008) but also applied to several multi criteria decision making problems such as supplier selection, inventory classification, quality perception, preference voting, location selection, etc. (see Karagiannis (2021) and the references therein), has the advantages and disadvantages of DEA models. Among them is the flexibility of weights that may vary across DMUs and indicators in such a way so as to maximize the overall achievement of each evaluated Decision-Making Unit (DMU). This flexibility may in some cases imply that DMUs are rated as efficient by doing well only on a single performance dimension. In such a case, zero weights are assigned to all but one indicator. However, more often, DMUs are evaluated based only on a subset of the considered indicators (most notably those in which they perform relatively better), implying that the rest have no effect on the CI. As this subset of indicators may differ across DMUs, it makes their multilateral comparison difficult or even inappropriate.[25,26]

---

[25] For example, consider two countries A and B being evaluated on the basis of two indicators, namely I_1 (patents) and I_2 (research grants, in thousand \$). If country A outperforms B in terms of patents but is outperformed by B in terms of research grants, the BoD model will base the composite indicator of country A only on the patents indicator and assign a zero weight on the research grants indicator, while the reverse will occur for country B. Comparing the performance of the two countries using these composite indices, would be deemed inappropriate.

[26] The benefit-of-the-doubt weighting might also be criticized for dismissing one of the three basic requirements in social choice theory in response to Arrow's theorem, namely anonymity or the assignment of equal weights to all indicators. Nevertheless, OECD (2008, p. 105) argue that anonymity

To avoid this kind of issues when implementing the BoD model several researchers have restricted weight flexibility by imposing different type of restrictions. These may take the form of (*i*) absolute weight bounds (see e.g., Rogge and Self, 2019), (*ii*) relative restrictions in the form of assurance regions (Gaaoul and Khalfallah, 2014), (*iii*) ordinal ranking of indicators' importance (see e.g., Cherchye *et al.* 2007b), and (*iv*) pie shares (Cherchye *et al.,* 2007a; Gonzalez *et al.,* 2018). Such forms of weights restrictions require experts, stakeholders or social planners to set the absolute or relative bounds or shares, a task that can be proved to be difficult and time-consuming since usually experts may have diverging opinions regarding the relative importance of indicators.

An alternative way to incorporate value judgments in BoD is by means of Value Efficiency Analysis (VEA), developed by Halme *et al.* (1999). This alternative has been suggested in OECD (2008, p. 92), where it is noted that "the benchmark could also be determined by a hypothetical decision maker … who would locate the target in the efficiency frontier with the most preferred combination of individual indicators'', but to the best of our knowledge it has not been implemented so far. In VEA, the views of an expert, a decision maker (DM) or a supervising agency are reflected in the choice of a "model" DMU that determines the Most-Preferred Solution (MPS) from their point of view. This *alternative in gauging preferences* might be proven to be more appealing to Decision Makers, as the latter might be more keen on choosing one DMU to serve as a benchmark, rather than engage in the task of selecting weight restriction bounds (Korhonen *et al.,* 2002).[27] The choice of the "model" DMU restricts the weights that the evaluated DMUs can select by determining a preferred range of indicator bundles.

In this chapter we use the VEA-BoD model to re-estimate the United Nations (UN) Human Development Index (HDI) using data for 2015. For these purposes we rely on the notion of uniformity, namely the intension for equal prioritization among the considered indicators, to choose the "model" country. Based on this objective and

---

is not an essential requirement in the construction of a composite indicator, as equal weighting is usually only one of the possible weighting schemes.

[27] We should emphasize that the chapter's aim is to provide an *alternative* approach to that of weight restrictions in incorporating DM preferences to the conventional BoD model, rather than an approach that performs *better* in restricting the flexibility of weights in conventional BoD, compared to weight restrictions.

normative argument, countries with a relatively balanced prioritization among health, education and income would be promoted as peers for improving human capabilities in the rest of the countries. The latter may need a shift of focus towards policies aimed at improving their most deprived human development dimensions.

The rest of the chapter proceeds as follows: In the next section we briefly discuss the conventional BoD model and we introduce the VEA-BoD model. In the third section, we present the empirical results for the HDI based on the VEA-BoD model. Concluding remarks follow in the last section.

3.2. The Conventional and the VEA BoD models

The BoD is a model facilitating the (linear) aggregation of a number of quantitative indicators into a single CI when exact knowledge of the weights is not available.[28] The model endogenously selects the best possible weights for each DMU, assuming implicitly that the DMUs attach less (more) importance to those indicators on which they perform relatively weak (strong) compared to the other evaluated units. The model is a special case of the input-oriented constant-returns-to-scale DEA model with a single constant input that takes the value of one for all DMUs (see Karagiannis, 2021). Its multiplier and envelopment forms are given as:

$$
\begin{array}{llll}
\max\limits_{u_j^o} & \sum\limits_{j=1}^{J} u_j^o y_j^o & & \\
s.t. & \sum\limits_{j=1}^{J} u_j^o y_j^k \leq 1 & k = 1, \dots, K & \\
 & u_j^o \geq 0 & j = 1, \dots, J &
\end{array}
\qquad
\begin{array}{llll}
\min\limits_{\lambda_k^o} & \sum\limits_{k=1}^{K} \lambda_k^o \ \ (= \theta^o) & & \\
s.t. & \sum\limits_{k=1}^{K} \lambda_k^o y_j^k \geq y_j^o & j = 1, \dots, J & \\
 & \lambda_k^o \geq 0 & k = 1, \dots, K &
\end{array}
\qquad (3.1)
$$

[28] The BoD is one of the four approaches proposed by OECD (2008) for constructing composite indicators. However, CI construction is a constantly expanding research field, in which several new methodological advancements exist. Some of these contributions are related to the BoD model, others make use of multicriteria decision-making approaches, such as goal-programming and non-compensatory approaches, while there are also mixed or hybrid approaches combining different methodologies to construct a composite indicator. A review of these approaches is a task out of the scope of this chapter, and the interested reader is referred to Greco *et al.* (2019) and El Gibari *et al.* (2019) for recent reviews.

where y refers to the component indicators, u to their relative weights (multipliers), θ to the efficiency score, λ to the intensity variables, $j = 1, \dots, J$ is used to index indicators and k ($k = 1, \dots, o, \dots, K$) to index DMUs.

Expert judgments, DM's mandates or public opinion and social planner norms regarding the performance of the evaluated DMUs are not *a priori* incorporated in the conventional form of the BoD model in (3.1). As a result, each DMU has the benefit-of-the-doubt in the selection of its relative weights in order to maximize the value of its CI. This allows DMUs that dominate all others in a single indicator to be rated as efficient even though they perform relatively weak in terms of all other indicators. The CI values resulting from (3.1) are thus the most optimistic for each DMU.

VEA takes into account DM's preferences or public opinion about desired norms and managerial or social goals by means of a pseudo-concave value function (i.e., an indifference curve) that becomes tangent to the DEA efficient frontier at a point referred to as the MPS. This point, ultimately chosen by a DM or a supervising agency, corresponds to a virtual or real DEA-efficient DMU, which is viewed as the "model" DMU having the most preferred input/output bundle. The VEA frontier is constructed by extending towards the axes the hyperplanes of the DEA efficient facets intercepting at the MPS, a process that naturally results in efficiency scores that are lower or equal to those of the conventional DEA model. This is depicted in Figure 3.1 for the case of two indicators. Choosing for example DMU B as the MPS extends facets AB and BC towards the axes, creating the VEA frontier (the red kinked line). This defines a range of preferred bundles given between rays OA and OC. As a result, the DEA benchmark profiles complying with the desired norms are now limited to facets AB and BC. For all inefficient DMUs which are radially projected in these two facets, the CI value that results from the VEA-BoD model is equal to that of the conventional BoD model while inefficient DMUs projected elsewhere on the BoD frontier and thus using a bundle outside of the preferred range are "penalized" and their CI value is less than that obtained from the conventional BoD model.

Following Halme *et al*. (1999), the VEA formulation of the BoD model in (3.1) is given, in its multiplier and envelopment form, as follows:

$$
\begin{array}{llll}
\max\limits_{u_j^o} & \sum\limits_{j=1}^{J} u_j^o y_j^o & & \\
s.t. & \sum\limits_{j=1}^{J} u_j^o y_j^k \leq 1 & k = 1, \dots, K, k \neq r \\
& \sum\limits_{j=1}^{J} u_j^o y_j^k = 1 & k = r \\
& u_j^o \geq 0 & j = 1, \dots, J
\end{array}
\qquad
\begin{array}{llll}
\min\limits_{\lambda_k^o} & \sum\limits_{k=1}^{K} \lambda_k^o & \\
s.t. & \sum\limits_{k=1}^{K} \lambda_k^o y_j^k \geq y_j^o & j = 1, \dots, J \\
& \lambda_k^o \geq 0 & k = 1, \dots, K, k \neq r \\
& \lambda_k^o \; free & k = r
\end{array}
\qquad (3.2)
$$

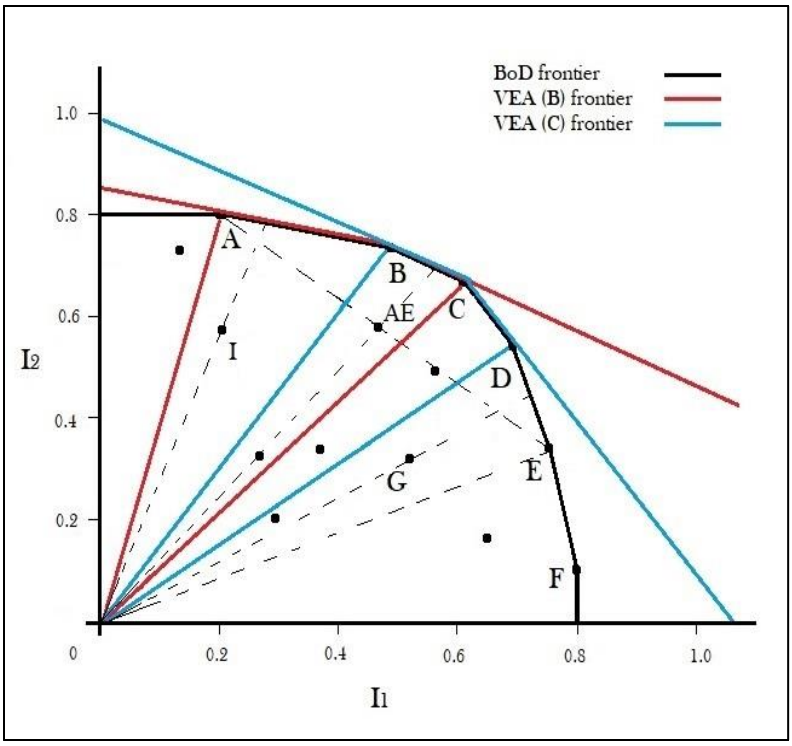
in which *r* refers to the DMU chosen as the MPS.

As we can see there are only slight differences between (3.1) and (3.2): in the multiplier form, the restriction corresponding to the MPS is turned from an inequality to a strict equality in order the range of acceptable weights to be restricted into those that are optimal for the "model" DMU. This in turn restricts the preferred range of bundles to be between rays OA and OC in Figure 3.1, if DMU B is chosen as the MPS. This corresponds in to removing the non-negativity restriction from the intensity variable corresponding to the MPS in the envelopment form of (3.2.), thus forcing the MPS to be peer for all evaluated DMUs.

In practical settings, the most crucial step in VEA is the choice of the MPS, as it affects the resulting frontier and, consequently, the derived efficiency scores. Nevertheless, no general rule of thumb exists for choosing the MPS, but several suggestions have been proposed. These involve the choice of either a real (usually DEA-efficient) DMU or a combination of DEA-efficient DMUs. The latter case can be operationalized as long as the combined DMUs generate at least one common facet, in which case the resulting VEA frontier expands only those common facets. If the chosen MPS units do not generate a common facet, then their average will not be DEA efficient and thus its DEA efficient peers would be used as the MPS. The same is true if the DM chooses a DEA-inefficient DMU as the MPS. For example, in Figure 3.1, the average of DMUs B and C lies on facet BC and limits the preferred input/output bundles between rays OB and OC. On the other hand, the average of DMUs A and E, denoted as AE, is DEA-inefficient and its peers (i.e., DMUs B and C) are used as the MPS.

Many of the proposed suggestions for choosing the MPS involve the subjective judgments of a DM. Such examples include using the DM's personal judgments to

Figure 3.1: Composite indicator construction based on BoD and VEA-BoD models



choose the MPS (Halme *et al.*, 2014), choosing the DMU performing the best in a particular model dimension (Marshal and Shortle, 2005), and using interactive multiobjective algorithms (see Halme *et al.*, 1999). This inherent subjectivity makes the task of choosing an MPS less transparent and raises concerns as it might compromise the evaluation process in the case of malevolent DMs who wish to curb the results in favor of certain DMUs.[29]

Nevertheless, there are other, relatively objective alternatives, the use of which can make the MPS choice as transparent as possible from the viewpoint of stakeholders or the public. They can also provide compromise solutions in cases where a DM is absent or unable to point at a preferred DMU and in cases of disagreement among a board of DMs.[30] These include: *first*, averaging inputs and outputs over more than one

---

[29] We emphasize that such subjectivity is also inherent in several stages of the composite indicator construction process, such as the choice of the relevant indicators to be included in the composite and the normalization scheme. It is frequently present in the choice of weight bounds in weight-restricted BoD as well. Thus, malevolent DMs can also choose weight bounds that will curb the BoD efficiency frontier, resulting in an evaluation process that favors certain DMUs.

[30] Some MPS choices might prove to be as time-consuming as the process of choosing weight restriction bounds. Nevertheless, as Korhonen *et al.* (2002) ague, DMs are more keen on simply pointing at a DMU

DMUs selected by the same or different DMs and using the resulting artificial DMU as a compromise solution MPS (Korhonen *et al.*, 2002). *Second,* using a participatory approach such as the Analytic Hierarchy Process (AHP, Saaty, 1988) or the Budget Allocation Process (BAP, see OECD, 2008).[31] *Third,* using an established criterion for ranking the DEA-efficient DMUs such as superefficiency scores (see Halme and Korhonen, 2015). *Fourth,* choosing the MPS on the basis of a strong normative argument, which mandates what the preferred performance of DMUs "ought to be" within the particular evaluation context. This alternative is followed in the empirical application of this chapter, in which MPS choice is based on the notion of uniformity. *Fifth,* using an ideally-performing virtual DMU, i.e., one that utilizes the maximum observed indicator values across DMUs. As such an Ideal DMU usually lies beyond the DEA efficient frontier, its DEA-efficient peers should be identified through a superefficiency model and be used as the MPS its place.[32]

3.3. Re-estimating the Human Development Index

3.3.1. Variables and modeling choices

In this section we use the VEA-BoD model to re-estimate UN's HDI for the year 2015. The HDI is a CI reflecting country achievements in human development, the underpinnings of which can be found in Sen's capability approach.[33] The capability approach views people as the main recipients (the "ends") of the development process and development itself as a process which expands people's choices, thereby placing the emphasis on "what people get from development, not only what they put into it" (Anand and Sen, 2000b, p. 84). The HDI contains three basic and universally valued capabilities, namely to be knowledgeable, to live a healthy life, and to have adequate command over resources in order to enjoy a decent standard of living (Anand and Sen,

---

rather that engaging in the task of choosing weight bounds, meaning that the concept of the MPS is generally easier to understand and to select, compared to absolute or relative weight bounds.

[31] The use of AHP for choosing the MPS was proposed in Korhonen *et al.* (1998).

[32] This alternative is inspired from the multicriteria TOPSIS (Technique for Ordered Similarity to Ideal Solution, see Huang and Yoon, 1981) technique. TOPSIS also involves an Anti-Ideal DMU, namely one utilizing the minimum observed indicator values across DMUs, but such a benchmark choice is not suggested as an MPS as it would be more likely to represent the *least* rather that the *most* preferred solution.

[33] For a recent review of the underpinnings and development of the HDI see Hirai (2017).

2000b). Ever since the first Human Development Report in 1990, there has been a quite long literature regarding (*i*) the choice of the capabilities to be included in the index, (*ii*) their relevant proxy variables, (*iii*) the normalization of these variables, (*iv*) the choice of the aggregator function, and (*v*) the selection of aggregation weights. We next consider these steps in sequence.

First, several important aspects of human development such as environmental sustainability, political rights and freedom, income or gender inequality as well as other demographic factors are not included in the current specification of the HDI (see e.g., Desai (1991), Ranis *et al.* (2006) and Klugman *et al.* (2011)) and several attempts (see e.g., Hicks (1997), Sagar and Najam (1998), Neumayer (2001) and Herrero *et al.* (2019)) have been made to incorporate them. For the purpose of this chapter, we keep the current HDI specification for both the capabilities considered and the variables used to proxy them. That is, we use life expectancy at birth as a proxy for living a healthy life, the arithmetic average of the mean and the expected years of schooling as a proxy for being knowledgeable, and the logarithm of GNI per capita in 2011 $ PPP to a proxy for the standard of living.[34]

Second, the HDI is based on the min-max normalization with the goalposts (minimum and maximum) values for each indicator being those of 1994.[35] This has been criticized as the normalized indicators and the resulting CI depend on the choice of these minimum and maximum values (Noorbakhsh, 1998a; Panigrahi and Sivramkrishna, 2002). Several alternatives have been proposed: in particular, Mazumdar (2003) and Chakravarty (2003) used sample minimum and maximum goalpost values, Noorbakhsh (1998a; b) employed the z-score normalization, Herrero *et al.* (2012) relied on the distance-to-the-leader normalization (i.e., divide each variable with its maximum value across countries), while Luque *et al.* (2016) suggested a normalization with two reference points, an aspiration point reflecting the desired level and a reservation point beyond which performance is not acceptable.[36] For the BoD model, the distance-to-the-leader is the appropriate normalization in order to ensure

---

[34] There is a long discussion in the literature about the logarithmic transformation of the income variable; see Kelley (1991), Chakravarty (2011), Ravallion (2012), and Herrero *et al.* (2012).
[35] Prior to 1994, the goalposts were set by the sample minimum and maximum values.
[36] For comparative results regarding the first three of these normalizations for the HDI see Karagiannis and Karagiannis (2020).

unit invariance, required in any DEA model. Notice that unit invariance is violated with the min-max normalization (Filippetti and Peyrache, 2011).

Third, the UNDP initially (1990-2009) used an arithmetic aggregation function but as of 2010 it has switched to a geometric aggregation function.[37] The main reason behind this switch is the implied perfect substitutability between the component indicators in the arithmetic aggregation (see Klugman *et al.* (2011) and the references therein). However, as argued by Ravallion (2012), arithmetic aggregation implies perfect substitutability between the considered indicators but not between capabilities due to the logarithmic transformation of the income variable. Both the BoD and the VEA-BoD models assume arithmetic aggregation of the component indicators. Previous attempts to estimate the HDI by means of a multiplicative BoD model, with the logarithm instead of the actual values of the component indicators, have been criticized (Tofallis, 2014) as they do not satisfy unit invariance.[38]

Fourth, choice of the weights for aggregating the component indicators is probably the most debated step. The UNDP used equal weights, which implies that each indicator and its corresponding capability are of equal importance to human development (Klugman *et al.*, 2011). Even though this has been criticized as arbitrary (see e.g., Desai, 1991), there are several studies that support the equal weights scheme either on the basis of the principle of parsimony (Hopkins, 1991) or empirical evidence based on Principal Components Analysis (Owgang, 1994; Noorbakhsh, 1998a, b; Owgang and Abdou, 2003; Nguefack-Tsangue *et al.*, 2011), expert opinion surveys (Chowdhury and Squire, 2006), or statistical criteria from Information Theory (Stapleton and Garrod, 2007). On the other hand, several other studies have called for variable weights: Srinivasan (1994, p. 240) noted that relative weights "need not be the same across individuals, countries, and socioeconomics groups". Along the same line, Fukunda-Parr (2003, p.306) referred that "the relative importance of capabilities can

---

[37] Sagar and Najam (1998), Prados de la Escosura (2010), Herrero *et al.* (2010), and *Zhou et al.* (2010) have also used geometric aggregation while Noorbakhsh (1998a; b) used the L_2 distance of each country from an ideal country that has the sample maximum value of indicators, Luque *et al.* (2016) and Krishnakumar (2018) set the HDI equal to the minimum of the three indicators (a scheme that allows for no substitutability), and Noorbakhsh (1998a) used the Borda's aggregation rule.

[38] Tofallis (2013) used a multiplicative BoD model that satisfies both unit and scale invariance but which, according to van Puyenbroeck and Rogge (2017), can be no longer considered as a geometric weighted average of indicators, as it violates the linear homogeneity property.

vary with social context-from one community or country to another and from one point of time to another", Klugman *et al.* (2011, p. 261) suggested that in an ideal situation the relative weights "should be traced either to individual preferences, some collective social choice process or to a strong normative argument'', and Noorbakhsh (1998a, p. 593) argued that "an alternative way is to derive these weights from the data".

Following these considerations, several studies favored the use of variable weights that vary either across countries or across both indicators and countries. In the former case, Lind (2010) used revealed preferences to obtain such a set of common across countries but unequal weights, Pinar *et al*. (2013, 2017) employed non-parametric stochastic dominance techniques, Karagiannis and Karagiannis (2020) relied on Shannon entropy, Despotis (2005a), Hatefi and Torabi (2010) and Sayed *et al*. (2015) used goal programming, and Tofallis (2013) a regression model with the CI obtained from the conventional BoD model as dependent variable and the component indicators as independent variables. [39] In the latter case, Mahlberg and Obensteiner (2001) used a normalized variant of the BoD model, Despotis (2005) employed the conventional BoD model in (3.1), Bougnol *et al*. (2010) considered a BoD model with weighted restrictions, and Lozano and Gutierrez (2008) relied on the range-adjusted-measure (RAM) BoD model.[40]

The VEA-BoD model used in this chapter allows weights to vary across both indicators and countries but only within a certain range, which is determined by the "model" country. This is in line with Sen (1999, p. 78) who mentioned that weights for each capability can be chosen from a specified range on which there is agreement. We base our choice of the "model" country on the notion of uniformity. Following Mishra and Nathan (2018), uniformity implies that, between two countries with the same average attainment across indicators, the CI should favor the most balanced country, i.e., the country with the minimum dispersion across indicators. Palazzi and Lauri (1998, p.196) also favored such a choice by postulating that "there are explicit or potential endogenous forces working to move the values of the single variables towards a more balanced relation". In Figure 3.1, by choosing a country such as C, which

---

[39] The procedure is repeated in Lind (2019) using world data for the years 1990-2017. The findings differentiate from the 2010 study in that the weight of income is now the lowest.

[40] In the normalized variant of the BoD model, $\sum_{j=1}^{J} u_j^k = 1$ in addition to other restrictions in (3.1).

displays balanced performance, implies that the preferred range of input/output bundles lies between rays OB and OD. For any country within this range, the conventional and the VEA-BoD model scores coincide while the farther a country's bundle is from those between OB and OD rays, the lower its VEA-BoD score will be compared to its BoD score.

We consider three alternatives for choosing a balanced MPS country: *first,* the country that is ranked first in 2015 UN HDI, namely Norway; *second*, the country with the minimum dispersion across indicators, namely Lithuania; and *third*, an artificial country with all indicators set at 0.5.[41] Norway is also a BoD-efficient country and thus it can serve as MPS by its own. The other two alternatives, namely Lithuania and the artificial country, are BoD-inefficient but share the same peers, namely Norway and Australia, and thus result in the same VEA-BoD model.

3.3.2. Empirical Results

The empirical results for the conventional BoD model in (3.1) and the two VEA-BoD models with Norway and Norway and Australia as MPS are presented in Table 3.1. The average CI value for the BoD model is 0.861, with five countries receiving CI scores of one, namely Norway, Australia, Singapore, Hong-Kong and Qatar.[42] As it was expected, VEA-BoD results on average into relatively lower scores and less DMUs as being efficient. From the BoD-efficient countries, Hong-Kong drops from the list when Norway is chosen as MPS while Hong-Kong and Qatar drop from the list when Norway and Australia are chosen as MPS. Qatar had an extremely unbalanced bundle that implicitly places higher importance on the "command over resources" indicator, for which it ranks 1st compared to education (82nd) and longevity (39th). Hong-Kong, on the other hand, implicitly places a higher importance on the longevity indicator, for which it ranks 1st (see Table 3.2).

---

[41] The artificial country with all indicators set at 0.5 is a multiple of the "Ideal DMU" country, for which all indicator values are equal to one. Hence, the radial projection of the Ideal DMU on the efficient frontier and, consequently, its DEA-efficient peers, coincide with those of the artificial country (i.e., Norway and Australia). Thus, the use of an "Ideal DMU" country as the MPS will produce the same results with our second and third proposed alternatives.

[42] Notice that, as Karagiannis (2017) has shown, the average accurately reflects the aggregate in the case of the BoD and thus, the numbers in the following Tables and Figures can be seen as aggregate values.

Table 3.1: CI estimates and efficient countries, BoD and VEA models

Model	BoD	VEA (a)	VEA (b/c)
composite indicator estimates			
maximum	1.000	1.000	1.000
minimum	0.611	0.574	0.574
average	0.861	0.834	0.833
median	0.880	0.856	0.856
standard deviation	0.094	0.106	0.106
Q1	0.793	0.748	0.747
Q3	0.926	0.913	0.913
efficient countries			
# of efficient countries (CI=1)	5	4	3
efficient country names	Norway, Australia, Singapore, Hong-Kong, Qatar	Norway, Australia, Singapore, Qatar	Norway, Australia, Singapore

The HDI frequency distributions of the three estimated models are portrayed in Figure 3.2. Based on Banker tests (see e.g., Banker and Natarajan, 2011) we can confirm that both the VEA-BoD distributions of efficiency scores differ, in a statistically significant way, from that of the conventional BoD. The same is not however true when we are comparing the efficiency scores from the two VEA-BoD models to each other, for which there are no statistically significant differences. This is also evident from the average rank shift, given as $R = (\frac{1}{N}) \sum_{j=1}^{N} \left| rank_1^j - rank_2^j \right|$ (Saisana *et al.*, 2005), which is roughly 1.4 positions when we are comparing the scores of the two VEA-BoD models while it is around 9 positions when we comparing the BoD and the VEA-BoD efficiency scores (see Table 3.3). In addition, relatively large rank shifts (more than ten positions) occur for 70 and 68 countries respectively when we are comparing the BoD with the two VEA-BoD scores, whereas it is limited to only three countries when comparing the two VEA-BoD scores. This rank variability is all but uniform across countries: country-specific Mean Absolute Deviation in ranks (Ḉillingirtürk and Koçak, 2018) in Figure 3.2(b) shows that countries in the middle rank positions, as identified by the BoD model, exhibit relatively higher rank variability compared to top or bottom ranked countries. In order to verify the latter, we constructed rolling country subsamples of size 40. More specifically, the first subsample consisted of the top-40 ranked countries by the BoD model. From that, we constructed the second subsample by dropping the country ranked 1st and including the country ranked 41th. Each following subsample was constructed likewise, and the last one consisted of the

Table 3.2: HDI and indicator values and ranks, BoD-efficient DMUs

DMU	Longevity	education index	income	HDI 2015	times as peer[b]
Norway	0.971 (17)	0.905 (6)	0.945 (6)	0.949 (1)	23
Australia	0.981 (9)[a]	1.000 (1)	0.906 (20)	0.939 (2)	49
Singapore	0.989 (4)	0.803 (37)	0.957 (2)	0.925 (6)	54
Hong-Kong (HK)	1.000 (1)	0.811 (30)	0.926 (10)	0.917 (12)	145
Qatar	0.931 (39)	0.689 (82)	1.000 (1)	0.856 (35)	30

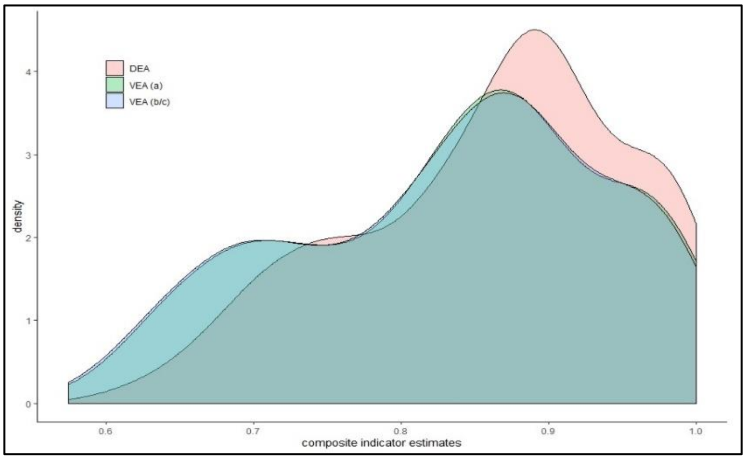
Notes: (a). Numbers in parentheses denote the country's rank position, (b). The last column denotes the time that each efficient country serves as a peer for inefficient ones in model (2) calculations.

bottom-40 ranked countries by the BoD model. The average rank shift between pairs of models for each subsample is plotted in Figure 3.2(c), where we see that the average rank variability between the BoD and the two VEA-BoD scores is considerably higher in subsamples including mostly middle-ranked countries, whereas this pattern is absent when we are comparing the two VEA-BoD scores, which on average displays minor rank differences.

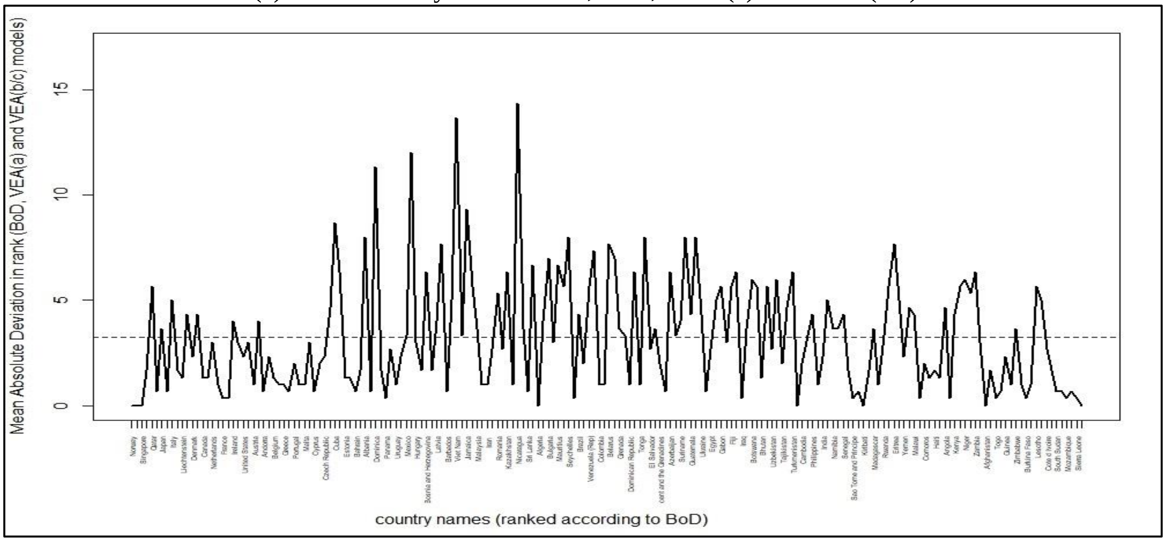
The above results suggest that the VEA-BoD model has a moderate impact on HDI scores compared to the conventional BoD model but a significant impact on country rankings, which is magnified for middle-ranked countries. This finding may however be affected by the choice of MPS, for which so far we have based on the notion of uniformity, i.e., relatively balanced achievements. We next examine the sensitivity of our results to MPS choices that go beyond balanced achievements. In the absence of a general consensus, potential candidates for MPS might be all countries found to be BoD-efficient: namely, Norway, Australia, Singapore, Hong-Kong and Qatar. Summary results of the VEA models using each or combinations of the above countries as MPS are given in Table 3.4 and their frequency distributions are portrayed in Figure 3.3, where are plotted against the BoD distribution of efficiency scores.

Consider first the cases where each of the BoD-efficient countries is chosen as the MPS. The differences between the BoD and the VEA-BoD scores depends on the extent of the preferred range of indicator bundles implied by each MPS, which in turn is closely related to the number of times an efficient country is used as a peer. For example, in the case of Hong-Kong, which serves as a peer for 145 of the 193 BoD-inefficient countries (see Table 3.2), the differences between the conventional BoD HDI and the VEA-BoD HDI using Hong-Kong as the MPS are minimal (see Figure 3.3) and in fact, statistically insignificant (see Table 3.5). The same is essentially true when

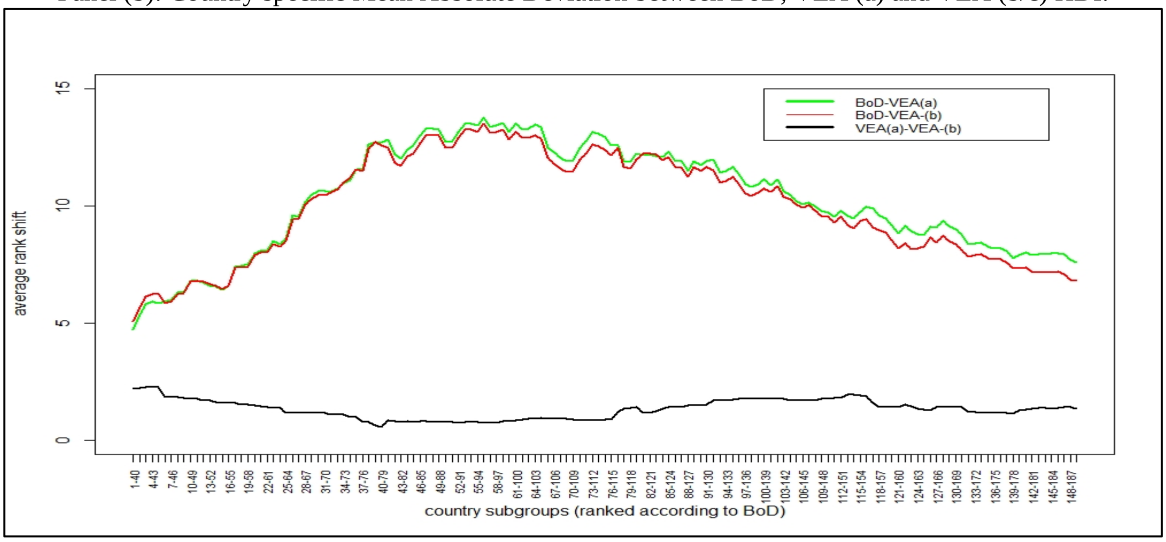
Figure 3.2: Comparison of distributions and rankings, BoD, VEA (a) and VEA (b/c) HDI



Panel (a): kernel density distributions, BoD, VEA (a) and VEA (b/c) HDI.



Panel (b): Country specific Mean Absolute Deviation between BoD, VEA (a) and VEA (b/c) HDI.



Panel (c): Average rank shift between pairs of models, moving country subsamples (n=40)

Table 3.3: Average rank shifts, large rank shifts, and statistical tests of equality between pairs of BoD, VEA(a) and VEA(b/c) models

Pair	Average rank shift	Large rank shifts (>10 positions)	Banker's F1	Banker's F2
BoD-VEA(a)	9.202	70	1.214**	1.438***
BoD-VEA(b/c)	8.952	68	1.229**	1.467***
VEA(a)-VEA(b/c)	1.388	3	1.012	1.020

Note: The F1 (F2) test assumes an exponential (half-normal) distribution of the inefficiency scores, following Banker et al (2010). Both tests compare the initial DEA to the respective VEA distribution and the alternative hypothesis for all cases was that the respective VEA model exhibits higher inefficiency scores. Three, two and one stars denote statistical significance at 1%, 5% and 10% respectively.

Australia (which serves as a peer 49 times) or Singapore (which serves as a peer 54 times) are considered as the MPS.[43] On the other hand, MPS choices such as Norway and Qatar result in statistically significant differences between the conventional BoD and the VEA-BoD models (see Figure 3.3 and Table 3.5). Norway's bundle displays, as we have mentioned, a relatively high balance that is absent from the majority of countries while Qatar's mix placed considerably higher importance on the "command over resources" indicator.[44] In Figure 3.4, we plot the values of Mean Absolute Deviation for the VEA-BoD models using each BoD-efficient country as the MPS. The average value of 4.59 (dashed line) indicates that varying the MPS can induce relatively moderate shifts in ranking. Nevertheless, rank variability appears to be higher for countries in the middle of the rankings, whereas top and bottom ranked countries appear to be relatively less affected by the chosen MPS.

When considering cases with jointly efficient pairs and triads of BoD-efficient countries as MPS (see Table 3.4 and Figure 3.3), several interesting findings emerge from these results: *first*, VEA-BoD models with two or three countries forming the MPS resemble more or less the behavior of the country with the most extreme indicator bundle. See for example the VEA-BoD models based on Norway alone and on Norway and Hong-Kong as the MPS. *Second*, VEA-BoD models with pairs and triads of

[43] In terms of Figure 3.1, we may think of Hong-Kong as being DMU A, for which the preferred range of mixes (between the I_2 axis and OB) is very wide and unbalanced. On the other hand, we may think of Singapore and Australia as being DMUs B and D respectively, the preferred mix ranges of which are slightly less wide but relatively more balanced compared to that of DMU A.

[44] In terms of Figure 3.1, we may think of Norway as being DMU C that displays the most balanced performance but has a relatively narrow preferred mix range, potentially serving as a peer for a few inefficient DMUs and of Qatar as being DMU F whose mix favors extremely indicator 2.

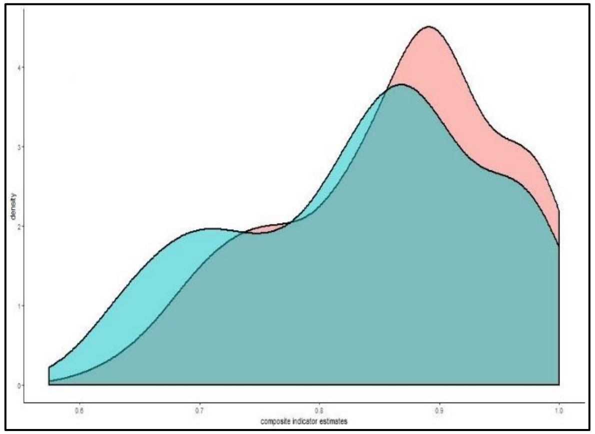
Table 3.4: Composite indicator estimates, model (3.2), alternative MPS selections

MPS selection	maximum	minimum	average	median	standard deviation	Q1	Q3	efficient countries (CI=1)
Norway	1.000	0.574	0.834	0.856	0.106	0.748	0.913	4
Australia	1.000	0.596	0.852	0.875	0.100	0.780	0.923	4
Singapore	1.000	0.605	0.852	0.872	0.096	0.786	0.921	5
Hong-Kong (HK)	1.000	0.611	0.856	0.875	0.097	0.790	0.924	3
Qatar	1.000	0.589	0.836	0.850	0.099	0.759	0.905	3
Norway-Australia	1.000	0.574	0.833	0.856	0.106	0.747	0.913	3
Norway-Singapore	1.000	0.574	0.833	0.856	0.106	0.747	0.913	4
Norway-Qatar	1.000	0.565	0.823	0.840	0.109	0.739	0.905	3
Australia-Singapore	1.000	0.589	0.843	0.864	0.103	0.763	0.916	4
Australia-HK	1.000	0.596	0.848	0.871	0.102	0.772	0.917	3
Singapore-HK	1.000	0.605	0.848	0.866	0.099	0.776	0.918	3
Singapore-Qatar	1.000	0.589	0.835	0.850	0.099	0.757	0.905	3
Norway-Australia-Singapore	1.000	0.574	0.831	0.856	0.107	0.747	0.913	3
Norway-Singapore-Qatar	1.000	0.565	0.822	0.840	0.109	0.729	0.905	3
Australia-Singapore-HK	1.000	0.589	0.841	0.862	0.104	0.762	0.914	3

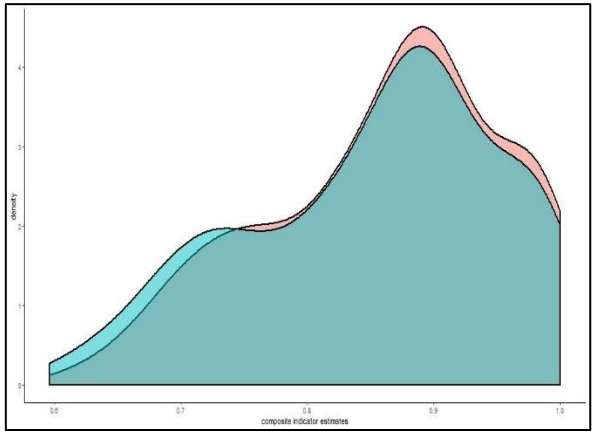
countries constituting the MPS do not differ in a statistically significant sense with VEA-BoD models with the most extreme (in terms of indicator bundle) of these countries as the MPS but they statistically differ from VEA-BoD models with other BoD-efficient countries as the MPS if the pair or triad includes a country with a relatively extreme bundle compared to the rest of the sample (i.e., Norway and Qatar) (see Table 3.6). *Third*, pairing countries with similar preferred ranges of indicator bundles to form the MPS (e.g., Hong-Kong and Singapore) seems to result in negligible differences compared to the VEA-BoD models with each of these countries as a single MPS (see Figure 3.3 and Table 3.6).

This demonstrated sensitivity of the models' estimates to the chosen MPS might pose difficulties to select among alternative evaluation results those that will be ultimately presented to stakeholders or the public and used for policy-designing purposes. As this situation is similar to the initial choice of the MPS, a first option for indecisive practitioners or DMs would be to use the evaluation results stemming from an objective and transparent MPS choice among those presented in the previous section, such as AHP or BAP. A second option would be to choose the evaluation results that fit the most the DMs' perceptions of "good" and "bad" performing DMUs in the sample. Lastly, MPS choice can also be based on the variability between the BOD and VEA-BoD estimates. For example, DMs opting for the least (most) rank variability between HDI estimates of the two models would select the evaluation results based on

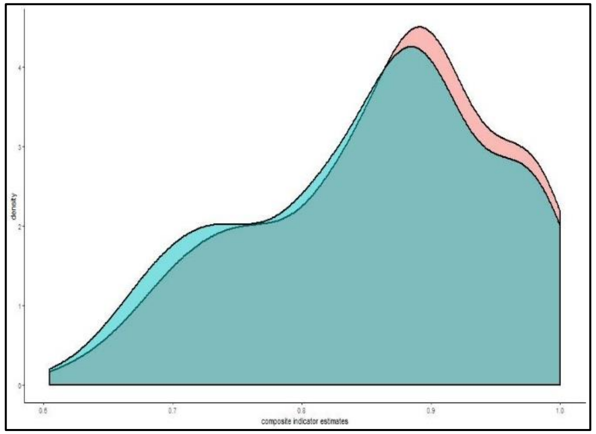
Figure 3.3: Kernel density distributions, model (3.1) vs. model (3.2), alternative MPS choices



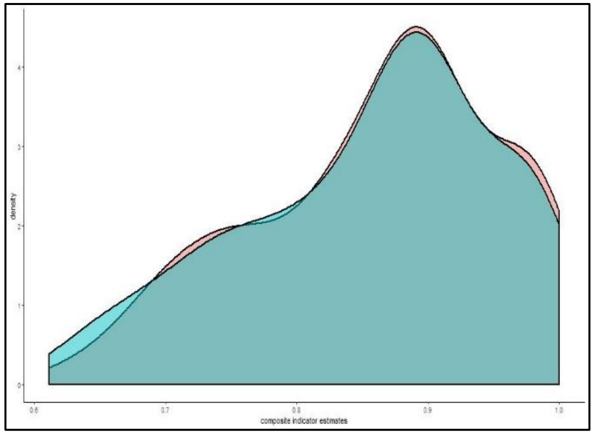
(1) Norway



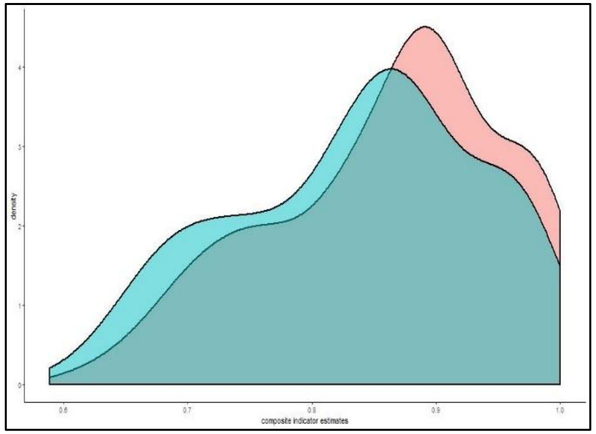
(2) Australia



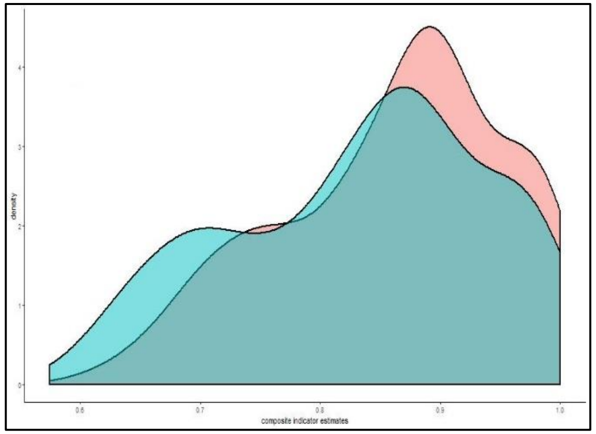
(3) Singapore



(4) Hong-Kong (HK)

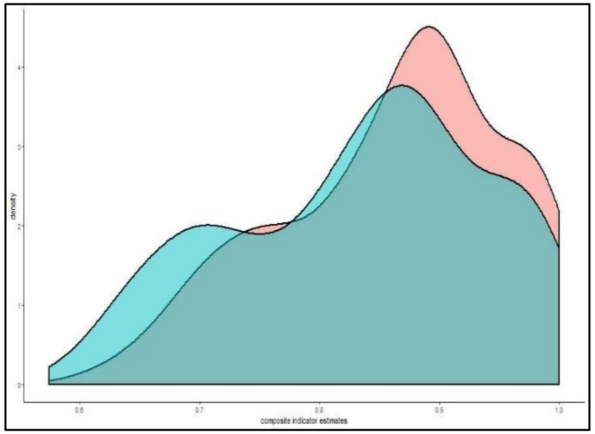


(5) Qatar

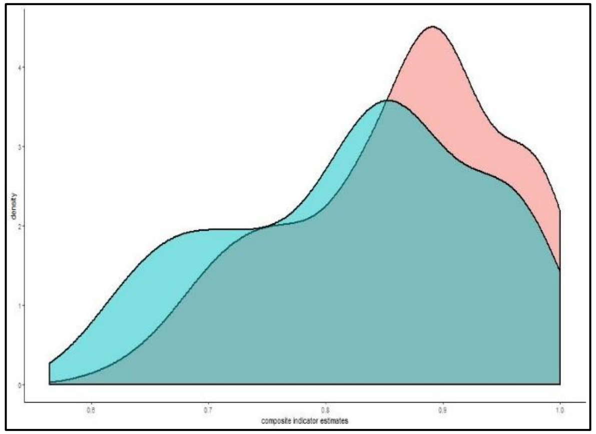


(6) Norway-Australia

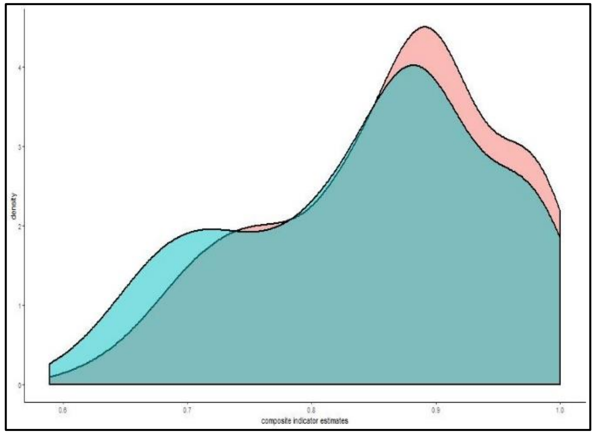
Figure 3.3 (cont.)



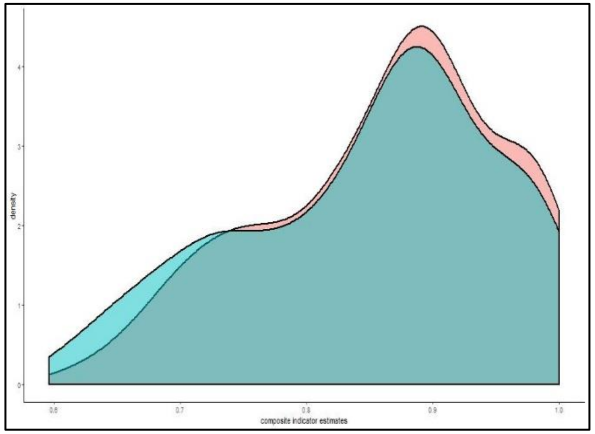
(7) Norway-Singapore



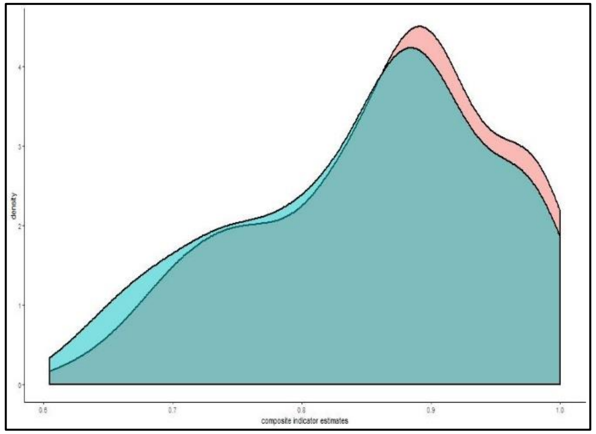
(8) Norway-Qatar



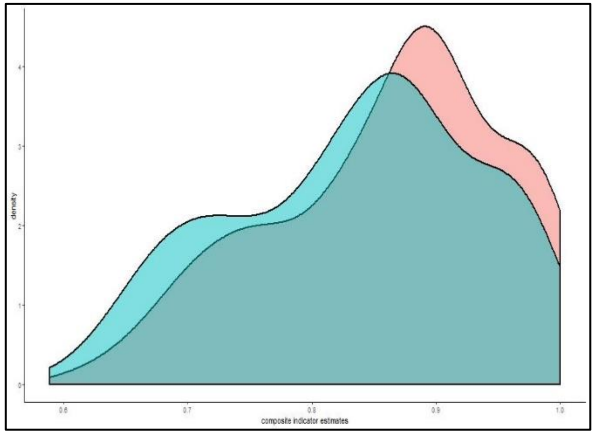
(9) Australia-Singapore



(10) Australia-HK

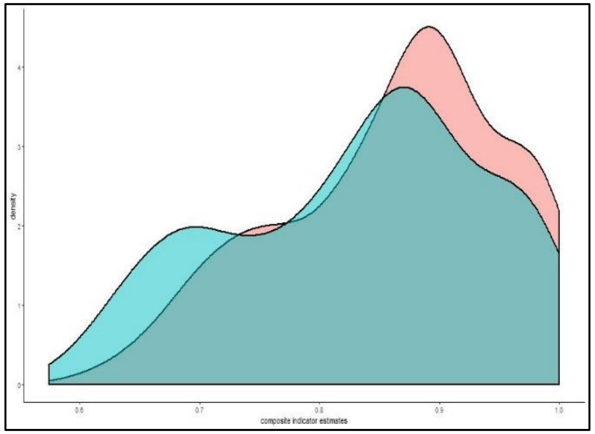


(11) Singapore-HK

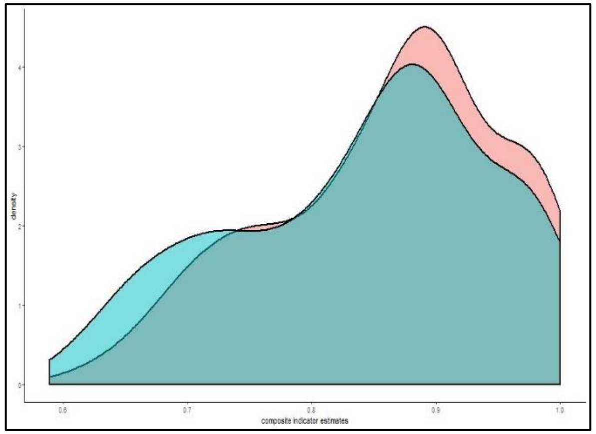


(12) Singapore-Qatar

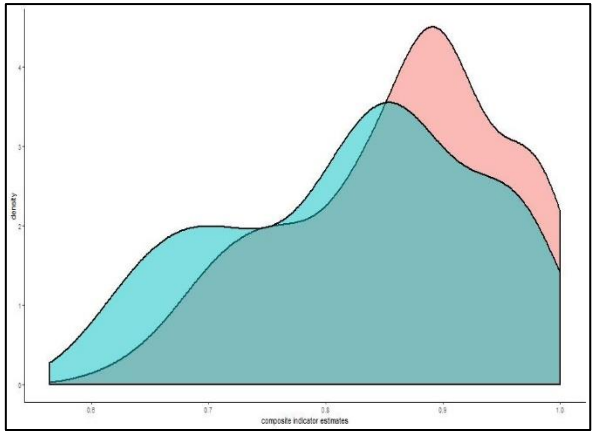
Figure 3.3 (cont.)



(13) Norway-Australia-Singapore



(14) Australia-Singapore-HK



(15) Norway-Singapore-Qatar

Table 3.5:Average rank shifts, large rank shifts, and statistical tests of equality, model (3.1) vs. model (3.2) for alternative MPS selections

MPS selection	Average rank shift	Large rank shifts (>10 positions)	Banker's F1	Banker's F2
Norway	9.202	70	1.214**	1.438***
Australia	3.362	4	1.074	1.159
Singapore	4.532	13	1.068	1.118
Hong-Kong (HK)	4.144	14	1.043	1.091
Qatar	10.537	84	1.195**	1.341**
Norway-Australia	8.952	68	1.229**	1.467***
Norway-Singapore	8.872	66	1.222**	1.452***
Norway-Qatar	12.559	88	1.309***	1.630***
Australia-Singapore	6.128	30	1.148*	1.300**
Australia-HK	4.963	14	1.103	1.222*
Singapore-HK	6.133	24	1.101	1.190
Singapore-Qatar	10.101	80	1.204**	1.360**
Norway-Australia-Singapore	8.654	65	1.241**	1.495***
Norway-Singapore-Qatar	12.271	87	1.316***	1.644***
Australia-Singapore-HK	6.899	34	1.163*	1.334**

Note: The F1 (F2) test assumes an exponential (half-normal) distribution of the inefficiency scores, following Banker and Natarajan (2011). Both tests compare the initial DEA to the respective VEA distribution and the alternative hypothesis for all cases was that the respective VEA model exhibits higher inefficiency scores. Three, two and one stars denote statistical significance at 1%, 5% and 10% respectively.

Australia (Norway and Qatar) as the MPS.

Last but not least, we examine the sensitivity of our results with respect to different modeling choices regarding the education indicator. Several authors (e.g., Mahlberg and Obensteiner, 2001; Lozano and Gutiérrez, 2008; Sayed *et al.*, 2015) suggested using the mean and the expected years of schooling as separate indicators while Herrero *et al*. (2012) proposed using only the expected years of schooling.[45] The comparative results concerning these two alternative formulations of the education variable are presented in Table 3.7. There seem to be no significant differences with our benchmark formulation of using the average of the mean and the expected years of schooling. The most notable difference is that now the VEA-BoD models are based on different countries for the MPS, namely Norway, Australia and Singapore and Norway

[45] The former choice is also supported by empirical findings indicating that using the average of the two variables results in substantial information loss (Canning *et al*., 2013).

Figure 3.4: Country-specific Mean Absolute Deviation between VEA models based on different BoD-efficient DMUs as the MPS

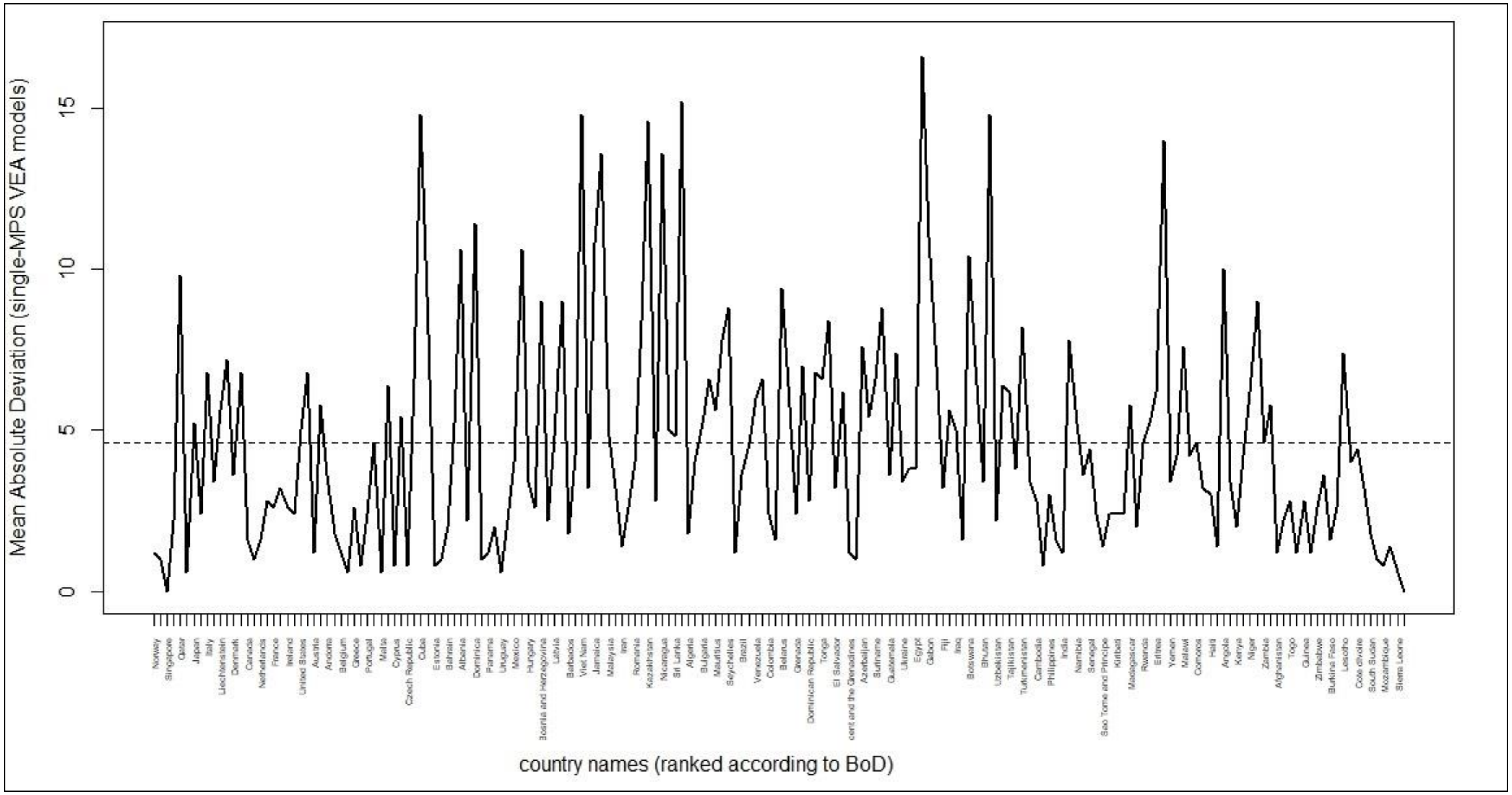


Table 3.6: Average rank shifts and statistical tests of equality, model (3.2) for alternative MPS selections

	average rank shift														
MPS selection	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
1.Norway		8.617	6.170	11.452	4.016	1.388	0.830	3.835	5.596	10.122	8.346	4.314	2.559	3.665	6.771
2.Australia			4.660	4.505	10.463	7.888	8.213	12.197	4.702	2.814	5.229	9.973	7.580	11.888	5.186
3.Singapore				6.261	7.037	5.739	5.766	9.186	3.202	5.654	3.133	6.388	5.090	8.846	3.984
4.Hong-Kong (HK)					12.777	10.649	11.037	14.745	6.633	2.777	4.309	12.170	9.968	14.426	5.691
5.Qatar						4.702	4.335	3.883	7.548	12.053	9.117	1.074	5.383	3.755	8.511
6.Norway-Australia							1.239	4.723	4.686	9.309	7.447	4.681	1.266	4.457	5.777
7.Norway-Singapore								4.303	5.074	9.686	7.910	4.080	1.803	3.931	6.229
8.Norway-Qatar									8.920	13.670	11.160	4.404	5.511	0.681	9.883
9.Australia-Singapore										5.250	3.676	6.878	3.878	8.559	1.516
10.Australia-HK											3.894	11.447	8.649	13.351	4.298
11.Singapore-HK												8.383	6.649	10.777	2.723
12.Singpore-Qatar													4.755	4.032	7.798
13.Norway-Australia-Singapore														5.117	4.936
14.Norway-Singapore-Qatar															9.511
15.Australia-Singapore-HK															
	Banker's F1 test														
1.Norway		1.130	1.137	1.164*	1.016	1.012	1.006	1.078	1.058	1.101	1.103	1.008	1.022	1.083	1.044
2.Australia	1.241*		1.006	1.030	1.112	1.143*	1.137	1.218**	1.068	1.026	1.025	1.121	1.155*	1.224**	1.083
3.Singapore	1.286**	1.037		1.024	1.119	1.150*	1.144*	1.225**	1.075	1.033	1.031	1.128	1.162*	1.232**	1.089
4.Hong-Kong (HK)	1.318**	1.063	1.025		1.146*	1.178*	1.171*	1.255**	1.100	1.057	1.056	1.155*	1.190**	1.261**	1.115
5.Qatar	1.072	1.157	1.199	1.229*		1.028	1.022	1.095	1.041	1.084	1.085	1.008	1.039	1.101	1.027
6.Norway-Australia	1.020	1.265*	1.312**	1.345**	1.094		1.006	1.065	1.071	1.114	1.116	1.020	1.010	1.071	1.056
7.Norway-Singapore	1.010	1.253*	1.299**	1.331**	1.083	1.01		1.071	1.065	1.108	1.110	1.014	1.016	1.077	1.050
8.Norway-Qatar	1.133	1.406***	1.458***	1.494***	1.215*	1.111	1.122		1.140	1.187*	1.188**	1.086	1.054	1.005	1.125
9.Australia-Singapore	1.107	1.121	1.162	1.192	1.032	1.128	1.117	1.254*		1.041	1.042	1.049	1.082	1.146*	1.014
10.Australia-HK	1.177	1.054	1.092	1.120	1.098	1.201	1.189	1.334**	1.064		1.002	1.092	1.126	1.193**	1.055
11.Singapore-HK	1.209*	1.026	1.064	1.091	1.127	1.233*	1.220*	1.370**	1.092	1.027		1.094	1.127	1.195**	1.056
12.Singpore-Qatar	1.058	1.173	1.216*	1.247*	1.014	1.078	1.068	1.198	1.046	1.113	1.143		1.031	1.092	1.035
13.Norway-Australia-Singapore	1.039	1.289**	1.337**	1.37**	1.115	1.019	1.029	1.090	1.150	1.224*	1.256*	1.099		1.060	1.067
14.Norway-Singapore-Qatar	1.143	1.418***	1.470***	1.507***	1.226*	1.121	1.132	1.009	1.265*	1.346**	1.381**	1.209*	1.100		1.131
15.Australia-Singapore-HK	1.078	1.151	1.193	1.223*	1.005	1.100	1.089	1.222*	1.026	1.092	1.121	1.020	1.121	1.232*	
	Banker's F2 test														

Note: The F1 (F2) test assumes an exponential (half-normal) distribution of the inefficiency scores, following Banker and Natarajan (2011). The tests in this table compare the efficiency distributions of the respective row and column VEA models. Three, two and one stars denote statistical significance at 1%, 5% and 10% respectively.

Table 3.7: Robustness tests: Alternative formulations for the education indicator

	formulation		
	average of two education variables	two separate education variables	only expected years of schooling
	BoD		
Average	0.861	0.864	0.861
# of countries with CI=1	5	8	5
	VEA (a) (MPS: country ranked the highest in the official HDI-ranking of 2015)		
MPS	Norway	Norway	Norway
Average	0.834	0.840	0.836
# of countries with CI=1	4	6	4
	VEA (b) (MPS: country with minimum dispersion)		
MPS	Lithuania*	Ireland*	Ireland*
	(Norway-Australia)	(Norway-Australia-Singapore)	(Norway-Australia-Singapore)
Average	0.833	0.836	0.833
# of countries with CI=1	3	4	3
	VEA (c) (MPS: virtual country with all normalized indicators equal to 0.5*)		
MPS	(Norway-Australia)	(Norway-Australia)	(Norway-Australia)
Average	0.833	0.839	0.835
# of countries with CI=1	3	4	3

Note: An asterisk denotes an inefficient country based on the BoD model. The countries in parentheses below it are its efficient peers and are used in its place as the MPS.

and Australia. Nevertheless, this change affects only slightly the HDI values.

3.4. Concluding Remarks

In this chapter, we use the VEA formulation of the BoD model, which integrates DM or expert opinion to the conventional BoD model through the choice of a “model“ DMU that serves as benchmark for all evaluated units. The “model” DMU defines a preferred range of indicator bundles and for DMUs operating within (outside of) this preferred range, VEA-BoD scores are equal to (lower than) the BoD scores. The proposed model is sensitive to the choice of the MPS. Models with MPS BoD-efficient units featuring a wider range of indicator bundles (as indicated by the times they are used as peers) or with bundles closer to the majority of the evaluated DMUs result in VEA-BoD scores that differ less from the BoD scores. In addition, VEA-BoD scores tend to differ more (less) from each other if their chosen MPSs have highly dissimilar (similar) bundles, while VEA-BoD models with more than one DMU as the MPS resemble closely the pattern of the most extreme of those DMUs. In our empirical application regarding the HDI, the VEA-BoD model causes moderate changes regarding the scores but significant changes in country rankings compared to the conventional BoD model, especially for middle-ranked countries which displayed on average higher rank variability compared to top and bottom performing countries.

The proposed model can be applied to a wide range of social and economic indicators and it could be useful to both the evaluated entities as well as DMs, since it allows pursuing the best-possible aggregation weights but to the extent that these weights comply with managerial goals. Nevertheless, the proposed model is not without limitations, as its current form inherits certain deficiencies of the conventional BoD and DEA models. More specifically, it is sensitive to the presence of outliers- which could also affect the MPS choice- and it fails to account for the effect of background 'contextual' variables which are not under the direct control of DMUs but can create favorable operating conditions for some of them and unfavorable for others. Hence, the present work could be further extended through a robust order-m framework (see Cazals *et al.*, 2002) to mitigate the impact of outlying observations and through a conditional DEA framework (see Daraio and Simar, 2005) in order to account for the effect of contextual variables. In addition, the present model can be readily extended to cases where DMUs select the worst possible aggregation weights by means of the inverted BoD model. In such a case, managerial goals regarding the *least preferred* indicator bundle would be considered.

CHAPTER 4

In search for the Most Preferred Solution in Value efficiency Analysis

4.1. Introduction

In several occasions where the performance of Decision Making Units (DMUs) is evaluated by means of Data Envelopment Analysis (DEA, Charnes *et al.*, 1978), it is desired or necessary to consider the preferences of central management, supervising agencies or Decision Makers (DMs) that coordinate the operation of DMUs.[46] This need might arise for purposes of performance monitoring, i.e., measuring the extent to which the performance of DMUs complies with overall behavioral or organizational objectives, as well as for performance control and future planning, namely designing mechanisms that redirect DMUs towards the achievement of managerial goals or normative performance standards.[47]

Preferences in DEA studies are often elicited by means commonly used in Multiple Objective Linear Programming (MOLP), namely by incorporating expert information on the desirable input and output values for the evaluated DMUs (Korhonen *et al.,* 2002). One form this might take is that of the Most Preferred Solution (MPS).[48] The MPS is a non-dominated (i.e., strongly DEA-efficient) DMU or a

---

[46] Such centrally managed and coordinated groups of DMUs (which may have either limited or enhanced control over the resources allocated to them, and autonomy in setting their own priorities) might include privately (e.g., bank branches, retail stores) or publicly owned entities (hospitals, education institutions). They may also be DMUs benefiting from a natural monopoly such as large infrastructure industries, e.g., water, electricity and gas networks (Afsharian *et al.,* 2019).
[47] These organizational goals might either be monetary (e.g., profit maximization) or non-monetary, such as targets set for overall output production. Normative standards for performance might arise, for example, from contract agreements signed by a group of DMUs and the supervising agency (Ruiz and Sirvent, 2019).
[48] Other forms might include targets set separately for each DMU, which may correspond to aspiration values or long-term goals set by management (Stewart, 2010).

combination of DMUs, which has the most desirable structure in DM's view, in the sense of maximizing his/her value (Korhonen, 2002) or utility function (Yang *et al.*, 2009).[49] It may represent the structure according to which management in a firm wishes to reorganize its branches or it might be viewed as a mentor from which other DMUs can learn. The MPS was incorporated into DEA by Halme *et al.* (1999), in an approach coined Value Efficiency Analysis (VEA). In VEA, the DMUs are assessed against a frontier consisting of the extended DEA efficient facets intercepting at the MPS, which is chosen by the DM in a prior step. In essence, the marginal rates of substitution (MRSs) of inputs or transformation (MRTs) of outputs imposed on the evaluated DMUs are those observed on the DEA frontier for the MPS.

Choosing the MPS is an important issue in VEA, as it affects the resulting efficiency frontier and, consequently, the DMUs' efficiency scores (Korhonen *et al.*, 2001). A suitably chosen MPS can yield valuable insights regarding the extent to which current DMUs' performance complies with managerial preferences or organizational goals, and provides the basis for a cost-saving or revenue-increasing restructuring. On the other hand, an inappropriate MPS choice might provide questionable efficiency scores, which may subsequently give rise to poor managerial decisions, such as an unnecessary and costly resource reallocation. Nevertheless, there seems to be no general rule for choosing the MPS in VEA. Instead, several suggestions have been made up to date. In many of these, the MPS is not chosen on the basis of some overall managerial objective and thus it is difficult to come up with an intuitive explanation for the DM's choice, while in others the chosen MPS may favor specialization in the production of a few outputs or in the use of a few inputs, which is often deemed unsatisfactory by managers (Epstein and Henderson, 1989). Other MPS choices may compare DMUs with exceptionally performing benchmarks, assess them against a DMU operating with non-technically optimal scale, or zero and undefined values for MRSs and MPTs.[50] In addition, no empirical work has been done so far on how alternatively chosen MPSs may affect the VEA efficiency scores.

---

[49] Strongly DEA-efficient DMUs are those not associated with input or output slacks (Charnes *et al.*, 1985).

[50] These occur when an optimal vector of input/output weights for the MPS in the multiplier form of a DEA model includes zero values (Olesen and Petersen, 2003), i.e., when a weakly efficiency facet is adjacent to the MPS.

The objective of this chapter is twofold: *First,* to expand the set of MPS choices. We first advocate that the DM could make a more informed choice of the MPS among the efficient DMUs, by paying attention on their position on the efficient frontier. In particular, we propose that prior to MPS choice the efficient DMUs are clustered based on whether they appear in the reference sets of other DMUs in DEA and whether they reside in frontier edges (Edvardsen *et al.,* 2008). This clustering can provide additional information to the DM about the DMUs for which there is strong evidence of good relative performance, those that are potentially overspecialized, and those that may be associated with zero and/or undefined marginal rates. Alternatively, one may choose the MPS among those with Most Productive Scale Size (MPSS), i.e., those achieving maximal average productivity for their input/output bundle. This will ensure that the DMUs are assessed against a technically optimal scale, the achievement of which is a long-term organizational goal, which interests both individual DMUs as well as central management (Førsund and Hjalmarsson, 1979). Our third proposed MPS is the combination of peers of the Average Production Unit (APU). The APU is an artificial DMU that operates with the group means quantities of inputs and outputs, and its technical efficiency score reflects the structural efficiency of the whole group of DMUs when resource allocation is centrally coordinated.[51] Its structure reflects the one that each DMU should have in order for the group as a whole to realize its full potential output production, and the resulting VEA scores may be particularly useful for guiding future performance planning. Another proposal is to assess DMUs using on a common vector of strictly positive input/output weights in VEA, by choosing a combination of DMUs that generate a unique Fully Dimensional Efficient Facet (FDEF) as the MPS. This results in evaluating all DMUs against a common standard and well-defined MRSs and MRTs and could be useful in several cases where the assessed DMUs perform essentially the same task or have limited autonomy in setting their own priorities and objectives (i.e., choose individually the values of input/output weights).

The chapter's second objective is to provide comparative empirical evidence on how alternative MPS choices may affect the estimated VEA efficiency scores. More specifically, using data for 526 Greek cotton farms, we compare the efficiency

---

[51] Structural efficiency was termed by Farrell (1957, p. 262) as 'the extent to which an industry keeps up with the performance of its own best firms'.

estimates obtained by the DEA model and those of VEA models with alternative MPS choices. The results of this analysis provide useful insights regarding the MPS choices that are more likely to result in excessive or negligible differences between the DEA and the VEA distributions of efficiency scores.

The rest of the chapter unfolds as follows: In the second section, we present the VEA model while in the third section, we review the MPSs proposed previously in VEA and suggest four new. In the fourth and the fifth section, we illustrate how the choice of the MPS may affect VEA efficiency scores. Concluding remarks follow in the last section.

4.2. Materials and methods

In VEA, DM preferences are reflected through an implicitly known pseudo-concave value function (i.e., an indifference curve), that becomes tangent to the DEA efficient frontier at the point where the MPS is located. This value function might reflect some organizational objective, i.e., be a cost or a profit function, but it might also reveal preferences other than those related with prices (Thanassoulis *et al.*, 2008, p. 73). The empirical VEA frontier is then constructed as the lower envelope of the extended efficient facets intercepting at the MPS. As DEA facets are generated by extreme-efficient DMUs, the MPS will in essence be either a single extreme-efficient DMU or a combination of extreme-efficient DMUs that are *jointly* efficient, in the sense that they generate at least one common facet. In the latter case, only those common efficient facets are extended to obtain the VEA frontier.

Introducing the MPS requires only slight modifications to the conventional DEA model. Let us consider a set of K DMUs $(k = 1, \dots, o, \dots, K)$, that operate under the same technology and use I $(i = 1, \dots, I)$ inputs to produce J $(j = 1, \dots, J)$ outputs. The input and output vectors of each DMU are assumed to be semi-positive, that is, each DMU uses at least one input to produce at least one output. Further, we assume that the DM has select a set $\mathcal{R}$ $(r = 1, \dots, R)$ of extreme-efficient DMUs as the MPS.[52]

[52] Note that the number of extreme-efficient DMUs constituting the MPS cannot be more than (I+J-1) in DEA models with constant returns to scale (CRS) and (I+J) in DEA models with variable returns to scale (VRS), as this is the maximum number of extreme-efficient DMUs that may generate an efficient facet of the DEA surface (see Olesen and Petersen, 2003).

An output-oriented, variable-returns-to-scale (VRS) VEA model in its multiplier and envelopment form is given as (Halme and Korhonen, 2015):

$$
\begin{aligned}
&\min_{u_j^o, v_i^o, u^o} \sum_{i=1}^{I} v_i^o x_i^o - u^o \\
&s.t.\ \ -\sum_{j=1}^{J} u_j^o y_j^k + \sum_{i=1}^{I} v_i^o x_i^k - u^o \geq 0 \quad k = 1, \dots, K, k \neq r \\
&\quad\ \ -\sum_{j=1}^{J} u_j^o y_j^r + \sum_{i=1}^{I} v_i^o x_i^r - u^o = 0 \quad r = 1, \dots, R \\
&\quad\ \ \sum_{j=1}^{J} u_j^o y_j^o = 1 \\
&\quad\ \ u_j^o \geq 0 \qquad j = 1, \dots, J \\
&\quad\ \ v_i^o \geq 0 \qquad i = 1, \dots, I \\
&\quad\ \ u^o\ free
\end{aligned}
\qquad
\begin{aligned}
&\max_{\varphi_{VEA}^o, \lambda_k^o} \ \varphi_{VEA}^o \\
&s.t.\ \sum_{k=1}^{K} \lambda_k^o y_j^k \geq \varphi_{VEA}^o y_j^o \quad j = 1, \dots, J \\
&\quad\ \ \sum_{k=1}^{K} \lambda_k^o x_i^k \geq x_i^o \qquad i = 1, \dots, I \\
&\quad\ \ \sum_{k=1}^{K} \lambda_k^o = 1 \\
&\quad\ \ \lambda_k^o \geq 0 \qquad k = 1, \dots, K, k \neq r \\
&\quad\ \ \lambda_r^o\ free \qquad r = 1, \dots, R \\
&\quad\ \ \varphi_{VEA}^o\ free
\end{aligned}
\qquad (4.1)
$$

where *x* and *y* are input and output quantities, $1/\varphi_{VEA} \in (0,1]$ is the efficiency score, λ the intensity variables, and v, u and u^o are parameters to be estimated. The input-oriented counterpart of (4.1) is given as:

$$
\begin{aligned}
&\max_{u_j^o, v_i^o, u^o} \sum_{j=1}^{J} u_j^o y_j^o + u^o \\
&s.t.\ \ \sum_{j=1}^{J} u_j^o y_j^k - \sum_{i=1}^{I} v_i^o x_i^k + u^o \leq 0 \quad k = 1, \dots, K, k \neq r \\
&\quad\ \ \sum_{j=1}^{J} u_j^o y_j^r - \sum_{i=1}^{I} v_i^o x_i^r + u^o = 0 \quad r = 1, \dots, R \\
&\quad\ \ \sum_{i=1}^{I} v_i^o x_i^o = 1 \\
&\quad\ \ u_j^o \geq 0 \qquad j = 1, \dots, J \\
&\quad\ \ v_i^o \geq 0 \qquad i = 1, \dots, I \\
&\quad\ \ u^o\ free
\end{aligned}
\qquad
\begin{aligned}
&\min_{\theta_{VEA}^o, \lambda_k^o} \ \theta_{VEA}^o \\
&s.t.\ \sum_{k=1}^{K} \lambda_k^o y_j^k \geq y_j^o \qquad j = 1, \dots, J \\
&\quad\ \ \sum_{k=1}^{K} \lambda_k^o x_i^k \geq \theta_{VEA}^o x_i^o \quad i = 1, \dots, I \\
&\quad\ \ \sum_{k=1}^{K} \lambda_k^o = 1 \\
&\quad\ \ \lambda_k^o \geq 0 \qquad k = 1, \dots, K, k \neq r \\
&\quad\ \ \lambda_r^o\ free \qquad r = 1, \dots, R \\
&\quad\ \ \theta_{VEA}^o\ free
\end{aligned}
\qquad (4.2)
$$

where $\theta_{VEA} \in (0,1]$ is the efficiency score. The constant-returns-to-scale (CRS) form of (4.1) and (4.2) are obtained by removing the free variable and the convexity constraint from their multiplier and envelopment forms, respectively.[53]

---

[53] The CRS counterpart of (4.2), in its multiplier form, appears for the first time in Oral and Yolalan (1990) and Oral *et al.* (1992), where it is used to compare every DMU's performance to that of a particular efficient DMU, which is selected at a previous step.

The envelopment form of the models in (4.1) and (4.2) differs from those of conventional DEA in that the sign of the intensity variable corresponding to the MPS is free instead of restricted to be non-negative (Halme *et al.,* 1999). In the multiplier form of the models, this corresponds to turning from inequality to equality the restriction referring to the MPS. This restricts the choice of input/output weights for the evaluated DMUs only to those that are optimal for the MPS.[54] In essence, the choice of the MPS results in evaluating every DMU using the MRSs and MRTs that are observed on the DEA frontier in the neighborhood of the MPS.[55] The DM may view these marginal rates as adequate enough to apply globally as they reflect his/her own valuation of inputs and outputs. The evaluated DMUs for which at least one optimal vector of weights in DEA is also optimal for the MPS receive the most optimistic VEA score possible, namely one that is equal to their DEA efficiency score. The remaining DMUs, the input/output structure of which "deviates too much" from the one of the MPS (Korhonen *et al.,* 2002, p. 59), are forced to accept less favorable weights in VEA compared to DEA, and their VEA scores are lower than the corresponding DEA ones.

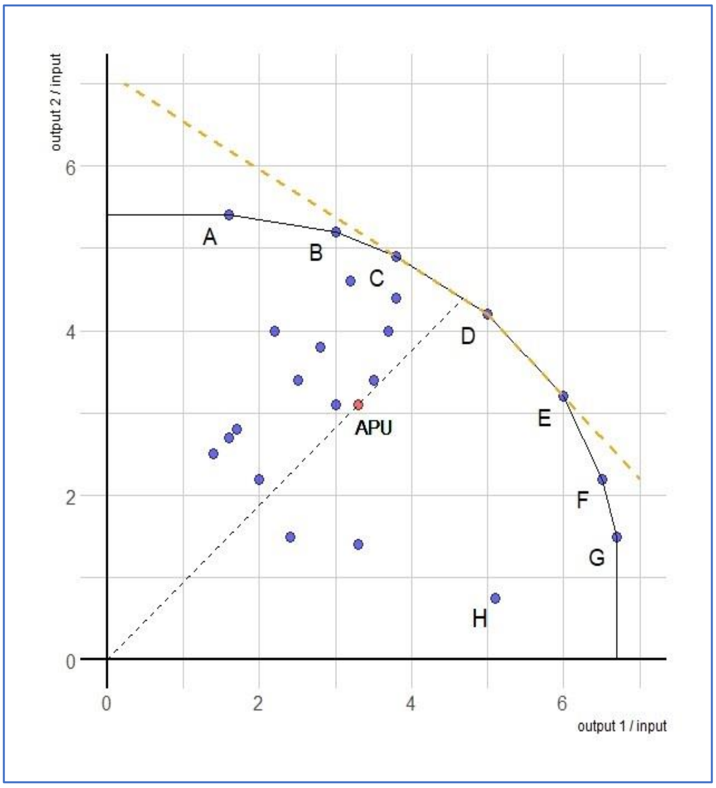
The facet extensions in VEA are illustrated in Figure 4.1 in the case of one-input-two-outputs technology. Choosing DMU D as the MPS implies the dashed line frontier by extending facets CD and DE. If the output price ratio ranges between the slopes of the two facets intercepting at D, the resulting VEA scores might be viewed as providing approximate estimates of overall (i.e., cost, revenue, or profit) efficiency(Joro and Korhonen, 2015 p. 100). If the DM wishes to prioritize the production of the second output compared to that of the first one, he/she might choose

---

[54] Such equality restrictions have been used for incorporating expert views in DEA in other studies as well, without referring explicitly to VEA. Zhu (2001) uses the CRS counterpart of (4.2) in its multiplier form to benchmark the quality of life of 20 cities against a set of peer DMUs that would necessarily contain three pre-selected cities identified by Fortune magazine as the top-three best cities in terms of quality of life (see his equation (8)). Furthermore, Cook *et al.* (2004) used input-oriented CRS and VRS VEA models under the name "fixed benchmark model" in order to measure the performance of out-of-sample DMUs (see their equations (9) and (10)). Also, Wang and Luo (2006) used a model that is equivalent to the input-oriented CRS VEA model, in which the frontier projection of an artificially constructed ideal DMU (IDMU), namely one that consumes the minimum sample quantities for each input while producing the maximum sample quantities for each output, corresponds to the MPS (see their equation (4)). The DEA frontier projection of the IDMU was obtained via a super-efficiency model.

[55] Ratios of optimal values of input and/or output DEA weights reflect marginal rates of substitution between inputs, transformation among outputs, and marginal products between inputs and outputs that are observable on the frontier (Charnes *et al.,* 1985).

Figure 4.1:Extending efficient facets through VEA



DMU B as MPS, which would extend facets AB and BC. On the other hand, if the DMUs operate in a relatively uniform environment (e.g., are employees within the same organizational department), the DM may wish to evaluate them based on a common value system. This could be done, for instance, by choosing both C and D as MPS. Then, only the common facet between C and D is expanded, and DMUs are evaluated by using a common vector of weights (the one that is normal to facet CD). If, however, DMU G is chosen as the MPS, the VEA frontier will also include the vertical segment (weakly efficient facet) between DMU G and the horizontal axis, for which the MRT between the two outputs is undefined. This will allow the inefficient DMU H to assign a zero value to the weights of the first output in its evaluation by VEA.

4.3. MPS choice

This section is divided into a literature review subsection, where we present and evaluate previously suggested MPSs, and a subsection where we make four new suggestions for MPS choice.

4.3.1. Literature review

In this section we present nine suggestions used previously in the literature and discuss the rationale associated with each one and characteristics that may encourage or discourage its use by managers.

4.3.1.1. DM personal judgment

In this case, DMs fully exert their judgments and obtain evaluation results that correspond to their legitimate priorities. The DM may choose a single or several DMUs for this purpose. In the latter case, Korhonen *et al.* (2002) suggested to form a virtual DMU by averaging over the input and output quantities of the chosen DMUs.[56] The so constructed DMU may be inefficient, indicating that the DM has conflicting preferences, and in this case its set of DEA peers should be used as MPS (see Joro and Korhonen, 2015, p. 124).

4.3.1.2. Prior or external information

Prior or external information, regarding previous evaluation results or achievements, may be used by DMs to choose a single or a set of MPS. Marshall and Shortle (2005) made this suggestion but its usefulness depends on the accuracy of the relevant information. As these might refer to a different sample of DMUs, another set of inputs and outputs, and different "environmental" conditions, they might not be representative for the evaluated DMUs at hand.

4.3.1.3. Best-in-input or best-in-output DMUs

Korhonen *et al.* (1998) suggested choosing as the MPS a best-in-input DMU, namely one that uses the smallest quantity of a particular input, or a best-in-output DMU, i.e., one that produces the largest quantity for a given output.[57] In a multiple-input-multiple-

[56] This is suggested if the DM views the most preferred structure as a combination of the chosen DMUs (Korhonen *et al.*, 2002). If the DM views each of the chosen DMUs as representing a different type of good performance, this may indicate that there are non-homogeneous sub-samples of DMUs, which may be fairer to evaluate separately from each other.

[57] Marshall and Shortle (2005) defined as "super-achievers" those DMUs that have the largest (smallest) sample quantity for a particular output (input) but also outperform the DMU with the second largest (smallest) quantity by a large margin. As the extent of this margin was not formally defined, we consider only best-input and best-in-output DMUs from now on.

output setting, there will be more than one best-in-input and best-in-output DMUs, in which case one of them should be chosen as the MPS. The choice may be facilitated if the DM views a particular input or output as overwhelmingly more important than all other inputs or outputs (as e.g., is the case with employee salaries in public services, see Joro and Viitala, 2004). Such views could however be reflected directly in the specification of inputs and outputs by excluding all other inputs or outputs from the analysis.

The use of a best-in-input or a best-in-output DMU may result in assessing the DMUs against a technically non-optimal scale. This is because a best-in-input DMU is usually of very small size and possibly of sub-optimal scale, and a best-in-output DMU is often large-sized and has supra-optimal scale. Also, the choice of a best-in-output MPS might imply a management directive towards increasing production disregarding the costs this may incur, while a best-in-input MPS might reflect the need for urgent budget cuts, without considering whether the resulting decreased production will be able to meet demand in the future.

4.3.1.4. IDMU

The IDMU uses the sample minimum quantities of each input to produce the sample maximum quantities of each output. It is thus "best" in all inputs and outputs. If it is not among the evaluated DMUs, it cannot be used as the MPS, but its DEA frontier projection could be. For this purpose, one may estimate its efficiency score by means of a super-efficiency DEA model and use its efficient projection of inputs and outputs as MPS (Wang and Luo, 2006). Since the frontier projection of the IDMU may contain slacks, the set of IMDU peers may instead be used as MPS, to ensure that it is a non-dominated DMU. In several occasions, the IDMU may look as a suitable MPS choice but its input/output bundle is likely to differ from most of the evaluated DMUs. This in turn may result in VEA efficiency scores that differ significantly from the DEA efficiency scores.

4.3.1.5. Most frequent peer

In this case, the MPS is the efficient DMU appearing the most times as a peer in the DEA model. This DMU is an example-to-follow for most of the DMUs, and it may be viewed as reasonable benchmark or "global leader" (Oral and Yolalan, 1990); Oral *et* *al.*, 1992) for them. Then, the VEA efficiency scores for most of the DMUs will be equivalent to their DEA ones, and thus the use of VEA will not provide additional insights to central management compared to the results of the DEA model.[58] In addition, a DMU acting as a peer for a large number of DMUs could be a potential outlier if it performs extremely better in relative terms compared to the DMUs it influences (Bogetoft and Otto, 2011, p. 147), in which case it should be excluded from the sample rather than being used as MPS.

4.3.1.6. Maximum (or infinite) super-efficiency

Halme and Korhonen (2015) suggested choosing as the MPS the DMU with the maximum super-efficiency score. In a CRS setting, the DEA super-efficiency model always results in *finite* scores, in which case it is rather straightforward to choose the MPS. On the other hand, the VRS super-efficiency DEA model may result in an infeasible solution for some DMUs. One may then choose as the MPS either the DMU with the maximum *finite* super-efficiency score or one among the DMUs for which the super-efficiency model has an infeasible solution. The DMU with the maximum super-efficiency score will frequently be among those that exert the most influence on the other DMUs' efficiency scores (Wilson, 1995), in the sense that it already appears as a peer for quite many DMUs. Then, the VEA model is not likely to provide additional insights to management. Also, DMUs with very large super-efficiency scores are often regarded as outliers (Wilson, 1995; Banker and Chang, 2006) that showcase extraordinary or very specialized performance, in which case such a DMU should not be used as MPS. On the other hand, DMUs for which the VRS super-efficiency model has an infeasible solution are usually located at some "end-point" of the DEA frontier (Seiford and Zhu, 1999), i.e., are likely overly specialized and are associated with MRSs and MRTs that are not well-defined (as DMUs A and G in Figure 4.1). If they are used as MPS, the VEA efficiency scores are likely to differ significantly from those of the DEA model and one or more of the inputs and the outputs will likely be assigned zero weights.

---

[58] This is noted by Oral and Yolalan (1990) and Oral et al. (1992), who interpreted the presence of insignificant differences as the choice of the "global leader" was quite realistic.

4.3.1.7. Minimum average Coefficient-of-Variation of weight vectors

According to Gonzalez *et al.* (2010), the efficient DMU with the minimum variability across its different optimal weight vectors is chosen as the MPS. To identify it, one needs first to estimate a VEA model using in sequence every efficient DMU as the MPS. For each of these models, one should calculate the Coefficient-of-Variation (CV) for the optimal values of every input and output weight and then take their average value. The MPS is chosen as the efficient DMU for which the average CV in the corresponding VEA model is the minimum. This might be appealing for DMs that want to avoid highly dissimilar optimal weight vectors among the evaluated DMUs in the VEA model but it can be relatively time-consuming. Furthermore, a common vector of weights across DMUs, which would reflect the greatest possible congruence (Gonzalez *et al.*, 2010), i.e., the minimum variability, among DMUs in selecting their optimal weights, is not guaranteed.

4.3.1.8. AHP importance weights

The Analytic Hierarchy Process (AHP) is suggested as another means to choose the MPS in VEA (Korhonen *et al.*, 1998). It may be used to obtain the "best" combination among all the DEA-efficient DMUs, or among a subset of them. The chosen DMUs are used as alternatives in AHP and the DM performs pairwise comparisons among them. The MPS is then obtained as a combination of the chosen DMUs using the importance weights derived from AHP. This might be a time-consuming process if there is a large number of chosen DMUs. In addition, the resulting DMU might be inefficient, indicating poor judgment in the initial selection of DMUs. In this case, its set of DEA peers should be used as MPS.

4.3.1.9. Interactive optimization

Halme *et al.* (1999) suggested the use of multi-criteria interactive optimization algorithms to choose the MPS. These algorithms enable the DM to search the efficient frontier and identify different non-dominated solutions. Halme *et al.* (1999) use the Pareto Race (see Korhonen and Wallenius, 1988), in which a MOLP problem is iteratively solved to obtain an efficient input/output combination, which has the maximum (minimum) possible value for each output (input). In each iteration, the DM reviews the resulting combination and can prioritize which input (output) should be further decreased (increased) at the expense of others, i.e., determine the direction on which the next (and possibly more preferred) input/output combination will be searched for. The algorithm stops when the DM decides that the last identified input/output combination is the MPS. In most of the cases, this is a combination of efficient DMUs.

An alternative proposed by Korhonen *et al.* (2002) is the Visual Interactive Method for Discrete Alternatives (VIMDA) (see Korhonen, 1988). It is similar to Pareto Race but in each iteration it identifies an input/output combination corresponding to an existing DMU rather than a combination of DMUs. Such algorithms can be time-consuming and require a DM that is willing to participate and direct the algorithm according to his/her preferences (Thiele *et al.,* 2009). This may increase management workload and the risk of providing a poor judgment. Also, in practical applications DMs usually view the existing DMUs as more reliable benchmarks compared to combinations of DMUs (Korhonen *et al.,* 2002).

4.3.2. Some new suggestions

In this section we expand the set of MPS choices in VEA by suggesting four new, each of which may be useful to managers for certain reasons.

4.3.2.1. Informed personal judgment

In the first of our suggestions the DM exerts his/her personal judgments by explicitly considering the position of DMUs on the DEA efficient frontier. Some of the efficient DMUs reside closer to most of the sample DMUs while others use a somewhat more extreme input/output bundle. In addition, some efficient DMUs are associated with zero or undefined marginal rates while others are not, some can remain efficient even if their input/output bundle changes, and for some there do not exist DMUs with similar input/output structure in the sample. Classifying the DMUs based on such features may aid the DM in making a more informed personal judgment when choosing the MPS.

We consider two main classifications of the efficient DMUs based on their position on the frontier. In the first one, the DMUs are classified as either active or self-evaluators (Edvardsen *et al.*, 2008). The former are efficient DMUs that appear as peers for at least one inefficient DMU, while the latter appear as peers only for themselves.[59] Each of the active and self-evaluator DMUs can be further classified as an interior or an exterior. For an exterior DMU, at least one among its adjacent facets is weakly efficient, while for an interior this is not true.[60] An interior active DMU resides closer to most of the sample DMUs and its use as MPS might result in moderate (and even insignificant) differences between the DEA and the VEA efficiency scores. An exterior active DMU may use a more extreme input/output bundle, and if used as MPS, a zero weight will be assigned to one or more inputs and/or outputs for some of the evaluated DMUs. The interior self-evaluators are "alone in the crowd", while the exterior self-evaluators are "far out", located at an "end-point", i.e., use an extreme input/output bundle and be very small- or large-scaled (Edvardsen *et al.,* 2008). In both cases, significant differences should be expected between the DEA and VEA efficiency scores. In addition, some inputs and/or outputs are more likely to have a zero weight if an exterior self-evaluator is used as the MPS.

The second classification partitions the efficient DMUS into terminal and non-terminal ones (Krivonozhko *et al.*, 2015). A terminal DMU will remain efficient even if the quantity of one of its inputs (outputs) is increased (decreased), while for a non-terminal one this is not true.[61] Each terminal DMU may be further classified as being either interior or exterior, but all non-terminal DMUs are interior.[62] An exterior terminal DMU is more likely to be located on "end-points" of the frontier compared to an interior terminal DMU, but Krivonozhko *et al.* (2015) note that both classes may

[59] Self-evaluators are those for which the maximum optimal values of the intensity variables are equal to zero for every inefficient DMU. If at least one such value is positive, the efficient DMU is classified as active (Edvardsen *et al.,* 2008). An alternative is to estimate the referencing share (see Torgensen *et al.,* 1996) for each efficient DMU, which captures the relative contribution of an efficient DMU in the total output expansion (input contraction) of all the inefficient DMUs for each specific output (input). DMUs with a zero referencing share are classified as self-evaluators.
[60] The classification of efficient DMUs into exteriors or interiors is obtained by enveloping the efficient DMUs "from below" (Edvardsen *et al.,* 2008) through a modified version of the Additive DEA model in which inputs are treated as outputs and vice versa. A DMU with a zero (positive) optimal value is classified as an exterior (interior).
[61] Terminal DMUs are adjacent to at least one-dimensional facet (Krivonozhko *et al.,* 2015). They are identified by estimating a series of linear programs, one for each different input and output, each of which aims at maximizing the value of the intensity variable of a given extreme-efficient DMU while allowing for the particular input (output) of the DMU to increase (decrease) along a one-dimensional ray. A DMU is classified as terminal if the optimal value of its intensity variable equals one in at least one of those linear programs. Otherwise, it is non-terminal.
[62] Krivonozhko *et al*. (2015) show that the set of terminal DMUs contains that of exterior DMUs as a subset, i.e., each exterior DMU is also a terminal DMU, but a terminal DMU may be either an interior or an exterior.

contain quite normal efficient units. Thus, the use of an exterior terminal DMU may result in significant or insignificant differences between the DEA and VEA efficiency scores, and the same may be the case when an interior terminal DMU is the MPS. On the other hand, the use of a non-terminal DMU as MPS will not result in assessing the DMUs against unacceptable marginal rates, while when using a terminal DMU this is expected to occur.

4.3.2.2. Most Productive Scale Size

Our second suggestion is to choose a DMU with MPSS as MPS. Such DMUs operate with technically optimal scale, namely maximize average productivity for their input/output mix.[63] Each such DMU is efficient under both a CRS and a VRS DEA model, i.e., resides on a frontier segment in which CRS prevails and scale elasticity equals one (Banker, 1984). The use of an MPSS DMU as the MPS in VEA ensures that DMUs are assessed against a technically optimal scale. The resulting VEA scores could yield useful insights for central management. They might be used for reorganizing or incentivizing the DMUs so that they adjust to the optimal scale, the pursuit of which constitutes a long-term organizational goal.

In several cases, there are multiple MPSS DMUs, each of which operates with the technically optimal scale for its own input/output bundle (Banker and Thrall, 1992). In this case Banker (1984) noted that obtaining the overall optimal scale for the underlying technology requires the use of additional knowledge or information. This can be provided by the DM by means of choosing one DMU or a combination of DMUs among those with MPSS as the MPS. The chosen input/output bundle might be close to that of most DMUs in the sample, in which case the VEA efficiency scores may differ only moderately from their DEA counterparts. Alternatively, there might be significant differences between the DEA and VEA efficiency scores if the DM chooses an MPSS DMU with somewhat extreme mix of inputs and outputs.

4.3.2.3. Average Production Unit

---

[63] Technically optimal scale in production theory was first discussed in Førsund and Hjalmarsson (1979) in single-output-multiple-input settings and was generalized for multiple inputs and outputs in Banker (1984).

Our next suggestion is to use the combination of DMUs that are the peers of the Average Production Unit (APU), namely an artificially constructed DMU that operates with the sample means quantities of inputs and outputs, as the MPS. This reflects the objective of maximizing the structural efficiency of an overall entity that coordinates a set of DMUs.[64,65] This entity might be either a firm operating through a network of multiple branches or plants, or an industry of similar firms. Structural efficiency is a normative rather than a positive measure (Karagiannis, 2015), in the sense that it assesses the extent of potential improvement of the entity (firm or industry) as a whole, as if it were a single DMU utilizing and coordinating (through centralized resource allocation) the total quantities of inputs and outputs. The maximum potential output for the entity could be realized if each of the coordinated DMUs had the input/output structure of the APU and then removed its technical inefficiencies (Kittelsen and Førsund, 1992; Karagianis, 2015) as well as input and/or output slacks.

When the APU peers are used as the MPS, the VEA efficiency scores reflect the relative performance of DMUs from the perspective of fully centralized management and can provide useful insights to managers that coordinate a firms' branches or to authorities planning a sectoral reorganization. The APU input/output bundle is relatively close to that of many DMUs, and thus one might expect moderate changes in the efficiency scores in VEA compared to DEA. However, the efficiency scores of DMUs using extreme input/output bundles may decrease considerably. For example, in Figure 4.1 where the APU is radially projected on the efficient facet CD and thus its DEA-efficient peers are DMUs C and D, VEA evaluates all DMUs compared to the extended facet CD, and the DMUs A, G and H exhibit large decreases in efficiency compared to their corresponding DEA scores.

4.3.2.4. Common weights

Our fourth suggestion concerns evaluating all DMUs using a common vector of strictly positive input/output weights. This results in evaluating all DMUs based on a common

[64] Førsund and Hjalmarsson (1979) were the first to argue that the extent of structural efficiency in a sample of DMUs is equal to the technical efficiency score of the APU, an argument formally proved by Li and Ng (1995).

[65] Using the sample average DMU as the MPS generalizes in a sense the suggestion made by Korhonen *et al.* (2002) to obtain the MPS by averaging across a pre-selected subset of efficient DMUs.

standard (Kao and Hung, 2005) and might thus be preferred when DMs wish to prevent individual DMUs from setting and pursuing their own priorities.[66] This could be the case if the DMUs are homogeneous enough, operate under a common policy framework (Cook *et al.,* 2019), and/or in the same environment (e.g., professors engaging in teaching and research activities within the same university faculty). Potential discrepancies between the results using common weights and conventional DEA should then indicate the effect of special circumstances under which a DMU operates (Roll *et al.*, 1991), or a DMU that may be prioritizing its own objectives over those of the organization. This suggestion for choosing the MPS is, to the best of our knowledge, the only one securing the assessment of DMUs against well-defined MRSs and MRTs.

Common and strictly positive weights across DMUs are guaranteed in VEA when a single FDEF of the DEA frontier is extended. This will occur if the unique combination of $(I + J - 1)$ extreme efficient DMUs that supports an FDEF when CRS is assumed (or $(I + J)$ DMUs in VRS models) (Olesen and Petersen, 2003; 2015) is chosen as the MPS, provided that at least one FDEF exists.[67] In most cases, the DEA frontier is generated by multiple FDEFs. The DM should then choose one among those FDEFs to be extended in the VEA model. The choice can be facilitated if one identifies all the FDEFs of the DEA efficient surface and the combinations of DMUs spanning each, which is frequently done using mixed-integer linear programs (Olesen and Petersen, 2003; Fukuyama and Sekitani, 2012; Davtalab-Olyaie *et al.* (2014).[68] The DM can then review these results and choose the FDEF against which DMUs will be assessed. The use of common weights in VEA will more likely result in efficiency scores that differ, in a statistically significant sense, from those of the DEA model. More specifically, only DMUs which are already projected by DEA in the chosen

[66] Common weights are frequently adopted in the DEA literature (see Afsharian *et al.,* 2021 for a recent review) and they reflect the greatest possible congruence among DMUs in selecting their optimal weights.
[67] FDEFs are associated with a unique normal vector of input/output weights with strictly positive values (Olesen and Petersen, 2015). Olesen and Petersen (1996) referred to the absence of at least one FDEF in the DEA frontier as an indication of an ill-conditioned dataset.
[68] Olesen and Petersen (2003) proposed a cutting plane binary optimization algorithm, which requires, as a prior step, to identify the set of DMUs that can be rendered efficient in a convex combination with each of the extreme-efficient DMU. Mixed-integer linear programs are then solved to identify, for each extreme-efficient DMU, all FDEFs generated by it. Fukuyama and Sekitani (2012) and Davtalab-Olyaie *et al.* (2014) proposed similar binary optimization algorithms. The former identifies both FDEFs and non-FDEFs, while the latter does not require the prior step that is necessary in the approach of Olesen and Petersen (2003).

FDEF, i.e., those for which the combination of efficient DMUs generating the FDEF coincides with their set of peers, will retain the same efficiency score, while the remaining ones will exhibit at least slight decreases in efficiency. The differences can be on average large if a DMU with a relatively extreme input/output bundle is among those generating the chosen FDEF.

4.4. Data, variables and modeling choices

For our empirical application we use data for 526 Greek cotton farms obtained from the Farm Accounting Data Network (FADN). The FADN covers large entrepreneurial farms as defined in the farms structure survey of the EU, in which each farm is classified by commodity according to its main source of revenue. That is, a farm is classified as a cotton producer if at least two thirds of its revenue come from the production of cotton.

Output orientation is usually considered as the more appropriate choice when measuring efficiency in agriculture, in which input choices are made well in advance of output realization. (Karagiannis, 2014). We also assume that input and output prices are uniform across DMUs, since the agricultural sector is widely considered as a rather competitive one, where there is usually a large number of farmers specializing on the production of a particular commodity and facing similar prices for the resources used and their final product. In this case, input and output data expressed both in terms of quantities and in terms of values (i.e., costs and revenues) can be used to assess technical efficiency (see Portela, 2014). We use four inputs, namely land measured in ha, labor (including family and hired workers) measured in annual working hours, intermediate inputs (i.e., fertilizer, pesticides, etc.) measured in euros, and capital stock (including machinery and buildings) measured in terms of the end-of-the-year book values (in euros) and a single output, measured in terms of total gross revenue (in euros).

Average values of the model variables are given in Table 4.1. In that, we also include information on additional farm characteristics. These are farm size, the farmer's age, the geographic region in which each farm is located, the percentages of own and irrigated land, the percentage of family labor employed, as well each farm's

Table 4.1:Sample average estimates of model variables

revenue (in euros)	6434.103
land (in ha)	1188.139
labor (in annual working hours)	2045.654
intermediate inputs (in euros)	2369.776
capital (in euros)	5703.852
number of farms in the sample	526
farms from Northern Greece	123
farms from Western Greece	7
farms from Central Greece	294
farms from South-Eastern Greece	102
small size farms	45
medium size farms	207
large size farms	274
farms owned by younger farmers	67
farms owned by middle-aged farmers	386
farms owned by older farmers	73
own land (%)	0.662
irrigated land (%)	0.829
family labor (%)	0.870
specialization index	0.736

degree of specialization in the production of cotton.[69] Such variables account for important factors which affect the operating conditions of farms and consequently, their input/output structure and can provide insights regarding the closeness of the MPS's structure compared to that of the majority of the sample farms

Most of our sample farms are located in Central Greece (i.e., Thessaly, 55.9% of the sample), while the rest are almost equally divided between Northern (Macedonia and Thrace) and Southeastern (namely Sterea Ellada and Aegean Islands) Greece (23.4% and 19.4% respectively). Only a small fraction (1.3%) of farms is located at Western (Epirus and Peloponnesus) Greece. On average, the sample farms are relatively specialized in the production of cotton, rent about 44% of their land, while

[69] In FADN, farm size is defined in terms of gross value added. FADN defines nine size classes, which are grouped here into three categories, namely small, medium, and large farms. We also define three different age bands, namely younger (less than 40 years old), middle-aged (between 40 and 60 years) and older farmers (over 60 years old). The degree of specialization is measured by the Herfindhal concentration index (defined as $H_k = \sum_j s_{jk}^2$, where s_{jk} is the share of the j[th] output in total production of the k[th] farm). A value of H equal to unity indicates complete specialization, whereas smaller values reflect increased diversification

most of them are of large size according to FADN standards and are operated by middle-aged farmers (see Table 4.1).

4.5. Empirical results

This section provides the first thorough comparative empirical analysis of the variability in VEA efficiency scores for alternative MPS choices. For these purposes several models were estimated. More specifically, technical and scale efficiency scores for the sample DMUs (including the APU) were obtained by estimating conventional CRS and VRS DEA models. We find 12 farms to be both technical and scale efficient (i.e., have MPSS), while there were 21 technically efficient farms, most of which (17) operating with increasing returns-to-scale (RTS).[70] The complete set of the efficient farms is given in Table 4.2. On average, inefficiency is more due to producing below the frontier rather than operating at non-optimal scale (average technical and scale efficiency equal 0.598 and 0.947, respectively), while supra-optimal scale farms appear to operate closer to optimal scale compared to sub-optimal scale farms.

In addition, we estimated super-efficiency CRS and VRS DEA models. For eight farms the VRS model resulted in an infeasible solution. Separate super-efficiency DEA models were estimated for the IDMU by including it in the sample, among which the one assuming VRS resulted in an infeasible solution. We also estimated CRS and VRS VEA models using each of the efficient farms as the MPS, to identify the farm for which the variability across the optimal input and output weight vectors is minimum, as suggested in Gonzalez *et al.* (2010). The FDEFs generating the CRS and VRS DEA frontiers (14 FDEFs in the CRS frontier and 68 FDEFs in the VRS one) were identified using the mixed integer binary optimization algorithm of Davtalab-Olyaie *et al.* (2014).

4.5.1. Choice of MPS

A two-step procedure was used to choose the MPS for the CRS and VRS VEA models. In that, we considered all the MPS choices discussed in the third section apart from external information, interactive optimization, personal judgments and the AHP. This

[70] An additional DMU having an efficiency score of one with CRS was identified as weakly efficient, namely having positive slacks, and was not further considered. All other DMUs with an efficiency score of one either with CRS or with VRS are extreme-efficient.

Table 4.2: Extreme-efficient DMUs selected as the MPS by alternative choices

	farm (in coded numbers)																																	
MPS choice	32	37	69	91	119	130	142	143	145	147	154	160	178	183	197	201	203	216	226	235	241	252	267	273	275	293	314	368	404	410	411	415	524	sum[a]
CRS																																		
most times as peer in DEA																						■[b]												1
maximum finite superefficiency													■																					1
infeasible superefficiency																																		-
best-in-output																						■												1
best-in-input																																		0
minimum average CV														■																				1
IDMU		■											■															■						1
interior active											□			□								□				■		□						5
interior self-evaluator																																		-
exterior active		□				■						□	□			□											□					□		7
exterior self-evaluator																																		-
interior terminal											■			□								□						□						4
exterior terminal		□				■						□	□			□											□					□		7
non-terminal																										■								1
MPSS		□				■					□	□	□	□		□						□				□	□	□				□		12
APU																						■				■	■					■		1
FDEF		□				■					□	□	■	□		□						■				□	□	□				■		14
times suggested as MPS		5				4					4	4	6	5		4						7				5	5	5				5		
VRS																																		
most times as peer in DEA																						■												1
largest finite superefficiency													■																					1
infeasible superefficiency			■		□				□						□				□	□			□						□					8
best-in-output																						■												1
best-in-input			■		□															□									□					4
minimum average CV							■																											1
IDMU																																		-
interior active						■		□	□	□	□		□	□			□		□				□	□	□	□	□	□	□	□	□	□		19
interior self-evaluator							□											■																2
exterior active	■	□	□	□								□			□	□				□	□	□											□	11
exterior self-evaluator					■																													1
interior terminal						■		□	□		□		□	□			□	□	□				□	□	□	□	□	□	□	□	□	□		19
exterior terminal	□	□	□	□	□							□			□	□				□	■	□											□	12
non-terminal							□			■																								2
MPSS		□				■					□	□	□	□		□						□				□	□	□				□		12
APU						■																■				■	■							1
FDEF		□				■	□	□	□	□	□	□	□	□	□	□	□		□	□	□	■	□	□	□	■	■	□	□	□	□	■	□	68
times suggested as MPS	2	4	4	2	4	5	4	3	4	4	4	4	5	4	4	4	3	2	4	5	3	7	4	3	3	5	5	4	5	3	3	4	3	

Notes: (a) In the case of the IDMU, the APU, and the FDEF, the last column refers to the number of combinations of DMUs that serve as peers for the IDMU and the APU (one) and the number of FDEFs identified in the CRS and VRS DEA models (14 with CRS and 68 with VRS). (b) Each rectangle highlights that the corresponding farm is identified as MPS by the respective choice. The filled rectangles refer to the DMUs or combinations of DMUs chosen as the MPS in the application of this chapter.

is because external information is not available while the other three suggestions require the presence of a DM. In the first step, we identified the DMUs that could be the MPS in each case assuming CRS and VRS. These are indicated by a rectangle in their corresponding cell in Table 4.2, the last column of which shows the number of different MPSs indicated by each choice.

More specifically, the farm appearing the most times as peer in the CRS and VRS DEA models, the best-in-output farm and the four best-in-input farms were identified. With CRS, only the best-in-output farm is efficient, while with VRS the best-in-input farms are efficient as well. We also identified the farms with the minimum average CV, those having the maximum finite super-efficiency score with CRS and VRS, as well as those for which the VRS DEA super-efficiency model resulted in an infeasible solution.[71] The peers of the APU and the IDMU were identified, albeit for the latter only with CRS. Each farm was also classified based on its position on the DEA frontier, following the two classification schemes presented in the third section. From Table 2 we see that at least one farm is included in every group with VRS, while with CRS there are no self-evaluators. Also, a farm may be classified in a different group with CRS and with VRS. In addition, we identified the combinations of efficient farms generating the 14 FDEFs of the CRS frontier and the 68 FDEFs of the VRS one. With CRS, each efficient farm generates at least one FDEF, while this is not the case with VRS.

The second step involved choosing *one* DMU or a combination of DMUs to use in the empirical application when more than one DMUs or combination of DMUs could be the MPS. This is more likely to be the case when the DM chooses the MPS among (i) interior active, (ii) exterior active, (iii) self-evaluators, (iv) interior terminal, (v) exterior terminal, (vi) non-terminal, (vii) MPSS, (vii) best-in-input and best-in-output DMUs, (viii) the DMUs for which the VRS super-efficiency model has an infeasible solution and (ix) the combinations of DMUs generating an FDEF. See the last column in Table 4.2, where for most of these choices there are multiple alternatives for the MPS.[72] For each of these choices, we chose as the MPS the farm for which land

[71] This could not occur with CRS, and thus there is a dash in the corresponding cell in the last column of Table 4.2.

[72] On the other hand, there is one combination of peers for the IDMU and the APU. It is also likely that only one DMU appears the most times as a peer in the DEA model, one DMU has the maximum finite

quantity was closest to the average quantity of land among those farms indicated as potential MPSs by the choice. A similar procedure was followed in the case of common weights, namely to choose one combination of farms generating an FDEF. We ranked the efficient farms in terms of their deviation from the sample average land quantity and chose the farm with the minimum deviation. If the farm ranked second shared a common facet with the one ranked first, we considered it for the combination. Otherwise, we bypassed it and moved to the next farm in the ranking. This process ended when a combination of farms generating an FDEF was obtained. After the MPS choice, a VEA model was estimated for each of the alternative MPSs with CRS and with VRS.

The farms or the combination of farms chosen as the MPS are indicated by a filled rectangle in the respective cell of Table 4.2. From that we see that a particular MPS choice may result in choosing a different MPS with CRS and with VRS (see the average CV). In addition, some farms are frequently suggested as the MPS: with CRS two farms are suggested as the MPS (either solely or within a combination of farms) six and seven times respectively, while with VRS case one farm is suggested as the MPS seven times while six farms are suggested five times each. This can be attributed to the fact that for many of the MPS choices multiple DMUs could be the MPS.

The economic and socio-demographic characteristics of the chosen MPS are given in Table 4.3. Most of these are medium or large in size, are located in Central Greece and operated by middle-aged farmers (ages 40 to 60). On the other hand, only a few chosen MPSs are located in Northern Greece. More specifically, farm #32 is a medium-sized farm located in Northern Greece that is chosen as an exterior active MPS with VRS. It operates with a sub-optimal scale and it owns and irrigates very low percentages of its land compared to the average. Farm #69 is located in Northern Greece, operates with sub-optimal scale and uses the lowest quantity of land in the sample (best-in-input), while it is also chosen as MPS among the farms with an infeasible VRS super-efficiency model. It is thus possibly located at an "end point" of the frontier. The same is likely the case for farm #119, which is a sub-optimal scale farm located in Northern Greece and chosen as an exterior self-evaluator MPS. On the other hand, the chosen interior

super-efficiency score and one DMU is selected based on the minimum "average CV", but ties among DMUs are also possible for these suggestions.

Table 4.3:Economic and socio-demographic characteristics of the MPSs

farm (in coded number)	revenue (in euros)	land (in ha)	labor (in annual working hours)	intermediate inputs (in euros)	capital (in euros)	region	farm size	farmer age	own land (%)	irrigated land (%)	family labor (%)	specialization index[c]	RTS
32	3160	1050	554	717	6384	Northern	medium	41	0.457	0.267	1.000	0.541	irs
37	6090	1620	1673	1975	458	Northern	large	59	0.133	0.620	0.727	0.762	crs
69	2029	360	321	689	3091	Northern	small	56	1.000	1.000	1.000	0.500	irs
119	1067	255	1267	279	1499	Northern	small	48	1.000	0.686	1.000	0.446	irs
130	7760	1240	3070	1070	3510	Western	large	46	0.290	1.000	0.961	0.607	crs
142	1670	210	1323	386	487	Central	small	64	1.000	1.000	1.000	0.693	irs
147	6741	670	955	1717	1527	Central	medium	55	0.701	1.000	1.000	1.000	crs
154	13578	1480	1029	3663	3337	Central	large	45	0.243	0.946	0.979	0.976	crs
178	4050	300	2222	776	162	Central	medium	33	1.000	1.000	0.734	0.546	crs
183	3393	420	820	1344	320	Central	medium	43	0.833	1.000	1.000	1.000	crs
216	9839	1360	640	2726	3971	Central	large	51	0.279	0.735	1.000	0.784	irs
241	13470	970	3679	5872	8257	Central	large	39	0.381	1.000	0.952	0.280	drs
252	31726	3130	1494	5808	10037	Central	large	60	0.284	0.831	0.871	0.907	crs
293	11799	1000	1800	2146	4703	Central	large	54	1.000	1.000	1.000	1.000	crs
314	13123	1060	1700	2406	8300	Central	large	57	1.000	1.000	1.000	1.000	crs
368	4902	695	917	1367	723	Central	medium	27	0.115	0.806	1.000	0.648	crs
415	3661	410	1664	531	1750	Central	medium	55	0.756	1.000	0.784	0.856	crs
average	6434.103	1188.139	2045.654	2369.776	5703.852				0.662	0.829	0.870	0.736	

self-evaluator farm #216 is of sub-optimal scale but large in size and located in Central Greece, as most of the sample DMUs. It is thus more likely to be "alone in the crowd".

Farm #130 is a large size, optimal-scale farm located in Western Greece and operated by a middle aged-farmer that is chosen as the MPS multiple times (as exterior active and exterior terminal with VRS, as interior active and interior terminal with VRS, and as MPSS). Its input-output bundle is somewhat similar to the average, suggesting that it is located close to most of the farms in the sample. The same is likely for farm #293, which has a similar structure, operates with technically optimal scale and is chosen as interior active and non-terminal MPS with CRS. On the other hand, the non-terminal MPS with VRS (farm #147) is a medium-sized one, although it has similar socio-demographic characteristics with farm #293 and is also of optimal scale. The other farms chosen based on their frontier location (farm #154 as interior terminal with CRS and farm #241 as exterior terminal with VRS) are both large-sized and located in Central Greece. Farm #241 is however of supra-optimal scale and is operated by a young farmer, while farm #154 operates with technically optimal scale.

The best-in-output farm #252 is a large-sized, relatively capital-intensive farm located in Central Greece which appears the most times as a peer with CRS and with VRS. It is thus a very influential peer, as is likely the case for farm #178, which is the one with the maximum finite super-efficiency score for both model specifications. It is located in Central Greece but is of medium size and relatively more labor-intensive. On the other hand, the two farms suggested as the MPS with the minimum variability in their optimal weights with CRS (farm #183) and with VRS (farm #142) have a very small scale compared to the average. Both are located in Central Greece but the latter operates with a technically sub-optimal scale and appears as a peer only for itself, suggesting that it is located at an "end-point" of the frontier.

In the case of the APU, a combination of four farms is the MPS either with CRS or with VRS. Each farm in these combinations is MPSS, while most of these are large-sized farms located Central Greece and operated from middle-aged farmers. The same is the case for the combinations of farms (four with CRS and five in VRS) selected as MPS in the case of common weights. Most of the farms in these combinations have an input/output bundle relatively close to the average. For the case for common weights this is a result of the process we followed to select the associated FDEF. On the other hand, the IDMU peers are three MPSS farms, which utilize very low capital quantities compared to the average. One of these is located in Northern Greece, while two of them own excessively low proportions of their cultivated land.

4.5.2. Comparative results between DEA and VEA models

The VEA efficiency scores are always less than or equal to the corresponding DEA scores. This implies a decrease in average efficiency compared to the DEA model (see Table 4.4), and a left ward shift of the VEA distribution of efficiency scores compared to that of DEA (see Figure 4.2). For some of the MPS choices these shifts are large, for others there are only moderate, while for some MPS choices the VEA distribution of efficiency scores is not statistically different from that of the DEA scores (see Table 4.5) based on average shifts in rank (given as $R = \frac{1}{k}(\sum_{k=1}^{K}|rank_A(y^k) - rank_B(y^k)|)$, see Saisana *et al.*, 2005) and distribution equality tests (Banker and Natarajan, 2011).

More specifically, the use as the MPS of (i) the farm that appears the most times as a peer, (i) the one with the maximum finite super-efficiency score, and (iii) the best-in-output farm results in efficiency distributions that do not differ, in a statistically significant way, between DEA and VEA, irrespective of the RTS assumption (see Table 4.5). The same is essentially true with CRS for the non-terminal and the interior-active MPS, which are the same farm. In these cases, the results from the VEA model do not offer some additional insights to managers compared to those of the DEA model. This should be expected for the first two MPS choices, as they are based on influential DMUs appearing as peers for a large proportion of farms. For the other choices, it is explained by the fact that the farms chosen as MPSs have an input/output bundle that is close to that of most of the sample farms. Note also that when the non-terminal DMU is the MPS, all inputs are important for the estimation of efficiency, in the sense that all farms assign a positive value to the weights attached to each input.

On the other hand, when the MPS is an (interior or exterior) self-evaluator there are statistically significant (see Table 4.5) differences between the VEA and the DEA distributions of efficiency scores, and the same holds with VRS for the minimum "average CV" choice. In these cases, we observe the largest left ward shifts in the VEA distribution of efficiency scores compared to that of DEA. This is expected to occur in most cases where the MPS is either an interior self-evaluator that is located "alone in the crowd", or an exterior self-evaluator located on an "end-point" of the frontier, as these DMUs appear as peers only for themselves. It also occurs in our case for the CRS

Table 4.4: Efficiency scores for alternative MPS choices

model		average		minimum		median		standard deviation		efficient farms	
		CRS	VRS	CRS	VRS	CRS	VRS	CRS	VRS	CRS	VRS
DEA		0.562	0.598	0.091	0.091	0.575	0.607	0.207	0.225	12	33
VEA	MPS choice										
1	most times as peer in DEA	0.555	0.586	0.090	0.091	0.565	0.593	0.206	0.219	11	23
2	maximum finite superefficiency	0.551	0.576	0.091	0.091	0.559	0.580	0.205	0.219	12	26
3	infeasible superefficiency	-	0.363	-	0.043	-	0.346	-	0.188	-	5
4	best-in-output	0.555	0.586	0.090	0.091	0.565	0.593	0.206	0.219	11	23
5	best-in-input	-	0.363	-	0.043	-	0.346	-	0.188	-	5
6	minimum average CV	0.373	0.288	0.038	0.044	0.346	0.282	0.194	0.145	5	6
7	IDMU	0.373	-	0.039	-	0.351	-	0.192	-	4	-
8	interior active	0.540	0.524	0.090	0.078	0.557	0.537	0.200	0.195	6	7
9	interior self-evaluator	-	0.208	-	0.016	-	0.168	-	0.159	-	5
10	exterior active	0.508	0.417	0.077	0.060	0.520	0.416	0.190	0.185	4	4
11	exterior self-evaluator	-	0.201	-	0.027	-	0.194	-	0.118	-	4
12	interior terminal	0.415	0.524	0.046	0.078	0.391	0.537	0.202	0.195	6	7
13	exterior terminal	0.508	0.498	0.077	0.087	0.520	0.515	0.190	0.196	4	6
14	non-terminal	0.540	0.477	0.090	0.061	0.557	0.463	0.200	0.210	6	11
15	MPSS	0.508	0.524	0.077	0.078	0.520	0.537	0.190	0.195	4	7
16	APU	0.498	0.499	0.079	0.078	0.513	0.514	0.184	0.186	4	5
17	FDEF	0.482	0.497	0.071	0.078	0.494	0.514	0.184	0.185	4	5

Table 4.5: Statistical tests between DEA and VEA

		average rank shift		Mann Whitney		Banker F1[a]		Banker F2	
VEA	MPS choice	CRS	VRS	CRS	VRS	CRS	VRS	CRS	VRS
1	most times as peer in DEA	9.167	12.522	0.639	0.814	1.021	1.03	1.031	1.032
2	maximum finite superefficiency	11.579	17.512	0.959	1.600	1.032	1.062	1.046	1.081
3	infeasible superefficiency	-	78.899	-	16.262***[b]	-	1.915***	-	2.847***
4	best-in-output	9.167	12.522	0.639	0.814	1.021	1.03	1.031	1.032
5	best-in-input	-	78.899	-	16.262***	-	1.915***	-	2.847***
6	minimum average CV	75.662	55.558	13.865***	20.985***	1.714***	2.264***	2.507***	3.686***
7	IDMU	75.579	-	14.283***	-	1.709***	-	2.488***	-
8	interior active	16.924	49.307	1.675*	5.512***	1.061	1.212***	1.086	1.279***
9	interior self-evaluator	-	99.032	-	23.529***	-	3.027***	-	6.655***
10	exterior active	40.820	68.930	4.455***	13.209***	1.153**	1.631***	1.223**	2.107***
11	exterior self-evaluator	-	62.406	-	24.320***	-	2.864***	-	5.589***
12	interior terminal	68.685	49.307	11.418***	5.512***	1.528***	1.212***	2.051***	1.279***
13	exterior terminal	40.820	60.338	4.455***	7.359***	1.153**	1.314***	1.223**	1.472***
14	non-terminal	16.924	58.622	1.675*	8.909***	1.061	1.415***	1.086	1.706***
15	MPSS	40.820	49.307	4.455***	5.512***	1.153**	1.212***	1.223**	1.279***
16	APU	39.169	55.605	5.329***	7.452***	1.183***	1.294***	1.265***	1.401***
17	FDEF	41.169	55.442	6.647***	7.564***	1.236***	1.298***	1.365***	1.409***

Notes: (a) F1 (F2) test compares the DEA and VEA distributions of efficiency scores, assuming an exponential (half-normal) distribution of the efficiency scores (see Banker and Natarajan, 2011). (b) Three, two and one stars denote statistical significance at 1%, 5% and 10% respectively.

Figure 4.2: DEA and VEA distributions of efficiency scores.

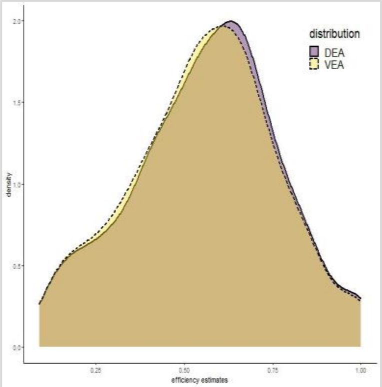
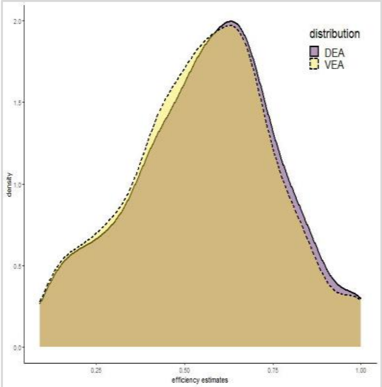
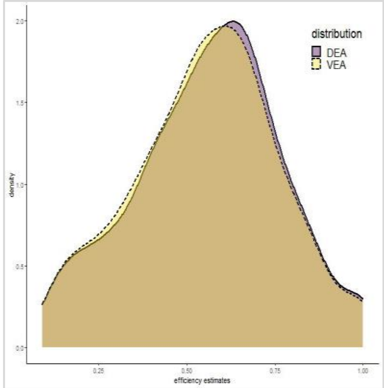
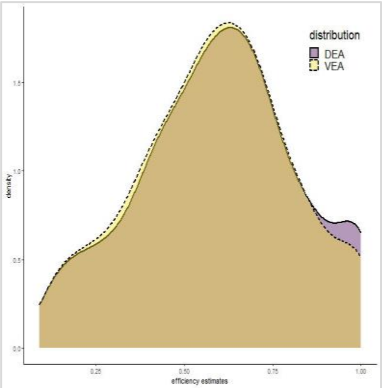
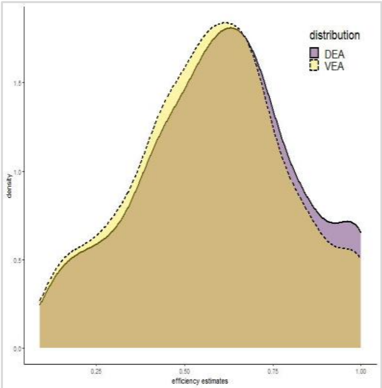
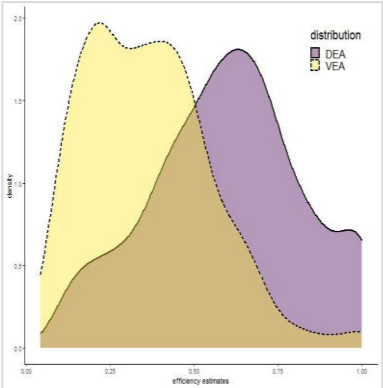
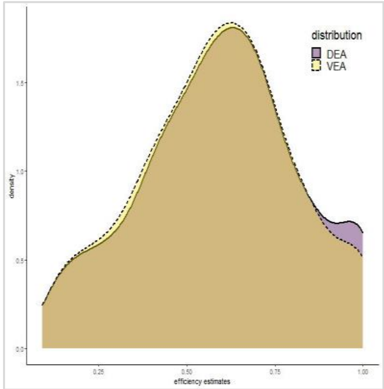
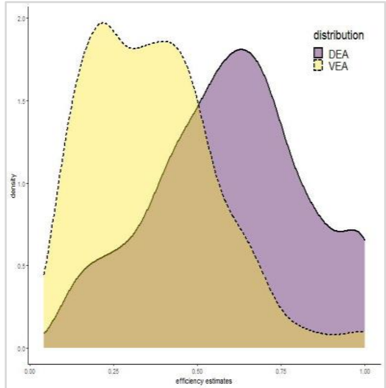
	MPS choice				
	1. most times as peer in DEA	2. maximum finite superefficiency	3. infeasible superefficiency	4. best-in-output	5. best-in-input
CRS			-		-
	MPS: {252}	MPS: {178}		MPS: {252}	
VRS					
	MPS: {252}	MPS: {178}	MPS: {69}	MPS: {252}	MPS: {69}

Figure 4.2 (Cont.)

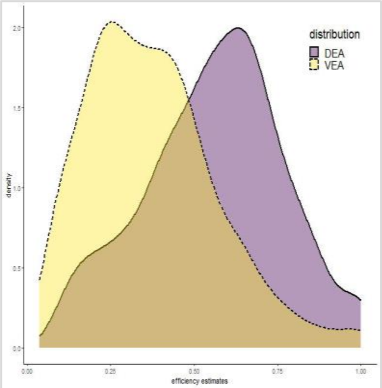
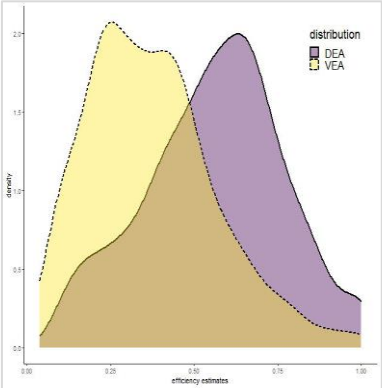
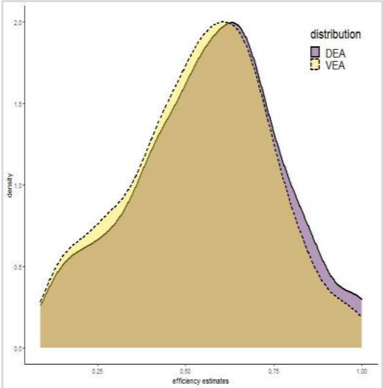
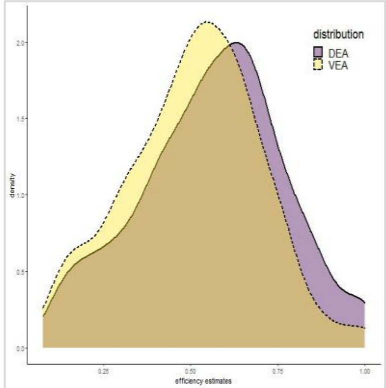
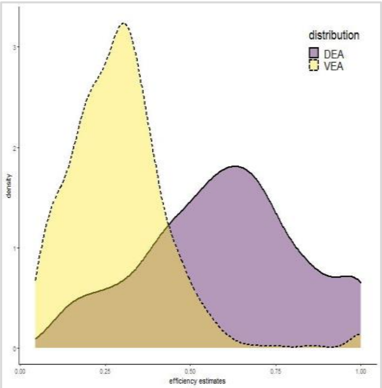
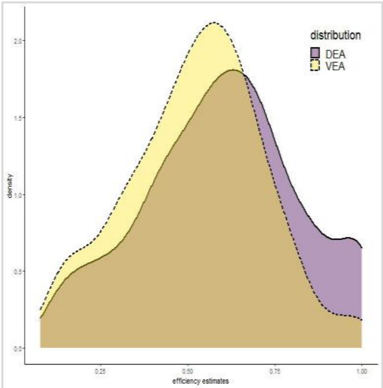
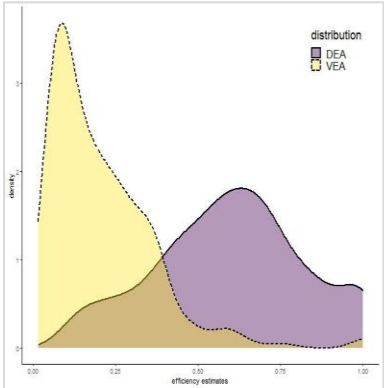
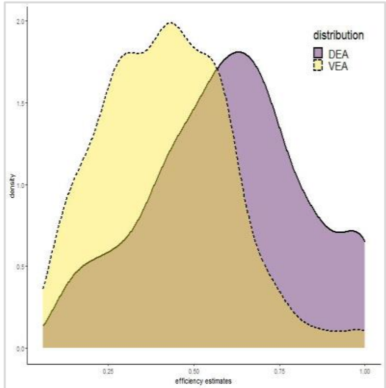
	6. minimum average CV	7. IDMU	8. interior active	9. interior self-evaluator	10. exterior active
CRS	 MPS: {183}	 MPS: {37,178,368}	 MPS: {293}	-	 MPS: {130}
VRS	 MPS: {142}	-	 MPS: {130}	 MPS: {216}	 MPS: {32}

Figure 4.2 (Cont.)

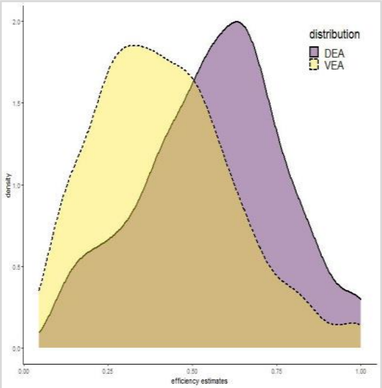
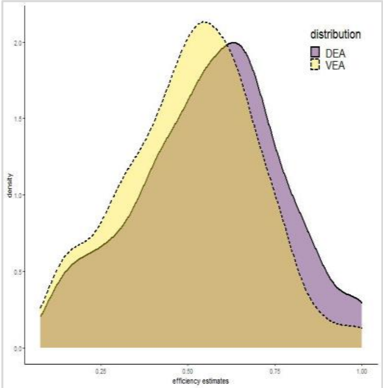
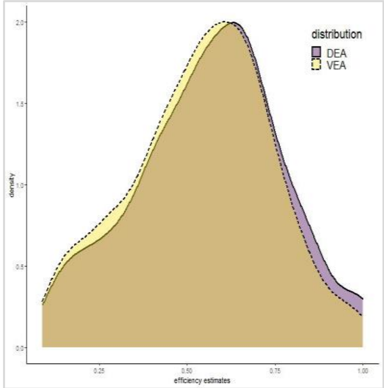
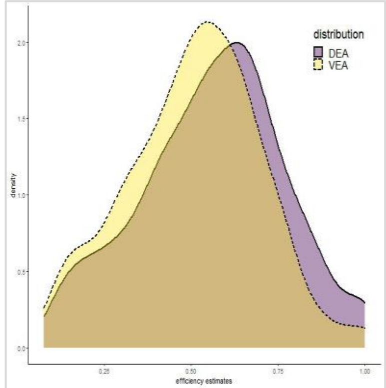
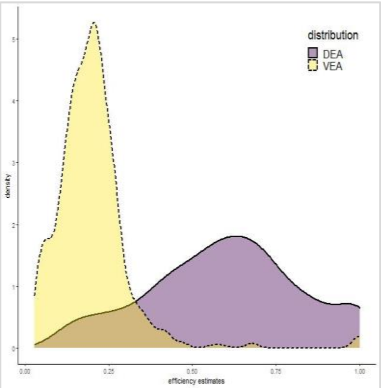
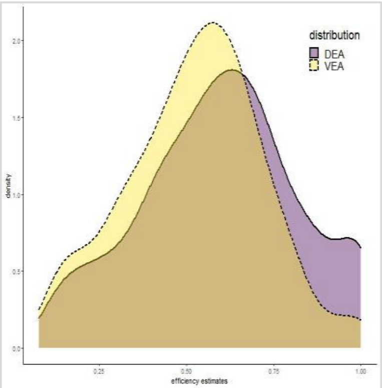
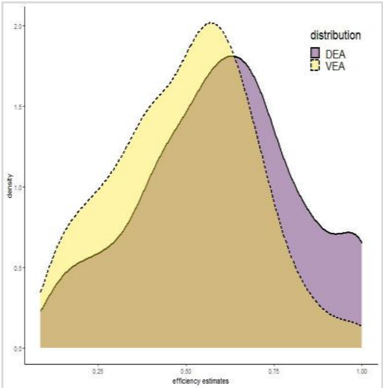
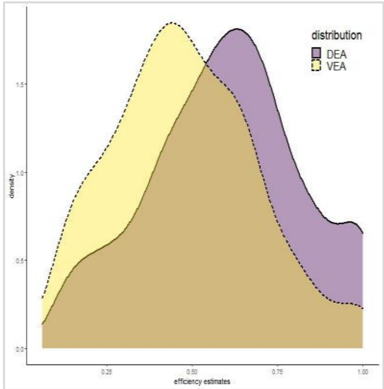
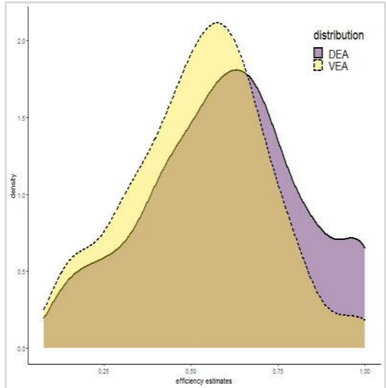
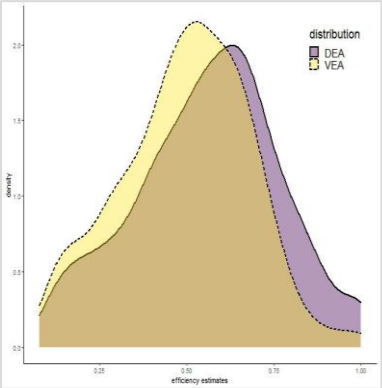
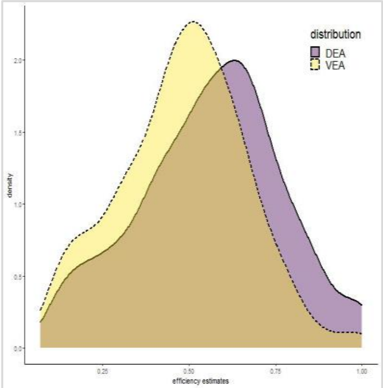
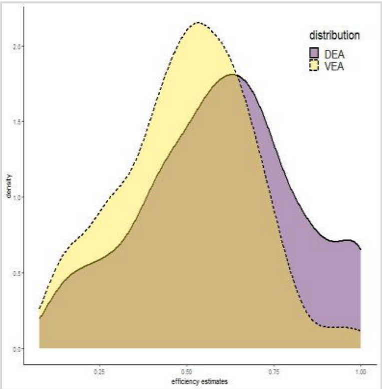
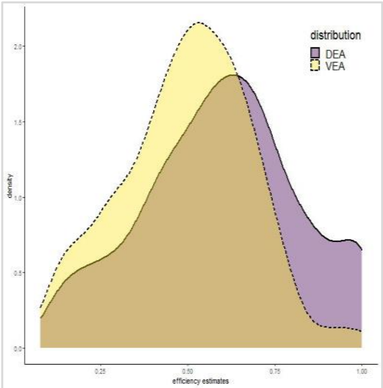
	11. exterior self-evaluator	12. interior terminal	13. exterior terminal	14. non-terminal	15. MPSS
CRS	-	 MPS: {154}	 MPS: {130}	 MPS: {293}	 MPS: {130}
VRS	 MPS: {119}	 MPS: {130}	 MPS: {241}	 MPS: {147}	 MPS: {130}

Figure 4.2 (Cont.)

	16. APU	17. FDEF
CRS	 MPS: {252,293,314,415}	 MPS: {130,178,252,415}
VRS	 MPS: {130,252,293,314}	 MPS: {130,252,293,314,415}

VEA model with the minimum "average CV" MPS choice as well, as the chosen farm is also an interior self-evaluator (see Table 4.2). In all three cases, the correlation between the VEA and the DEA efficiency scores is particularly low (see Table 4.6). In addition, when an exterior self-evaluator farm is the MPS, some of the inputs are irrelevant for the estimation of efficiency. More specifically, a zero value is assigned to the weights attached to land and capital by all farms.

Large leftward shifts in the VEA distribution of efficiency scores compared to DEA are observed for a series of other MPS choices. These are (i) the best-in-input farm and the farm for which the VRS super-efficiency model results in an infeasible solution (which in this case is the same farm), (ii) the farm with the minimum average CV with CRS, (iii) the IDMU peers, (iv) an interior terminal farm when CRS is assumed, and (iv) an exterior active farm with VRS. In all these cases, the VEA efficiency scores decrease, on average, by more than 30% compared to DEA (see Table 4.4). This suggests that the input/output bundle used by the MPS in each case is quite dissimilar from the bundles used by most of the farms. For MPSs with infeasible super-efficiency scores or the IDMU peers, this may often be expected, as the former are usually located at an "end-point" of the frontier, while the latter is likely to use a rather extreme input/output bundle.[73] In these two cases some of the inputs are irrelevant for the estimation of efficiency. This is true for capital in the case of the MPS with infeasible super-efficiency score and for land in the case of the IDMU peers, indicating that the farms are assessed by means of non-well defined marginal rates. For the remaining choices in this group, large differences between the DEA and the VEA efficiency scores may or may not be the case. For example, in our case there are large differences between the VEA and DEA distributions of efficiency scores not only when the chosen best-input farm is the MPS, but also if some of the other three best-in-input farms are used as the MPS instead.[74] This however may not occur in a different sample and/or model specifications.

---

[73] The other seven efficient farms for which the VRS DEA super-efficiency model results in an infeasible solution (see Table 4.2) have similar economic and socio-demographic characteristics to the selected MPS. When each of these farms is used as the MPS, the VEA distribution of efficiency scores differs in a statistically significant way from the DEA one.

[74] The other three VRS efficient farms that use the lowest quantities of land, capital and intermediate inputs (see Table 4.2) are have similar economic and socio-demographic characteristics to the selected

Table 4.6:Simple and rank correlation coefficients between DEA and VEA models

		simple correlation		Spearman rank correlation	
VEA	MPS choice	CRS	VRS	CRS	VRS
1	most times as peer in DEA	0.998	0.996	0.992	0.974
2	maximum finite superefficiency	0.998	0.995	0.991	0.979
3	infeasible superefficiency	-	0.855	-	0.730
4	best-in-output	0.998	0.996	0.992	0.974
5	best-in-input	-	0.855	-	0.730
6	minimum average CV	0.839	0.938	0.781	0.863
7	IDMU	0.843	-	0.782	-
8	interior active	0.993	0.972	0.981	0.888
9	interior self-evaluator	-	0.669	-	0.603
10	exterior active	0.979	0.920	0.921	0.795
11	exterior self-evaluator	-	0.933	-	0.820
12	interior terminal	0.869	0.972	0.815	0.888
13	exterior terminal	0.979	0.914	0.921	0.836
14	non-terminal	0.993	0.925	0.981	0.850
15	MPSS	0.979	0.972	0.921	0.888
16	APU	0.979	0.960	0.922	0.847
17	FDEF	0.975	0.960	0.920	0.847

Lastly, the VEA distribution of efficiency scores differs only moderately from DEA for the following MPS choices: (i) an exterior active farm in the CRS model, (ii) either an interior active or an interior terminal farm with VRS, (iii) the MPSS choice for both model specifications, (iv) an exterior terminal and a non-terminal farm with VRS, (v) common weights and (vi) the APU peers with CRS and VRS. In the first three of these cases, the same farm #130 is used as the MPS, while in all of them the differences between the DEA and the VEA distributions of efficiency scores are significant in a statistical sense (see Table 4.5). This indicates that, even though the changes in efficiency are moderate, the use of VEA does result in additional insights to management with respect to the results obtained from the DEA model. Among those cases, significant differences between the DEA and the VEA efficiency scores may be expected when the APU's peers are used as the MPS and in common weights VEA, but not necessarily for the other cases. In common weights VEA, only 21 (with CRS) and seven farms (with VRS) have efficiency scores equal to their corresponding DEA

farm. When they are used as the MPS, large and statistically significant differences are observed between the VEA and DEA efficiency scores.

scores, while for the remaining farms VEA efficiency scores decrease at least slightly compared to their DEA counterparts. The differences are moderate as the farms forming the chosen combination have input/output bundles that are similar to those of most DMUs in the sample but could be larger if a farm with a somewhat extreme bundle was chosen in the combination. When the APU's peers are the MPS, farms with an input/output structure that is close to the average, i.e., that of the APU, exhibit slight or no decreases in efficiency in the VEA model, while the scores of farms with extreme bundles efficiency scores decrease considerably. In addition, in the case of a non-terminal MPS, no inputs are irrelevant for the for the estimation of efficiency, while when an (interior or exterior) terminal farm is the MPS (either with CRS or with VRS) some farms assign a zero weight to one or more of the land, labor, and/or capital inputs.

4.5.3. Comparative results among VEA models

Comparing the efficiency distributions among VEA models with alternative MPSs can provide additional insights. No doubt, the VEA distributions of efficiency scores are the same among those MPS choices for which the same farm is used as the MPS. In our case, these are (i) a farm with an infeasible super-efficiency scores and a best-in-input farm, (ii) an exterior active, an exterior terminal and an MPSS farm with CRS, (iii) an interior active, an interior terminal and an MPSS farm with VRS, (iv), an interior active and a non-terminal farm with CRS, and (v) the farm appearing the most times as a peer in DEA and the best-in-output farm, for both model specifications (see Table 4.2).

In addition, based on Banker's (see Tables 4.7 and 4.8) and Mann-Whitney tests (Table 4.9),[75] correlation analysis (Tables 4.10 and 4.11), and shifts in rank (Table 4.12), we can infer that there are no significant differences among the efficiency scores of VEA models when the MPS is either (i) the farm appearing the most times as a peer in DEA, (ii) the farm with the maximum finite super-efficiency score, (iii) the best-in-output farm (both with CRS and with VRS), (iv) an interior active or (v) a non-terminal farm with CRS. This is to be expected in our case, as the VEA distributions of

[75] In the case of pairwise comparisons among different VEA models, the Banker test statistics are calculated by placing in the numerator the VEA model for which the sum of the logarithms of its inefficiency scores is the largest. This guarantees that the test statistic is always greater than or equal to one.

Table 4.7: Banker statistical tests among VEA models, constant-returns-to-scale

		VEA																
VEA	MPS choice	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17
1	most times as peer in DEA		1.010	-	1.000	-	1.678***	1.674***	1.039	-	1.129**	-	1.496***	1.129**	1.039	1.129**	1.158***	1.210***
2	lowest finite superinefficiency	1.015		-	1.010	-	1.662***	1.657***	1.028	-	1.117**	-	1.481***	1.117**	1.028	1.117**	1.146**	1.198***
3	infeasible superinefficiency	-	-		-	-	-	-	-	-	-	-	-	-	-	-	-	-
4	best-in-output	1.000	1.015	-		-	1.678***	1.674***	1.039	-	1.129**	-	1.496***	1.129**	1.039	1.129**	1.158***	1.210***
5	best-in-input	-	-	-	-		-	-	-	-	-	-	-	-	-	-	-	-
6	minimum average CV	2.432***	2.397***	-	2.432***	-		1.003	1.616***	-	1.487***	-	1.122**	1.487***	1.616***	1.487***	1.449***	1.387***
7	IDMU	2.413***	2.379***	-	2.413***	-	1.008		1.612***	-	1.483***	-	1.119**	1.483***	1.612***	1.483***	1.445***	1.383***
8	interior active	1.054	1.039	-	1.054	-	2.307***	2.290***		-	1.087*	-	1.441***	1.087*	1.000	1.087*	1.115**	1.165***
9	interior self-evaluator	-	-	-	-	-	-	-	-		-	-	-	-	-	-	-	-
10	exterior active	1.187**	1.17**	-	1.187**	-	2.050***	2.034***	1.126*	-		-	1.325***	1.000	1.087*	1.000	1.026	1.072
11	exterior self-evaluator	-	-	-	-	-	-	-	-	-	-		-	-	-	-	-	-
12	interior terminal	1.989***	1.961***	-	1.989***	-	1.223**	1.213**	1.887***	-	1.677***	-		1.325***	1.441***	1.325***	1.292***	1.236***
13	exterior terminal	1.187**	1.170**	-	1.187**	-	2.050***	2.034***	1.126*	-	1.000	-	1.677***		1.087*	1.000	1.026	1.072
14	non-terminal	1.054	1.039	-	1.054	-	2.307***	2.290***	1.000	-	1.126*	-	1.887***	1.126*		1.087*	1.115**	1.165***
15	DMU with MPSS	1.187**	1.170**	-	1.187**	-	2.050***	2.034***	1.126*	-	1.000	-	1.677***	1.000	1.126*		1.026	1.072
16	APU	1.227***	1.209**	-	1.227***	-	1.982***	1.967***	1.164**	-	1.034	-	1.621***	1.034	1.164**	1.034		1.045
17	FDEF	1.324***	1.305***	-	1.324***	-	1.837***	1.822***	1.256***	-	1.116	-	1.502***	1.116	1.256***	1.116	1.079	

Notes: (a) F1 (F2) test compares the VEA efficiency distributions of efficiency scores with each other assuming an exponential (half-normal) distribution of the efficiency scores (see Banker and Natarajan, 2011). (b) Results from the F1 (F2) test are depicted in the upper (lower) diagonal. (c) Three, two and one stars denote statistical significance at 1%, 5% and 10% respectively.

Table 4.8: Banker statistical tests among VEA models, variable-returns-to-scale

		VEA																
MPS	selection criterion	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17
1	most times as peer in DEA		1.031	1.859***	1.000	1.859***	2.198***	-	1.177***	2.938***	1.583***	2.780***	1.177***	1.275***	1.373***	1.177***	1.255***	1.260***
2	lowest finite superinefficiency	1.048		1.803***	1.031	1.803***	2.132***	-	1.141**	2.850***	1.536***	2.697***	1.141**	1.237***	1.332***	1.141**	1.218***	1.222***
3	infeasible superinefficiency	2.759***	2.632***		1.859***	1.000	1.182***	-	1.580***	1.581***	1.174***	1.495***	1.580***	1.458***	1.354***	1.580***	1.481***	1.476***
4	best-in-output	1.000	1.048	2.759***		1.859***	2.198***	-	1.177***	2.938***	1.583***	2.78***	1.177***	1.275***	1.373***	1.177***	1.255***	1.260***
5	best-in-input	2.759***	2.632***	1.000	2.759***		1.182***	-	1.580***	1.581***	1.174***	1.495***	1.580***	1.458***	1.354***	1.58***	1.481***	1.476***
6	minimum average CV	3.572***	3.409***	1.295***	3.572***	1.295***		-	1.868***	1.337***	1.388***	1.265***	1.868***	1.724***	1.601***	1.868***	1.751***	1.745***
7	IDMU	-	-	-	-	-	-		-	-	-	-	-	-	-	-	-	-
8	interior active	1.240***	1.183**	2.225***	1.240***	2.225***	2.881***	-		2.497***	1.345***	2.363***	1.000	1.084*	1.167***	1.000	1.067	1.071
9	interior self-evaluator	6.450***	6.154***	2.338***	6.450***	2.338***	1.805***	-	5.202***		1.856***	1.057	2.497***	2.304***	2.14***	2.497***	2.341***	2.333***
10	exterior active	2.042***	1.948***	1.351***	2.042***	1.351***	1.75***	-	1.647***	3.159***		1.756***	1.345***	1.241***	1.153**	1.345***	1.261***	1.257***
11	exterior self-evaluator	5.417***	5.168***	1.963***	5.417***	1.963***	1.516***	-	4.368***	1.191**	2.653***		2.363***	2.18***	2.025***	2.363***	2.215***	2.207***
12	interior terminal	1.240***	1.183**	2.225***	1.240***	2.225***	2.881***	-	1.000	5.202***	1.647***	4.368***		1.084*	1.167***	1.000	1.067	1.071
13	exterior terminal	1.427***	1.361***	1.934***	1.427***	1.934***	2.504***	-	1.151*	4.521***	1.431***	3.797***	1.151*		1.077	1.084*	1.016	1.012
14	non-terminal	1.653***	1.577***	1.669***	1.653***	1.669***	2.161***	-	1.333***	3.902***	1.235***	3.277***	1.333***	1.159**		1.167***	1.094**	1.09**
15	MPSS	1.240***	1.183**	2.225***	1.240***	2.225***	2.881***	-	1.000	5.202***	1.647***	4.368***	1.000	1.151*	1.333***		1.067	1.071
16	APU	1.358***	1.295***	2.032***	1.358***	2.032***	2.631***	-	1.095	4.751***	1.504***	3.99***	1.095	1.051	1.218**	1.095		1.004
17	FDEF	1.365***	1.303***	2.021***	1.365***	2.021***	2.616***	-	1.101	4.724***	1.495***	3.967***	1.101	1.045	1.211**	1.101	1.006	

Notes: (a) F1 (F2) test compares the VEA efficiency distributions of efficiency scores with each other assuming an exponential (half-normal) distribution of the efficiency scores (see Banker and Natarajan, 2011). (b) Results from the F1 (F2) test are depicted in the upper (lower) diagonal. (c) Three, two and one stars denote statistical significance at 1%, 5% and 10% respectively.

Table 4.9: Mann-Whitney statistical tests among VEA models

		VEA																
VEA	MPS choice	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17
1	most times as peer in DEA		0.807	15.844***	0.000	15.844***	20.738***	-	4.471***	23.383***	12.668***	24.320***	4.471***	6.663***	8.273***	4.471***	6.747***	6.863***
2	lowest finite superinefficiency	0.321		15.322***	0.807	15.322***	20.408***	-	3.914***	23.171***	12.016***	24.117***	3.914***	5.869***	7.577***	3.914***	5.912***	6.025***
3	infeasible superinefficiency	-	-		15.844***	0.000	6.629***	-	13.070***	14.555***	5.064***	15.456***	13.070***	11.131***	8.957***	13.070***	11.671***	11.589***
4	best-in-output	0.000	0.321	-		15.844***	20.738***	-	4.771***	23.383***	12.668***	24.320***	4.771***\	6.663***	8.273***	4.771***	6.747***	6.863***
5	best-in-input	-	-	-	-		6.629***	-	13.070***	14.555***	5.064***	15.456***	13.070***	11.131***	8.957***	13.070***	11.671***	11.589***
6	minimum average CV	13.864***	13.690***	-	13.864***	-		-	19.015***	10.424***	12.123***	12.129***	19.015***	17.302***	15.396***	19.015***	18.045***	17.993***
7	IDMU	13.867***	13.689***	-	13.867***	-	0.887		-	-	-	-	-	-	-	-	-	-
8	interior active	1.028	0.687	-	1.028	-	13.207***	13.969***		22.298***	9.165***	23.492***	0.000	2.178**	4.292***	0.000	2.145**	2.274**
9	interior self-evaluator	-	-	-	-	-	-	-	-		18.189***	1.675*	22.298***	21.299***	20.171***	22.298***	21.738***	21.704***
10	exterior active	3.796***	3.489***	-	3.796***	-	11.483***	11.481***	2.796***	-		19.852***	9.165***	6.958***	4.534***	9.165***	7.402***	7.309***
11	exterior self-evaluator	-	-	-	-	-	-	-	-	-	-		23.492***	22.568***	21.538***	23.492***	23.049***	23.016***
12	interior terminal	10.920***	10.735***	-	10.920***	-	3.616***	3.559***	10.179***	-	8.144***	-		2.178**	4.292***	0.000	2.145**	2.274**
13	exterior terminal	3.796***	3.489***	-	3.796***	-	11.483***	11.481***	2.796***	-	0.000	-	8.144***		2.175**	2.178**	0.126	0.006
14	non-terminal	1.028	0.687	-	1.028	-	13.207***	13.969***	0.000	-	2.796***	-	10.179***	2.796***		4.292***	2.456**	2.346**
15	MPSS	3.796***	3.489***	-	3.796***	-	11.483***	11.481***	2.796***	-	0.000	-	8.144***	0.000	2.796***		2.145**	2.274**
16	APU	4.764***	4.341***	-	4.764***	-	10.926***	10.936***	3.678***	-	0.879	-	7.510***	0.879	3.678***	0.879		0.141
17	FDEF	6.014***	5.728***	-	6.014***	-	9.823***	9.815***	5.075***	-	2.386**	-	6.268***	2.386**	5.075***	2.386**	1.533	

Notes: (a) Results for constant-returns-to-scale models are depicted in the lower diagonal, while those of variable-returns-to-scale models are depicted in the upper diagonal.
(b) Three, two and one stars denote statistical significance at 1%, 5% and 10% respectively.

Table 4.10: Simple and rank correlation coefficients among VEA models, constant-returns-to-scale

		VEA																
VEA	MPS choice	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17
1	most times as peer in DEA		0.981	-	1.000	-	0.806	0.808	0.979	-	0.929	-	0.840	0.929	0.979	0.929	0.928	0.931
2	maximum finite superefficiency	0.995		-	0.981	-	0.813	0.812	0.980	-	0.906	-	0.845	0.906	0.980	0.906	0.914	0.923
3	infeasible superefficiency	-	-		-	-	-	-	-	-	-	-	-	-	-	-	-	-
4	best-in-output	1.000	0.995	-		-	0.806	0.808	0.979	-	0.929	-	0.840	0.929	0.979	0.929	0.928	0.931
5	best-in-input	-	-	-	-		-	-	-	-	-	-	-	-	-	-	-	-
6	minimum average CV	0.849	0.852	-	0.849	-		0.997	0.774	-	0.709	-	0.996	0.709	0.774	0.709	0.699	0.809
7	IDMU	0.853	0.855	-	0.853	-	0.999		0.771	-	0.727	-	0.991	0.727	0.771	0.727	0.710	0.824
8	interior active	0.992	0.992	-	0.992	-	0.837	0.839		-	0.918	-	0.814	0.918	1.000	0.918	0.944	0.932
9	interior self-evaluator	-	-	-	-	-	-	-	-		-	-	-	-	-	-	-	-
10	exterior active	0.981	0.974	-	0.981	-	0.817	0.828	0.976	-		-	0.740	1.000	0.918	1.000	0.985	0.976
11	exterior self-evaluator	-	-	-	-	-	-	-	-	-	-		-	-	-	-	-	-
12	interior terminal	0.879	0.883	-	0.879	-	0.997	0.995	0.870	-	0.846	-		0.740	0.814	0.740	0.738	0.835
13	exterior terminal	0.981	0.974	-	0.981	-	0.817	0.828	0.976	-	1.000	-	0.846		0.918	1.000	0.985	0.976
14	non-terminal	0.992	0.992	-	0.992	-	0.837	0.839	1.000	-	0.976	-	0.870	0.976		0.918	0.944	0.932
15	MPSS	0.981	0.974	-	0.981	-	0.817	0.828	0.976	-	1.000	-	0.846	1.000	0.976		0.985	0.976
16	APU	0.980	0.976	-	0.980	-	0.812	0.819	0.986	-	0.995	-	0.844	0.995	0.986	0.995		0.973
17	FDEF	0.978	0.975	-	0.978	-	0.881	0.889	0.977	-	0.991	-	0.903	0.991	0.977	0.991	0.989	

Note: Simple correlation coefficients are depicted in the lower diagonal, while Spearman rank correlation coefficients are depicted in the upper diagonal.

Table 4.11: Simple and rank correlation coefficients among VEA models, variable-returns-to-scale

		VEA																
VEA	MPS choice	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17
1	most times as peer in DEA		0.954	0.785	1.000	0.785	0.838	-	0.912	0.666	0.840	0.808	0.912	0.864	0.872	0.912	0.883	0.883
2	maximum finite superefficiency	0.991		0.696	0.954	0.696	0.901	-	0.878	0.572	0.764	0.791	0.878	0.870	0.895	0.878	0.852	0.852
3	infeasible superefficiency	0.871	0.842		0.785	1.000	0.670	-	0.699	0.961	0.965	0.680	0.699	0.737	0.779	0.699	0.703	0.709
4	best-in-output	1.000	0.991	0.871		0.785	0.838	-	0.912	0.666	0.840	0.808	0.912	0.864	0.872	0.912	0.883	0.883
5	best-in-input	0.871	0.842	1.000	0.871		0.670	-	0.699	0.961	0.965	0.680	0.699	0.737	0.779	0.699	0.703	0.709
6	minimum average CV	0.932	0.955	0.811	0.932	0.811		-	0.753	0.565	0.728	0.758	0.753	0.842	0.903	0.753	0.745	0.747
7	IDMU	-	-	-	-	-	-		-	-	-	-	-	-	-	-	-	-
8	interior active	0.976	0.968	0.846	0.976	0.846	0.911	-		0.533	0.821	0.892	1.000	0.786	0.760	1.000	0.979	0.979
9	interior self-evaluator	0.693	0.658	0.940	0.693	0.940	0.645	-	0.635		0.870	0.490	0.533	0.662	0.729	0.533	0.546	0.552
10	exterior active	0.930	0.909	0.977	0.930	0.977	0.872	-	0.932	0.850		0.820	0.821	0.754	0.789	0.821	0.813	0.819
11	exterior self-evaluator	0.930	0.927	0.830	0.930	0.830	0.890	-	0.969	0.595	0.927		0.892	0.618	0.649	0.892	0.845	0.848
12	interior terminal	0.976	0.968	0.846	0.976	0.846	0.911	-	1.000	0.635	0.932	0.969		0.786	0.760	1.000	0.979	0.979
13	exterior terminal	0.920	0.927	0.819	0.920	0.819	0.936	-	0.886	0.695	0.851	0.801	0.886		0.879	0.786	0.840	0.840
14	non-terminal	0.929	0.943	0.856	0.929	0.856	0.960	-	0.891	0.751	0.882	0.828	0.891	0.924		0.760	0.756	0.759
15	MPSS	0.976	0.968	0.846	0.976	0.846	0.911	-	1.000	0.635	0.932	0.969	1.000	0.886	0.891		0.979	0.979
16	APU	0.967	0.958	0.844	0.967	0.844	0.913	-	0.993	0.640	0.925	0.950	0.993	0.915	0.890	0.993		1.000
17	FDEF	0.967	0.958	0.847	0.967	0.847	0.914	-	0.993	0.642	0.927	0.952	0.993	0.914	0.890	0.993	1.000	

Note: Simple correlation coefficients are depicted in the lower diagonal, while Spearman rank correlation coefficients are depicted in the upper diagonal.

Table 4.12: Average shifts in rank among VEA models

		VEA																
VEA	MPS choice	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17
1	most times as peer in DEA		22.004	71.713	0.000	71.713	59.446	-	44.461	92.600	61.522	63.991	44.461	56.157	55.061	44.461	49.108	48.945
2	lowest finite superinefficiency	15.298		85.120	22.004	85.120	46.336	-	49.586	103.975	74.157	65.985	49.586	52.727	49.370	49.586	54.376	54.201
3	infeasible superinefficiency	-	-		71.713	0.000	92.746	-	84.448	33.015	27.440	88.220	84.448	81.169	75.389	84.448	83.841	83.180
4	best-in-output	0.000	15.298	-		71.713	59.446	-	44.461	92.600	61.522	63.991	44.461	56.157	55.061	44.461	49.108	48.945
5	best-in-input	-	-	-	-		92.746	-	84.448	33.015	27.440	88.220	84.448	81.169	75.389	84.448	83.841	83.180
6	minimum average CV	71.061	69.201	-	71.061	-		-	75.710	108.632	82.380	75.269	75.710	59.165	45.727	75.710	77.632	77.324
7	IDMU	70.715	69.283	-	70.715	-	7.552		-	-	-	-	-	-	-	-	-	-
8	interior active	17.188	17.036	-	17.188	-	76.040	76.400		112.011	62.782	46.581	0.000	73.068	76.254	0.000	20.106	19.905
9	interior self-evaluator	-	-	-	-	-	-	-	-		59.509	117.685	112.011	92.814	85.256	112.011	110.309	109.748
10	exterior active	38.732	44.040	-	38.732	-	86.395	83.059	41.387	-		64.774	62.782	78.584	73.564	62.782	64.076	63.165
11	exterior self-evaluator	-	-	-	-	-	-	-	-	-	-		46.581	97.507	92.850	46.581	56.626	55.844
12	interior terminal	63.998	62.207	-	63.998	-	9.778	14.615	68.431	-	81.214	-		73.068	76.254	0.000	20.106	19.905
13	exterior terminal	38.732	44.040	-	38.732	-	86.395	83.059	41.387	-	0.000	-	81.214		50.584	73.068	62.135	62.343
14	non-terminal	17.188	17.036	-	17.188	-	76.040	76.400	0.000	-	41.387	-	68.431	41.387		76.254	75.632	75.385
15	MPSS	38.732	44.040	-	38.732	-	86.395	83.059	41.387	-	0.000	-	81.214	0.000	41.387		20.106	19.905
16	APU	37.643	41.681	-	37.643	-	87.822	85.806	32.951	-	17.313	-	81.507	17.313	32.951	17.313		1.484
17	FDEF	38.288	40.674	-	38.288	-	69.400	66.190	38.516	-	22.302	-	64.334	22.302	38.516	22.302	24.203	

Note: Results for constant-returns-to-scale models are depicted in the lower diagonal, while those of variable-returns-to-scale models are depicted in the upper diagonal.

efficiency scores for these MPS choices do not differ significantly from that of DEA. In our case, the best-in-output farm appears also as a peer for most farms, and the interior active and the non-terminal farms used as MPS are the in fact same farm. Also, the VEA efficiency scores do not differ in a statistically significant sense when the chosen MPS is based either on (i) the APU, (ii) an MPSS farm, or (iii) common weights. This is explained by the fact that the combinations of farms in the common weights and the APU choices in our case are similar to each other and include in most cases the MPSS farm #130, which however may not be the case with other datasets.

On the other hand, the VRS VEA distributions of efficiency scores differ in a statistically significant way from one another when the MPS is either (i) an interior self-evaluator farm, (ii) an exterior self-evaluator farm or (iii) a farm for which the super-efficiency score is infeasible. This indicates that statistically significant differences between VEA distributions of efficiency scores were found when the MPS choices reflect DMUs with a rather extreme input/output bundles. In these cases, as we have seen before, the VEA efficiency scores for each of these MPS choices are significantly different from those of DEA.

4.6. Concluding remarks

VEA can be a very useful tool for performance evaluation, providing guidance towards informed decision-making. The efficient frontier against which the DMUs are assessed in VEA depends on the chosen MPS. In this chapter, we first reviewed several MPS choices previously used in the literature. For some of these, there is a difficulty to intuitively explain the DMs' choice, as they do not explicitly consider some overall organizational objective, while others may compare DMUs against exceptionally performing benchmarks or inappropriate MRSs and MRTs. We then made four new suggestions for choosing the MPS: *First,* to make a more informed personal choice by explicitly considering the relative position of efficient DMUs on the DEA frontier. *Second,* choose a DMU with MPSS as the MPS, which results in assessing the DMUs against the technically optimal scale in DM's view. *Third,* choose the set of APU's peers as the MPS. In this case the resulting VEA scores resemble the extent of efficiency from the perspective of fully centralized management, and can be useful for DMs who coordinate resource allocation and pursue the objective of structural efficiency maximization. *Fourth,* to evaluate all DMUs based on common and strictly positive (i.e., well defined) weights, by choosing as the MPS a unique combination of DMUs generating an FDEF. This results in evaluating the DMUs against a common standard, which may be a prerequisite when management wishes to fully limit the assessed DMUs' autonomy in setting their own objectives.

The empirical comparative analysis using data on Greek cotton farmers provides useful results on how MPS choice may affect the VEA efficiency scores: *First*, the use of an influential peer as the MPS (the DMU appearing the most times as a peer and the one with the maximum finite super-efficiency score) does not offer additional insights to managers compared to the results obtain from the DEA model. *Second*, MPSs that are frequently located on "end-points" of the DEA frontier (those with an infeasible super-efficiency score and interior- or exterior-self-evaluators) appear to result in large differences on efficiency scores between the DEA and VEA models and in some of the inputs being irrelevant for the estimation of the VEA efficiency scores. *Third*, the use of both an (interior or exterior) terminal as well as a non-terminal DMU as MPS may result in significant differences between the DEA and VEA efficiency scores, but in the latter case all inputs were important for the estimation of efficiency while in the former case, zero optimal weights were assigned to some inputs. *Fourth*, both MPS choices pursuing minimum variability among the DMUs' optimal weights (minimum average CV and common weights) resulted in significant differences between the DEA and VEA efficiency scores. This may often be the case for the common weights choice. *Fifth,* the VEA scores when the MPS is either the APU or an MPSS DMU differ significantly, in a statistical sense, from that of the DEA model, which may often be the case for the APU.

On the other hand, the same VEA efficiency scores were obtained from different MPS choices for which the same DMU was used as the MPS, while similar scores were obtained from alternative MPS choices in which an influential peer is the MPS, namely the DMU appearing the most times as a peer and the DMU with the maximum finite super-efficiency score. Similarly, the VEA efficiency scores when the MPS was chosen based either on the APU, an MPSS farm, or common weights were not statistically different from each other. However, choices in which the MPS may often be a DMU with a rather extreme input/output bundle, namely self-evaluators and DMUs with infeasible super-efficiency scores, resulted in significantly different VEA scores with one another.

The empirical analysis conducted in this study provided the first thorough overview on the effect of MPS choice on the VEA scores. As our empirical findings may be data specific, a promising task for future research would to empirically assess the effect of MPS choice on VEA scores using data from other sectors and countries. Such studies could provide valuable insights that would complement those of the present study. Furthermore, as the incorporation of the MPS in VEA models restricts the assessed DMUs' choice of optimal values of input/output weights in a manner similar to that of introducing weight restrictions in DEA models, another avenue for future research would be to explore the relationship between VEA and weight-restricted DEA models in more detail.

Part II: Theoretical essays

CHAPTER 5

On Value Efficiency Analysis and Cross Efficiency

5.1. Introduction

Peer appraisal of Decision Making Units (DMUs) is involved in several cases of performance evaluation. These include, among others, cases of (a) participatory decision making (Oral, 2012), such as budget allocation; (b) zero-sum type of decisional contexts, where each DMU evaluates the remaining DMUs based on its own "value system"; and (c) instances in which transparency matters considerably for stakeholders, as with decision-making in international organizations or national government bodies (Oral, 2010) and faculty or institution appraisal in higher education (Oral *et al.,* 2014). In addition, peer appraisal may be desirable or necessary in the cases of players evaluation in sports and in the assessment of alternative portfolios of financial institutions.

Within Data Envelopment Analysis (DEA), cross efficiency (see Sexton *et al.,* 1986) and Value Efficiency Analysis (VEA) (see Halme *et al.,* 1999) are two popular frameworks used for peer appraisal assessments. In the former, the whole set of evaluated DMUs is involved, while in the latter only a subset, which are considered as the Most Preferred Solution (MPS) by an external Decision Maker (DM) or a central planner. Both these alternative peer appraisal frameworks rely on the input/output multipliers estimated by the conventional DEA model. In cross efficiency, each DMU is evaluated by using the vector of optimal multipliers of all other DMUs. For some DMUs, this vector may not be unique and to resolve this problem several secondary goal formulations have been proposed. These include the popular benevolent, aggressive (Sexton *et al.,* 1986; Doyle and Green, 1994; 1995), and neutral formulations (Wang and Chin, 2010a), which select one vector of multipliers among those optimal for the "reference" DMU and use it to obtain the cross efficiency scores of the remaining DMUs, as well as the less known Targeted Benevolence (TB) (Oral *et al.*, 1991) formulation.[76] In this less often employed but definitely useful formulation, all vectors of optimal multipliers of the "reference" DMU are used in the evaluation process as the remaining DMUs have the option of selecting the one maximizing their cross efficiency score.[77] This affirmative and appreciative form of appraisal enhances fairness and transparency in the evaluation process and provides each DMU with the most optimistic cross efficiency score. However, the estimation of the TB cross efficiency matrix is more complicated compared to that of other secondary goal formulations, and for this reason Doyle and Green (1995) felt that the use of TB may increase if shortcuts on its implementation are introduced. On the other hand, in VEA, each DMU is evaluated by means of a vector of multipliers selected among those that are optimal for the DMU(s) considered as the MPS, i.e., those reflecting better the DM's preferences about input and/or output mixes.[78]

In this paper, we examine how cross efficiency and VEA may be related to each other. By doing so we show that these two seemingly unrelated frameworks of peer appraisal are equivalent to each other for a particular formulation of cross efficiency, namely the TB. More specifically, we verify that the TB formulation is equivalent to VEA if either (i) an efficient "reference" DMU, i.e., the one whose vector of optimal multipliers is used to evaluate all other DMUs in the TB formulation, is also chosen as the MPS in VEA, or (ii) the radial projection of an inefficient "reference" DMU on the DEA frontier is also chosen as the MPS in VEA. This result implies that alternative interpretations of the TB and VEA efficiency scores can be derived and also, that one can obtain the matrix of the TB cross efficiency scores through a series of envelopment

[76] The term Targeted Benevolence was coined by Doyle and Green (1995) but it has also been referred to as the "Most Resonated Appreciative (MRA)" model by Oral et al. (2015) and as the "positively targeted peer-evaluation" model by Davtalab-Olyaie *et al.* (2021).

[77] TB has been used for, among others, project selection (Oral *et al.,* 1991; Oral, 2010), faculty evaluation in higher education (Oral *et al.,* 2014) and players evaluation in sports (Oukil and Govindaluri, 2017). The TB cross efficiency scores have also been employed (albeit usually referred to as "maximum cross efficiency scores") as an input to (i) approaches aiming to obtain a complete ranking of DMUs via cross efficiency (see Yang *et al.*, 2012; Oukil, 2020) and (ii) refinements of the conventional aggressive and benevolent formulations (Wu *et al.,* 2016a).

[78] Recent applications of VEA include, but are not limited to, the evaluation of hospital departments (Halme and Korhonen, 2000), academic institutions (Korhonen *et al.,* 2001) as well as bank branches (Eskelinen *et al.,* 2014).

form VEA models, each of which uses a different DMU or its efficient projection as the MPS. Besides of using the envelopment instead of the multiplier form, we show that the estimation of the TB cross efficiency matrix can be simplified further --if there are inefficient DMUs that are projected on the strongly efficient frontier and have the same set of peers. We provide the necessary steps for obtaining the TB cross efficiency matrix by means of VEA and we illustrate the usefulness of our findings using a number of examples.

The rest of this paper is organized as follows: In the next section we present the materials and methods used in the paper. The paper's main results are reported in the third section. The steps for estimating the TB cross efficiency matrix are given in the fourth section, while concluding remarks follow in the last section.

5.2. Materials and methods

5.2.1. Cross efficiency

Let us consider a set of K DMUs $(k=1,\ldots o,\ldots,K)$ operating under the same technology and producing a set of J $(j=1,\ldots,J)$ outputs by utilizing I $(i=1,\ldots,I)$ inputs. The fractional programming form of an input-oriented constant-returns-to-scale (CRS) DEA model for the o^{th} DMU is given as (Charnes *et al.*, 1978):[79]

$$\begin{aligned}
&\max_{\xi_j^o,\omega_i^o} \sum_{j=1}^{J} \xi_j^o y_j^o \Bigg/ \sum_{i=1}^{I} \omega_i^o x_i^o \\
&s.t.\ \sum_{j=1}^{J} \xi_j^o y_j^k \Bigg/ \sum_{i=1}^{I} \omega_i^o x_i^k \le 1 \qquad k = 1,\ldots,o,\ldots,K \\
&\qquad \xi_j^o \ge 0 \qquad\qquad\qquad\qquad j = 1,\ldots,J \\
&\qquad \omega_i^o \ge 0 \qquad\qquad\qquad\qquad i = 1,\ldots,I
\end{aligned} \tag{5.1}$$

where x and y are respectively the quantities of inputs and outputs and ω and ξ are their multipliers. Using the Charnes and Cooper (1962) transformation, (5.1) can be converted into the following linear model:

[79] We focus on input-oriented CRS DEA models but the extension of our results to output-oriented and variable-returns-to-scale (VRS) models is straightforward.

$$
\begin{aligned}
&\max_{u_j^o, v_i^o} \sum_{j=1}^{J} u_j^o y_j^o \\
&s.t. \sum_{j=1}^{J} u_j^o y_j^k - \sum_{i=1}^{I} v_i^o x_i^k \leq 0 \quad k = 1, \dots, o, \dots, K \\
&\quad \sum_{i=1}^{I} v_i^o x_i^o = 1 \\
&\quad u_j^o \geq 0 \qquad\qquad\qquad j = 1, \dots, J \\
&\quad v_i^o \geq 0 \qquad\qquad\qquad i = 1, \dots, I
\end{aligned}
\tag{5.2}
$$

where $u_j^o = \beta \xi_j^o$, $v_i^o = \beta \omega_i^o$ and $\beta = (\sum_{i=1}^{I} \omega_i^o x_i^o)^{-1}$.

The vector of optimal input/output multipliers provided by (5.2) is used to obtain the self-appraisal DEA efficiency score of the o^{th} DMU, $E_o^o = \sum_{j=1}^{J} v_j^o y_j^o / \sum_{i=1}^{I} u_i^o x_i^o$, and the peer-appraisal or cross efficiency score of the remaining DMUs, i.e., $E_o^h = \sum_{j=1}^{J} v_j^o y_j^h / \sum_{i=1}^{I} u_i^o x_i^h$.[80] By estimating (5.2) for each DMU we obtain the elements of the cross efficiency matrix, where each row contains a particular DMU's self-appraisal efficiency score (diagonal element) and the peer-appraisal efficiency scores of all other DMUs when appraised by that DMU (off-diagonal elements). If the values of all the optimal input/output multipliers are strictly positive, E_o^h can be interpreted as the conventional efficiency score relative to an extended facet reference technology based only on the input/output combinations of the DMUs residing in the efficient facet that is normal to this particular multiplier vector (Olesen, 2018).

The vector of optimal multipliers may not however be unique for the efficient DMUs ($E_o^o = 1$), and, in rare occasions, for some inefficient DMUs ($E_o^o < 1$) as well (Cooper *et al.*, 2007, p. 32). In these cases, there are multiple cross efficiency scores for the remaining DMUs, each of which is based on a different vector of multipliers among those that are optimal for the "reference" DMU. This poses a problem on which vector of multipliers to be used for peer appraisal purposes. The use of the one obtained from (5.2) for the "reference" DMU, as proposed by Sexton *et al.* (1986), is rather

[80] Notice that in E_o^h the superscript denotes the DMU being evaluated and the subscript the "reference" DMU whose optimal multipliers are used for peer appraisal purposes.

unsatisfactory, because it depends on arbitrary factors such as the order with which the data are entered into the linear optimization software used (Doyle and Green, 1994).

Several approaches have been developed to resolve this problem: (i) the secondary goal formulations, which include the aggressive and benevolent, the TB, weight profiles and ranking optimization approaches, (ii) the game theoretic approaches which comprise of the pareto optimality, non-cooperative game theory and the bargaining approaches, and (iii) the prospect theory approach. An overview of these approaches including the main references is provided in Table 5.1.[81]

In the secondary goal approaches, such as the benevolent and aggressive formulations, introduced in Doyle and Green (1995), the problem of non-unique optimal multiplier vectors for the “reference” DMU is resolved by modifying accordingly the objective function in (5.2) in order to result in a unique vector of optimal input and output multipliers, which is then used for peer appraisal of the remaining DMUs. More specifically, the benevolent (aggressive) formulation selects the vector of multipliers that maximizes (minimizes) the average cross efficiency score of the remaining DMUs, or the efficiency score of a composite DMU that is obtained by aggregating the inputs and the outputs of the remaining DMUs. In the latter case, the following linear programming model is solved for the benevolent formulation:

$$
\begin{aligned}
\max_{u_j^o, v_i^o} \quad & \sum_{j=1}^{J} u_j^o Y_j^h \\
s.t. \quad & \sum_{j=1}^{J} u_j^o y_j^k - \sum_{i=1}^{I} v_i^o x_i^k \leq 0 \quad k = 1, \dots, o, \dots, K, k \neq h \\
& \sum_{j=1}^{J} u_j^o y_j^h - \sum_{i=1}^{I} v_i^o \hat{x}_i^h = 0 \\
& \sum_{i=1}^{I} v_i^o X_i^h = 1 \\
& u_j^o \geq 0 \qquad\qquad j = 1, \dots, J \\
& v_i^o \geq 0 \qquad\qquad i = 1, \dots, I
\end{aligned}
\tag{5.3}
$$

[81] Recent theoretical extensions of cross efficiency measurement to other performance evaluation problems include, among others, cases where input/output data are uncertain (Pan *et al.*, 2021), clustering (Chen *et al.*, 2022), and economic efficiency evaluation (Aparicio and Zofio, 2021).

Table 5.1: Overview of approaches for estimating the DEA cross efficiency matrix

1. Using the optimal set of input/output multipliers obtained from the DEA model	Sexton *et al.* (1986)
2. Secondary goal models	
2.1. Aggressive/benevolent	
min/max the sum of cross efficiency scores or the score of an aggregate DMU	Doyle and Green (1995)
max/min the sum of deviations of cross efficiency scores from unity	Liang *et al.* (2008a)
max/min the sum of deviations of cross efficiency scores from DEA scores	Wang and Chin (2010a)
min(max) the best (worst) cross efficiency score	Lim (2012)
min or max the sum of deviations of cross efficiency scores from the maximum or minimum possible scores	Wu *et al.* (2016a; b)
lexicographic min/max the cross efficiency scores	Chen (2018)
min/max the number of DMUs for which cross efficiency score equals the DEA score	Lam (2010); Davtalab-Olyaie (2019)
2.2. Targetted Benevolence	Oral *et al.* (1991; 2015); Oral (2010)
2.3. Neutral	Wang and Chin (2010b); Wang *et al.* (2011a)
2.4. Weights profiles methods	
least dissimilar weights	Ramon *et al.* (2010)
common least dissimilar weights	Ramon *et al.* (2011)
percentage deviation from the mean	Lam and Bai (2011)
goal programming	Örkcü and Bal (2011); Al-Siyabi *et al.* (2019)
Ideal/Anti-ideal DMU method	Wang *et al.* (2011b); Carrillo and Jorge (2018); Shi *et al.* (2019)
minimum disparity between weight vectors	Wang *et al.* (2012)
iterative method	Lin *et al.* (2016)
efficient facets approach	Dellnitz *et al.* (2021)
interval reference point method	Shi *et al.* (2021)
hypervolume maximization	Alcaraz *et al.* (2022)
2.5. Ordinal evaluations/ranking optimization	Wu *et al.* (2009a); Contreras (2012)
3. Game theoretic approaches	
3.1. Pareto optimality	Wu *et al.* (2016c); Davtalab-Olyaie *et al.* (2021)
3.2. Non-cooperative game theory	Liang *et al.* (2008b); Wu *et al.* (2009b); Liu *et al.* (2017)
3.3. Bargaining approaches	Wu *et al.* (2009c); Contreras *et al.* (2021)
4. Prospect theory	Liu *et al.* (2019)

in which the h^{th} DMU is the "reference" DMU, $Y_j^h = \sum_{k \neq h} y_j^k\, , j = 1, \dots, J\, , X_j^h = \sum_{k \neq h} x_i^k\, , i = 1, \dots, I$ and $\hat{x}_i^h = E_h^h x_i^h$. The second constraint in (5.3) forces the optimization procedure to maintain the efficiency score of the "reference" DMU on its self-appraisal level while maximizing the efficiency score of the aggregate DMU.[82] Essentially, (5.3) evaluates this aggregate DMU against a reference technology consisting of the lower envelope of the extended efficient facets of the DEA frontier that intercept to each other at the point where the "reference" DMU is (if it is efficient) or where it is projected by the DEA model, if it is inefficient (Olesen, 2018). The resulting efficiency score can be interpreted as the minimum proportional input reduction required for the aggregate DMU to reach the frontier but is however of no particular use. Instead, the optimal vector of multipliers obtained from (5.3) is used to compute the cross efficiency scores for the remaining DMUs when the h^{th} DMU is the "reference" DMU. For the corresponding aggressive formulation, the objective function in (5.3) is changed from maximization to minimization.

On the other hand, in the TB cross efficiency formulation, each DMU is allowed to use that vector of optimal input/output multipliers of the "reference" DMU, among those in (5.2), that maximizes its cross efficiency score. Oral *et al.* (1991) modelled this by means of the following fractional programming model:

$$
\begin{aligned}
&\max_{\xi_j^o,\omega_i^o} \sum_{j=1}^{J} \xi_j^o y_j^o \Big/ \sum_{i=1}^{I} \omega_i^o x_i^o \\
&s.t. \ \sum_{j=1}^{J} \xi_j^o y_j^k \Big/ \sum_{i=1}^{I} \omega_i^o x_i^k \leq 1 \quad k = 1, \dots, o, \dots, K, \ \ k \neq h \\
&\qquad \sum_{j=1}^{J} \xi_j^o y_j^h \Big/ \sum_{i=1}^{I} \omega_i^o x_i^h = E_h^h \\
&\qquad \xi_j^o \geq 0 \qquad\qquad\qquad j = 1, \dots, J \\
&\qquad \omega_i^o \geq 0 \qquad\qquad\qquad i = 1, \dots, I
\end{aligned}
\tag{5.4}
$$

which, using the Charnes and Cooper (1962) transformation, can be converted into a linear model as follows:

[82] Notice that using an "average" instead of an aggregate DMU, i.e., replacing Y_j^h with $\bar{y}_j^h = (1/K-1)\sum_{k\neq h} y_j^k$, $j = 1, \dots, J$ and X_j^h with $\bar{x}_j^h = (1/K-1)\sum_{k\neq h} x_i^k$, $i = 1, \dots, I$ will not affect the results, as long as CRS is maintained

$$
\begin{aligned}
&\max_{u_j^o, v_i^o} \sum_{j=1}^{J} u_j^o y_j^o \\
&s.t. \ \sum_{j=1}^{J} u_j^o y_j^k - \sum_{i=1}^{I} v_i^o x_i^k \leq 0 \qquad k = 1, \dots, o, \dots, K, \ \ k \neq h \\
&\sum_{j=1}^{J} u_j^o y_j^h - \sum_{i=1}^{I} v_i^o \hat{x}_i^h = 0 \\
&\sum_{i=1}^{I} v_i^o x_i^o = 1 \\
&u_j^o \geq 0 \qquad j = 1, \dots, J \\
&v_i^o \geq 0 \qquad i = 1, \dots, I
\end{aligned} \tag{5.5}
$$

where $\hat{x}_i^h = E_h^h x_i^h$. As in (5.3), the second constraint in (5.5) forces the optimization procedure to maintain the efficiency score of the "reference" DMU on its self-appraisal level. However, in contrast with (5.3), (5.5) aims at maximizing the cross efficiency score *of the evaluated DMU* (Oral *et al.*, 1991). Thus, (5.5) evaluates each of the remaining DMUs against the frontier which is used in (5.3) to evaluate the aggregate DMU. In other words, (5.5) applies the basic principle of DEA (i.e., selection of the most favorable multipliers) in cross efficiency as well, allowing for an affirmative, fair, and transparent peer appraisal. The cross efficiency scores obtained from (5.5) are greater than or equal to the corresponding scores from any other secondary goal formulation (Davtalab-Olyaie *et al.*, 2021); that is, the TB cross efficiency scores provide the most optimistic peer appraisal evaluation.[83] The resulted efficiency scores can also be interpreted as the minimum proportional input reduction required by each of the remaining DMUs to reach the frontier consisting of the lower envelope of the extended efficient facets of the DEA efficient frontier that are normal to the vectors of multipliers optimal for the "reference" DMU. In contrast to the benevolent and the aggressive secondary goal formulations, (5.5) needs to be solved $K \times (K-1)$ times to obtain the cross efficiency matrix (Oral, 2010). For this reason, Doyle and Green (1995) stressed the need to "spot where shortcuts may be taken" in estimating the TB cross efficiency matrix.

[83] See Davtalab-Olyaie *et al.* (2021) Theorem 1.

5.2.2. *Value Efficiency Analysis*

In VEA, peer appraisal is conducted by evaluating the performance of all other DMUs with respect to the chosen MPS. The MPS reflects DM preferences over the most desirable input/output structure, in that it maximizes the DM's implicitly known pseudoconcave value function (Korhonen *et al.*, 2002).[84] In practice, it is explicitly chosen by the DM, and several criteria have been used in the VEA literature for choosing the MPS (see the fourth chapter in this Thesis for a review). DM preferences on the most desirable structure are then incorporated in the VEA model by essentially forcing the chosen MPS to be in the set of peers for every DMU. This is accomplished by simply turning the inequality constraint corresponding to the MPS in (5.2) to a strict equality. Assuming that the h^{th} DMU has been chosen as the MPS, the input-oriented CRS VEA model for the o^{th} DMU is given as (Halme *et al.*, 1999):

$$\begin{aligned}
&\max_{u_j^o, v_i^o} \sum_{j=1}^{J} u_j^o y_j^o \\
&s.t. \ \sum_{j=1}^{J} u_j^o y_j^k - \sum_{i=1}^{I} v_i^o x_i^k \leq 0 \quad k = 1, \dots, o, \dots, K, \ \ k \neq h \\
&\qquad \sum_{j=1}^{J} u_j^o y_j^h - \sum_{i=1}^{I} v_i^o x_i^h = 0 \\
&\qquad \sum_{i=1}^{I} v_i^o x_i^o = 1 \\
&\qquad u_j^o \geq 0 \qquad\qquad j = 1, \dots, J \\
&\qquad v_i^o \geq 0 \qquad\qquad i = 1, \dots, I
\end{aligned} \tag{5.6}$$

The second constraint in (5.6) essentially forces the evaluated DMU to choose, among the (possibly multiple) vectors of input/output multipliers that are optimal for the MPS in (5.2), the one maximizing its efficiency score. As each of these vectors is normal to an efficient facet generated (partly) by the MPS, the resulting VEA frontier is in essence the lower envelope of the extended efficient facets intercepting at the MPS. If the o^{th} DMU shares at least one optimal vector of input/output multipliers with the MPS, its peer-appraisal score obtained from (5.6) will be the most optimistic, i.e., equal to its

[84] See Joro and Korhonen (2015) for a detailed treatment of VEA.

self-appraisal DEA efficiency score E_O^o. Otherwise, (5.6) will assign to the DMU a score lower than E_O^o.

Model (5.6) does not have a feasible solution if the chosen MPS is not a DEA-efficient unit. To avoid this, one may either use its peers identified by the dual form of (5.2) as the MPS (Halme *et al.*, 1999) and turn their respective inequality constraints in (5.2) into equalities or rely on its radial projection (Joro and Korhonen, 2015, p. 176), i.e., substitute x_i^h by $E_h^h x_i^h$ in the second constraint in (5.6). Both result in the same efficiency scores if the inefficient DMU chosen as the MPS is radially projected on a part of the strongly DEA efficient frontier. These DMUs can be identified by estimating the so-called Phase II DEA model than maximizes the sum of input and output slacks for each evaluated DMU while substituting x_i^o by $E_O^o x_i^o$ (see, e.g., Cooper *et al.*, 2007a, pp. 44-45). For each such DMU, the optimal sum of slacks will be equal to zero, meaning that the coordinates of its radial projection $\left(E_h^h x_i^h, y_j^h\right)$ are equal to a linear combination of the input and output values of its peers. Then turning the inequality constraint into an equality for the radial projection of the h^{th} DMU in (5.6) is equivalent to turning the inequality constraint into an equality for each of its peers. If instead the h^{th} DMU is projected on a part of the weakly efficient DEA frontier, each vector of optimal multipliers is associated with at least a zero value for an input or an output, and the coordinates of its radial projection is not equal to a linear combination of the input and output values of its peers. For the latter, there are optimal multiplier vectors in which all values are strictly positive. Then, using in (5.6) the radial projection of the h^{th} DMU and its peers as the MPS does not result in evaluating the DMUs using the same vectors of optimal multipliers, and, consequently, (5.6) will not necessarily produce the same efficiency scores for each DMU.

5.2.3. A motivating example

To illustrate the notion of peer appraisal through cross efficiency and VEA and the relations between them, let us consider a small numerical example with 8 DMUs each using two inputs to produce a single output. The relevant data are given in columns (2) to (4) of Table 5.2, the efficiency scores are given in column (5),), the vectors of the (normalized) input and output multipliers $\left(\frac{u_1^o}{v^o}, \frac{u_1^o}{v^o}, 1.000\right)$ are given in columns (6) to (8) along with the efficient facet to which each of these vectors is normal to (column

Table 5.2: Data and efficiency scores for the illustrative example

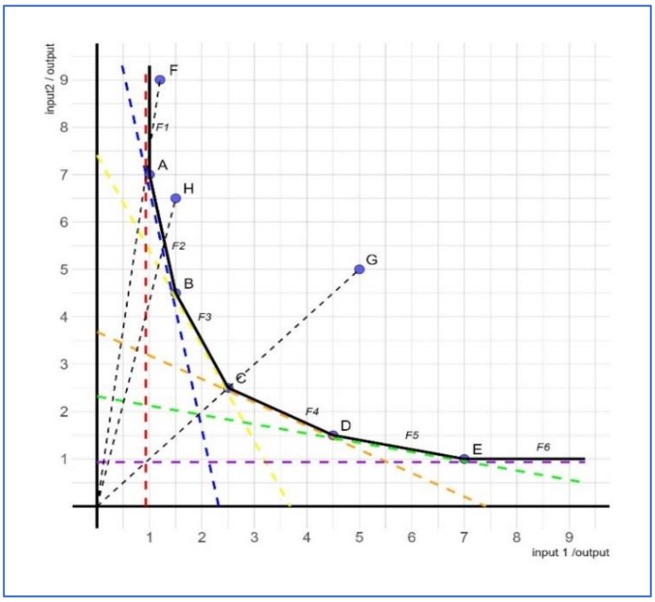
data				DEA optimal solution				
DMU	x_1	x_2	y	θ	u_1^o/v^o	u_1^o/v^o	facet	peers
A	1	7	1	1.000	0.417	0.083	F_2	A
B	1.5	4.5	1	1.000	0.267	0.133	F_3	B
C	2.5	2.5	1	1.000	0.267	0.133	F_3	C
D	4.5	1.5	1	1.000	0.133	0.267	F_4	D
E	7	1	1	1.000	0.083	0.417	F_5	E
F	1.2	9	1	0.833	1.000	0.000	F_1	A
G	5	5	1	0.500	0.267	0.133	F_3	C
H	1.5	6.5	1	0.857	0.417	0.083	F_2	A, B

(9)), while the DMUs' peers identified by estimating the dual of (5.2) are given in column (10). The resulting efficient frontier is portrayed in Figure 5.1a and consists of six facets identified by $F_l, l = 1, .., 6$. The supporting hyperplanes for these facets are depicted by the colored dashed lines. Estimating (5.2) for the extreme-efficient DMU A results in an optimal multiplier vector $\left(\frac{u_1^A}{v^A}, \frac{u_1^A}{v^A}, 1.000\right) = (0.417, 0.083, 1.000)$ that is normal to facet F_2. An alternative optimal solution for DMU A corresponds to the multiplier vector $(1.000, 0.000, 1.000)$ that is normal to facet F_1. In a similar manner, alternative optimal solutions exist for all the remaining extreme-efficient DMUs (i.e., *B*, *C*, *D*, and *E*) and the DEA-inefficient DMU *G*. For *G*, this occurs because it is projected on point *C*, in which two facets intercept. In contrast, a single optimal multiplier vector exists for the DEA-inefficient DMUs *F* and *H*.

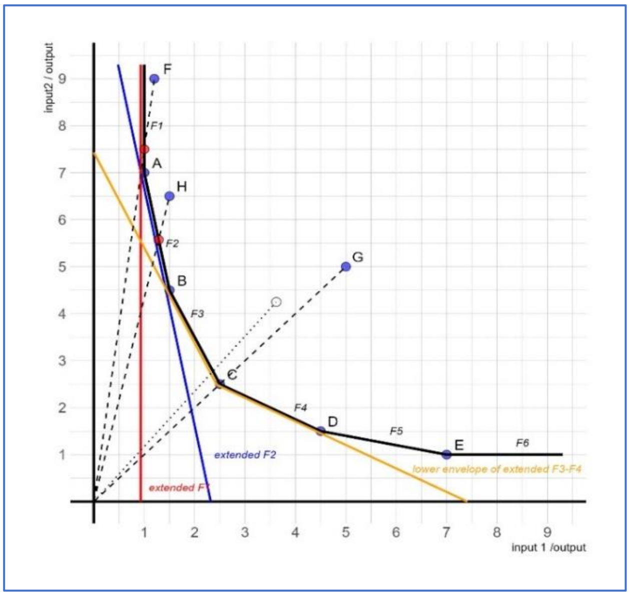
Using the optimal multiplier vectors obtained from (5.2) for each DMU to compute the peer appraisal cross efficiency scores for the remaining DMUs results in the cross efficiency matrix given in the upper panel of Table 5.3. For example, the first column corresponding to DMU A as the "reference" DMU, is obtained using the vector $(0.417, 0.083, 1.000)$ as $E_A^k = 1/\left(0.417x_1^k + 0.083y_2^k\right), k \neq A$. However, the alternative optimal multiplier vector for DMU A could have been used instead to obtain $E_A^k,\ k \neq A$. This is also true for DMUs *B*, *C*, *D*, *E*, and *G*.

The TB cross efficiency matrix is given in the lower panel of Table 5.3, while the procedure for obtaining the TB cross efficiency scores is portrayed in Figure 5.1b. For example, when the DEA-efficient DMU *C* is the "reference" DMU, (5.5) allows each of the remaining DMUs to choose among the multiplier vectors normal to facets

Figure 5.1: Efficient frontier and geometric representation of the TB secondary goal formulation



(a) Efficient frontier and efficient facets



(b) The TB formulation

F_3 and F_4, the one maximizing their cross efficiency score. For some DMUs (i.e., *A*, *B*, *F* and *H*) this is the vector normal to F_3 while for others (i.e., *D*, *E*) it is the one normal to F_4. Thus, the DMUs are in essence evaluated against the lower envelope of the extended facets F_3 and F_4 (the yellow solid piecewise linear frontier).[85] When the

[85] In the benevolent formulation instead, the DMUs in this case would be evaluated against only the extended facet F_3. Model (5.3) would be used to identify which of the multiplier vectors normal to F_3

Table 5.3: Cross efficiency matrices for the illustrative example

	reference DMU							
	A	B	C	D	E	F	G	H
multiplier sets obtained from DEA								
A	1.000	0.833	0.833	0.500	0.333	1.000	0.833	1.000
B	1.000	1.000	1.000	0.714	0.500	0.667	1.000	1.000
C	0.800	1.000	1.000	1.000	0.800	0.400	1.000	0.800
D	0.500	0.714	0.714	1.000	1.000	0.222	0.714	0.500
E	0.333	0.500	0.500	0.833	1.000	0.143	0.500	0.333
F	0.800	0.658	0.658	0.391	0.260	0.833	0.658	0.800
G	0.400	0.500	0.500	0.500	0.400	0.200	0.500	0.400
H	0.857	0.789	0.789	0.517	0.353	0.667	0.789	0.857
TB secondary goal								
A	1.000	1.000	0.833	0.500	0.143	1.000	0.833	1.000
B	1.000	1.000	1.000	0.714	0.222	0.667	1.000	1.000
C	0.800	1.000	1.000	1.000	0.400	0.400	1.000	0.800
D	0.500	0.714	1.000	1.000	0.667	0.222	1.000	0.500
E	0.333	0.500	0.833	1.000	1.000	0.143	0.833	0.333
F	0.833	0.800	0.658	0.391	0.111	0.833	0.658	0.800
G	0.400	0.500	0.500	0.500	0.200	0.200	0.500	0.400
H	0.857	0.857	0.789	0.517	0.154	0.667	0.789	0.857

DEA-inefficient DMU K is the "reference" DMU, we take its radial projection. This is on the interior of facet F_2 and thus there is only one optimal multiplier vector for DMU K. The associated TB cross efficiency scores reflect the radial reduction of inputs for the remaining DMUs to reach the extended facet F_2 (the blue solid line). In a similar manner, DMU G is projected on facet F_1 and thus the TB cross efficiency scores of the remaining DMUs are obtained as the radial distance to the extended facet F_1 (the red solid line). On the contrary, the projection of DMU H to the efficient frontier coincides with DMU C which is on the intersection of facets F_3 and F_4. Thus, when H is the "reference" DMU, the remaining DMUs can choose between the multiplier vectors that are normal to F_3 and F_4 the one maximizing their efficiency scores.

On the contrary, when VEA is used for peer appraisal purposes, one needs first

and F_4 maximizes the efficiency of the DMU constructed as the average input and output quantities of all sample DMUs except C (the white dot in Figure 5.1a). This artificial DMU is projected on the interior of F_3, and thus the associated multiplier vector would be used to compute the peer appraisal scores of all other DMUs when DMU C is the "reference" DMU. In a similar fashion, the aggressive formulation would evaluate all DMUs against the extended facet F_4.

to choose the MPS. This is frequently chosen among the DEA-efficient DMUs, i.e., DMUs *A*, *B*, *C*, *D*, and *E*. The VEA efficiency scores when each of these DMUs is the MPS are given in columns (2) to (6) of Table 5.4. The VEA efficient frontier is portrayed, when DMU *C* is chosen as the MPS, by the solid yellow piecewise line in Figure 5.2a. In this case, turning the inequality constraint corresponding to *C* into an equality in (5.6) forces the evaluated DMUs to maximize their efficiency score by choosing a multiplier vector among those optimal for *C* (i.e., a vector that is normal to either F_3 or F_4).

Alternatively, one of the DEA-inefficient DMUs *F*, *G*, and *H* could be chosen as the MPS, in which case their peers or their radial projection on the DEA frontier should be used as the MPS instead. DMUs *H* and *G* are projected on the strongly efficient frontier and thus, for these DMUs the two options result in the same VEA scores. DMU *H's* projection is on the interior of facet F_2 and if used as the MPS in (5.6) the resulting VEA frontier is the extended facet F_2 (the solid blue line in Figure 5.2a). Using DMU *H*'s peers (*A* and *B*) as the MPS, it results in the same efficiency scores (see columns (7) and (8) of Table 5.4). This is because turning in (5.6) the inequality constraints corresponding to *both A* and *B* into equalities means that the evaluated DMUs are forced to choose a multiplier vector that is optimal for *both* A and B.[86] There is only one such vector and is normal to F_2. Similarly, when DMU *G* is chosen as the MPS, using its projection as the MPS in (5.6) results in the same VEA scores (given in column (9) of Table 5.4) as using its peer (DMU *C*) as the MPS. On the other hand, DMU *F* is projected on facet F_1, that is not part of the strongly efficient frontier. The VEA scores when its radial projection is used as the MPS in (5.6) (given in column (10) of Table 5.4) are different from the ones resulting when its peer (DMU *A*) is chosen as the MPS. In the former case, the VEA frontier is the extended facet F_1 (red solid line, see Figure 2b) while in the latter case it is the lower envelope of the extended), facets F_1 and F_2 (piecewise linear blue line).

A comparison of Tables 5.3 and 5.4 shows that the TB cross efficiency scores are related to those resulting from VEA for particular MPS choices. More specifically,

---

[86] In other words, the equality constraint in (5.6) concerning the projection of DMU *H* can be expressed as a convex combination of (and replaced by) two similar equalities concerning DMUs *A* and *B*, as *H*'s projection is in essence a combination of *A* and *B*.

Table 5.4: VEA efficiency scores based on different MPS specifications for the illustrative example

	MPS								
	A	B	C	D	E	H'	$\{A,B\}$	G'	F'
A	1.000	1.000	0.833	0.500	0.143	1.000	1.000	0.833	1.000
B	1.000	1.000	1.000	0.714	0.222	1.000	1.000	1.000	0.667
C	0.800	1.000	1.000	1.000	0.400	0.800	0.800	1.000	0.400
D	0.500	0.714	1.000	1.000	0.667	0.500	0.500	1.000	0.222
E	0.333	0.500	0.833	1.000	1.000	0.333	0.333	0.833	0.143
F	0.833	0.800	0.658	0.391	0.111	0.800	0.800	0.658	0.833
G	0.400	0.500	0.500	0.500	0.200	0.400	0.400	0.500	0.200
H	0.857	0.857	0.789	0.517	0.154	0.857	0.857	0.789	0.667

Note: An apostrophe denotes a DMU's radial projection on the DEA frontier.

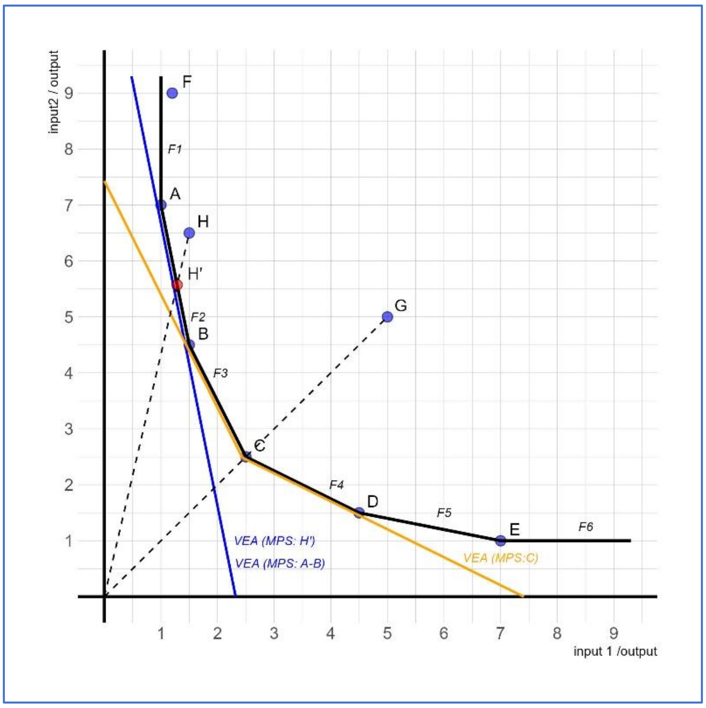
the TB cross efficiency scores when each of the DEA-efficient DMUs *A*, *B*, *C*, *D*, and *E* is the "reference" DMU are equal to the VEA scores when each of these DMUs is used as the MPS. The TB cross efficiency scores when the DEA-inefficient DMU *G* is the "reference" DMU are also equal to the VEA scores when *G*'s radial projections (or its peers) is the MPS, and the same holds for DMU *H*. These two DMUs are projected on a part of the strongly efficient DEA frontier. This is not however the case for the DEA-inefficient DMU *F*. When *F* --which is projected on a part of the weakly efficient frontier-- is the is the "reference" DMU, the TB cross efficiency scores are equal to the VEA scores when *F*'s projection is used as the MPS but are different from the scores obtained using *F*'s peer (DMU *A*) as the MPS.

5.3. Theoretical results and implications

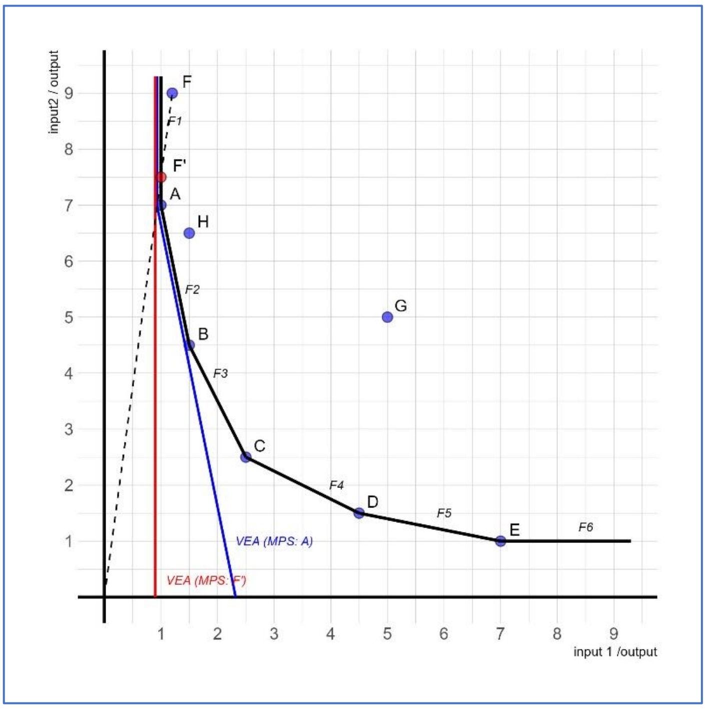
From the above example, it seems that the TB formulation in (5.5) and the VEA model in (5.6) result in the same efficiency score when the same DMU is chosen as the "reference" DMU in the former and as the MPS in the latter, provided that for a DEA-inefficient DMU its radial projection is used as the MPS in VEA. In this section, we provide a theoretical proof for this equivalence and explore its implications. In particular, we show that the following Proposition holds:

PROPOSITION 5.1: The TB formulation is equivalent to VEA, if either (i) an efficient "reference" DMU in the TB formulation is chosen as the MPS in VEA, or (ii) the radial projection of an inefficient "reference" DMU in the TB formulation is chosen as the MPS in VEA

Figure 5.2: Geometric representation of Value Efficiency Analysis



(a) VEA for alternative MPS specifications



(b) Using a DEA-inefficient DMU that is associated with slacks as the MPS

Proof: Case (i): If the "reference' DMU in the TB formulation is an efficient DMU then $\hat{x}_i^h = x_i^h$ and the equality constraint in (55.) may be re-written as:

$$\sum_{j=1}^{J} u_j^o y_j^h - \sum_{i=1}^{I} v_i^o x_i^h = 0 \tag{5.7}$$

Then, (5.5) and (5.6) are equivalent to each other. Case (ii): If the DMU chosen as the MPS in VEA is inefficient and its radial projection is used as the MPS, x_i^h in the first equality constraint in (5.6) should be replaced by $E_h^h x_i^h$ to ensure a feasible solution. Then, (5.5) and (5.6) are also equivalent to each other.

Notice that, if variable returns to scale are assumed (see Banker *et al.*, 1984), a free variable u^o is simply introduced in the objective function and the first two sets of constraints in (5.5) and (5.6) without affecting the above results.[87] In a similar fashion, if increasing (decreasing) returns to scale are assumed, then the introduced variable u^o is further restricted to be less-than-or-equal (greater-than-or-equal) to zero. This completes the proof. □

Two immediate implications of this result are the following: *first*, we can derive alternative interpretations of the VEA and TB efficiency scores. In particular, the VEA scores can be interpreted as the most favorable (i.e., the TB) cross-efficiency scores from the perspective of a particular "reference" DMU, namely the one chosen as the MPS in VEA, while the TB cross-efficiency scores when a particular DMU is used as a "reference" reflect also the judgements of a DM that this "reference" DMU has the most desirable input/output structure.

Second, the scores in the *o*th column of the TB cross efficiency matrix, i.e., the peer appraisal scores of the *o*th DMU by all the other DMUs, can be obtained by estimating a series of VEA linear programs using each DMU, if it is efficient *or* its radial projection on the DEA frontier if it is inefficient, as the MPS. Repeating this for all DMUs we can obtain the matrix of the TB cross efficiency scores. Consequently, the TB cross efficiency scores can be obtained from the envelopment form of VEA, namely:

---

[87] It has been shown that, in VRS models, some cross efficiency scores may be equal to zero or even take negative values (see Soares de Mello *et al.* (2013) for further elaboration on this matter). The occurrence of such peculiar efficiency scores does not affect the validity of our results. Instead, according to our Proposition, if the TB cross efficiency scores for a particular "reference" DMU include some peculiar values, these will also occur to the corresponding VRS VEA model. Korhonen *et al.* (2002) were the first to notice that negative scores may occur in VRS VEA.

$$
\begin{aligned}
&\min_{E_h^o,\lambda_o^k} \ E_h^o \\
&s.t. \ \sum_{k=1}^{K} \lambda_k^o y_j^k \geq y_j^o \qquad j = 1, \dots, J \\
&\qquad \sum_{k \neq h} \lambda_k^o x_i^k + \lambda_h^o x_i^h \leq E_h^o x_i^o \quad i = 1, \dots, I \\
&\qquad \lambda_k^o \geq 0 \qquad\qquad k = 1, \dots k, \dots, K, \ \ k \neq h \\
&\qquad \lambda_h^o \ free \\
&\qquad E_h^o \ free
\end{aligned}
\tag{5.8}
$$

if the MPS is a DEA-efficient DMU. In (5.8), the intensity variable corresponding to the MPS is a free variable instead of being non-negative as in conventional DEA models. If, on the other hand, the MPS is a DEA-inefficient DMU, then the envelopment form of VEA becomes:

$$
\begin{aligned}
&\min_{E_h^o,\lambda_o^k} \ E_h^o \\
&s.t. \ \sum_{k=1}^{K} \lambda_k^o y_j^k \geq y_j^o \qquad j = 1, \dots, J \\
&\qquad \sum_{k \neq h} \lambda_k^o x_i^k + \lambda_h^o \hat{x}_i^h \leq E_h^o x_i^o \quad i = 1, \dots, I \\
&\qquad \lambda_k^o \geq 0 \qquad\qquad k = 1, \dots k, \dots, K, \ \ k \neq h \\
&\qquad \lambda_h^o \ free \\
&\qquad E_h^o \ free
\end{aligned}
\tag{5.9}
$$

where $\hat{x}_i^h = E_h^h x_i^h$. This is, to the best of our knowledge, the first time that cross efficiency scores can be obtained from the envelopment form.

Using the envelopment instead of the multiplier form to obtain the TB cross efficiency scores allows to simplify the process of estimation. This is made possible by noticing that the columns of the cross efficiency matrix (namely, the peer appraisal scores of all DMUs when appraised by a particular DMU) are the same when DEA-inefficient DMUs, chosen as the MPS, are projected on the same part of the strongly DEA efficient frontier and thus, have the same peers. In this case, the cross efficiency scores may be obtained by estimating (5.9) for only one among those DMUs. This means that obtaining the TB cross efficiency matrix through VEA involves estimating *at most* the same number of linear programs as those required by the TB formulation, but using the envelopment form. This provides the necessary shortcuts, hoped for by Doyle and Green (1995), in estimating the TB cross efficiency matrix.

5.4. An algorithm for estimating the TB cross efficiency scores

The proposed algorithm for estimating the TB cross efficiency scores by means of VEA consists of the following steps:

Step 1: Estimate the DEA model in (5.2) and then classify the DMUs into three mutually exclusive groups: group $\mathbb{A}$ containing the DEA-efficient DMUs, group $\mathbb{B}$ containing the DEA-inefficient DMUs that are projected on a part of the strongly efficient DEA frontier, and group $\mathbb{C}$ containing the DEA-inefficient DMUs that are projected on a part of the weakly efficient DEA frontier. Let $\mathcal{A}, \mathcal{B}$ and $\mathcal{C}$ denote the number of DMUs in each of these groups.

Step 2: Using each of the DMUs in group $\mathbb{A}$ as the MPS, estimate the TB cross efficiency scores of the remaining DMUs by using (5.8).

Step 3: Obtain the set of peers for each of the DMUs in group $\mathbb{B}$. Let there be a number of $\mathcal{F} \leq \mathcal{B}$ different sets of peers $S_f, (f = 1, \ldots, \mathcal{F})$ each of which corresponds to possibly several DMUs. Let $\mathbb{B}_f$ $(f = 1, \ldots, \mathcal{F})$ be the subset of $\mathbb{B}$ containing the DMUs having S_f as their set of peers, where $\bigcup_{f=1}^{\mathcal{F}} \mathbb{B}_f = \mathbb{B}$. Estimate (5.7) using the DMUs in each of the sets S_f as the MPS. The resulting scores are the TB cross efficiency scores when each of the DMUs in group $\mathbb{B}_f$ is the "reference" DMU.

Step 4: For each DMU in group $\mathbb{C}$, obtain its radial projection on the DEA frontier, namely $\hat{h} = (\hat{x}_i^h, \hat{y}_j^h) = (E_h^h x_i^h, y_j^h)$. Then, the TB cross efficiency scores of the remaining DMUs when the h^{th} DMU is the "reference" DMU are obtained by (5.9).

The number of linear programs needed to obtain the TB cross efficiency matrix through this algorithm is equal to $(\mathcal{A} + \mathcal{F} + \mathcal{C}) * (K - 1))$. If $\mathcal{F} < \mathcal{B}$, namely when there are at least two inefficient DMUs that are projected on the same part of the strongly efficient DEA frontier, then this number is smaller than $(K * (K - 1))$. If, on the other hand, each DMU in class $\mathbb{B}$ is projected on a different part of the strongly efficient DEA frontier, i.e., $\mathcal{F} = \mathcal{B}$, or if there are no DMUs in class $\mathbb{B}$, i.e., $\mathcal{B} = 0$, then the same number of linear programs is needed to obtain the TB cross efficiency matrix by means of VEA and the TB formulation itself.

From the above, it should be clear that the "shortcuts" for obtaining the TB cross efficiency matrix via VEA depend on (i) the number of DMUs in group $\mathbb{B}$, (ii) the difference between this number and the number of different sets of peers for these DMUs, and (iii) the difference between the number of inputs and outputs and the number of evaluated DMUs. More specifically, the larger is the number of DMUs in group $\mathbb{B}$, the more likely is that some of them may share the same set of peers. On the other hand, the larger the difference between the number of DMUs in group $\mathbb{B}$ and the number of different sets of peers identified for them, the smaller the number of linear programs needed to obtain the TB cross efficiency matrix by means of VEA, compared to that of the TB formulation.

5.5. Empirical implementation and discussion

To demonstrate the simplifications in estimating the TB cross efficiency matrix by means of VEA, we use seven datasets referring to respectively nursing homes (Sexton *et al.*, 1986), university departments (Wong and Beasly, 1990), R&D programs (Oral *et al.* (1991), manufacturing systems (Shang and Sueyoshi, 1995; Baker and Talluri, 1997), university faculty members (Oral *et al.*, 2014), and cotton farms (see the fourth chapter in this Thesis). The number of DMUs (K), of inputs (I), and of outputs (J) involved in each dataset are given in columns (2) to (4) of Table 5.5. The number of DMUs in each of the groups $\mathbb{A}$, $\mathbb{B}$, and $\mathbb{C}$ is given in columns (5) to (7), while the number of different sets of peers for the DMUs in group $\mathbb{B}$ is given in column (8). The number of linear programs needed to obtain the TB cross efficiency matrix by means of (5.6) is given in column (9), while the respective figures when the TB cross efficiency scores are obtained by means of VEA, following the estimation steps outlined above, are given in column (10).

From these empirical results we see that, in four of the cases when assuming CRS and in five of them when assuming VRS, there are no DMUs in group $\mathbb{B}$, and thus the number of linear programs needed to obtain the cross efficiency matrix by means of VEA and the TB formulation is the same. For example, in the case of the 37 R&D projects considered in Oral *et al.* (1991), there are two DEA-efficient DMUs and 35 DEA-inefficient DMUs with CRS, each of which is associated with slacks, while in

Table 5.5: Number of linear models estimated to obtain the TB cross efficiency matrix

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
					DMUs in group				Number of linear programs	
	study	DMUs (K)	inputs (I)	outputs (J)	$\mathbb{A}$	$\mathbb{B}$	$\mathbb{C}$	sets of peers in group $\mathbb{B}$	TB model	VEA model
	constant returns to scale									
1	Sexton et a*l*. (1986)	6	2	2	4	2	0	1	30	25
2	Wong and Beasly (1990)	7	3	3	6	0	1	-	42	42
3	Oral *et al*. (1991)	37	1	5	2	0	35	-	1,332	1,332
4	Shang and Sueyoshi (1995)	12	2	4	7	0	5	-	132	132
5	Baker and Talluri (1997)	27	2	2	9	14	4	5	702	468
6	Oral *et al*. (2014)	32	1	5	11	0	21	-	992	992
7	Chapter 4, this Thesis	526	4	1	12	209	305	13	276,150	173,250
	variable returns to scale									
1	Sexton et a*l*. (1986)	6	2	2	5	0	1	-	30	30
2	Wong and Beasly (1990)	7	3	3	6	0	1	-	42	42
3	Oral *et al*. (1991)	37	1	5	12	0	25	-	1,332	1,332
4	Shang and Sueyoshi (1995)	12	2	4	10	0	2	-	132	132
5	Baker and Talluri (1997)	27	2	2	11	9	7	8	702	676
6	Oral *et al*. (2014)	32	1	5	16	0	16	-	992	992
7	Chapter 4, this Thesis	526	4	1	33	89	404	28	276,150	244,125

VRS the respective figures are equal to 12 and 25. In the remaining three datasets, we have found DMUs in group $\mathbb{B}$ and thus the number of different sets of peers is smaller than the number of DMUs. In these cases, the number of times the VEA model needs to be estimated to obtain the cross efficiency matrix is less than that for the TB model. For example, in the case of the 526 cotton farms considered in the fourth chapter of this Thesis, obtaining the cross efficiency matrix with CRS by means of VEA involves the estimation of 173,250 linear programs compared to 276,150 in the TB model.

The number of DMUs in group $\mathbb{B}$ is indirectly affected by the modelling choices related to the number of inputs and outputs and the nature of returns to scale, as these affect the number of efficient DMUs in DEA. Other things being equal, increasing the number of inputs and outputs naturally increases the number of efficient DMUs, and thus the number of DMUs in groups $\mathbb{B}$ and $\mathbb{C}$ is reduced, while the same occurs when VRS is assumed instead of CRS. In the case of Baker and Talluri (1997) and regardless of the returns-to-scale assumption, there are DMUs in group $\mathbb{B}$, while this does not occur for any DMU in Oral *et al.* (2014). These two studies have roughly the same number of DMUs, but the number of inputs and outputs in the latter is six compared to four in the former. On the other hand, when VRS is assumed rather than CRS, the only inefficient DMU in Sexton *et al.* (1986) is in group $\mathbb{C}$. Thus, the number of VRS linear programs needed to obtain the cross efficiency matrix by means of VEA and the TB formulation is the same, in contrast to when CRS is assumed. In addition, in the case of Baker and Talluri (1997), the number of DMUs in group $\mathbb{B}$ are reduced to eight in VRS compared to 14 when CRS is assumed, and thus more VRS VEA linear programs are needed to obtain the cross efficiency matrix compared to those in CRS (676 vs. 368). Nevertheless, these figures are smaller than those of the TB model irrespective of the returns to scale.

On the other hand, increasing the number of DMUs for a given number of inputs and outputs, is expected to increase the DMUs classified as inefficient, at least to a larger extent compared to those that are rendered efficient. This may or may not result in an increase in the number of DMUs that are projected on the same part of the strongly efficient frontier, as this depends on the DMUs' relative position in the input-output space and cannot be explicitly related to a modelling choice. For instance, the number of inputs and outputs is four in both Baker and Talluri (1997) and Sexton *et al.* (1986) (see Table 5.5). In the former, in which the number of DMUs is more than four times larger compared to the latter, the DMUs in group $\mathbb{B}$ are much more when CRS is assumed. On the other hand, in the cases of Wong and Beasly (1990) and Oral *et al.* (1991), the number of inputs and outputs is also the same, i.e., six, but the DMUs included in the latter are roughly five times more compared to the former. Nevertheless, in both cases and regardless of the returns-to-scale assumption, no DMU is included in group $\mathbb{B}$.

5.6. Concluding remarks

Cross efficiency and VEA are two DEA frameworks used for peer appraisal purposes, which so far have run in parallel lines in the efficiency literature. In this paper, we show that these two lines of literature meet each other for a particular formulation of cross efficiency, namely the TB. In particular, we show that the TB formulation of cross efficiency is equivalent to VEA if either (i) an efficient "reference" DMU, i.e., the one whose optimal multipliers are used to evaluate all other DMUs in the TB formulation, is chosen as the MPS in VEA, or (ii) the radial projection of an inefficient "reference" DMU on the DEA frontier is chosen as the MPS in VEA. The implication of this is that the matrix of the TB cross efficiency scores can be obtained by estimating a sequence of VEA envelopment rather than multiplier form models, as it is common in cross efficiency. This involves estimating at most the same number of linear programs and thus simplifies the problem of obtaining the TB cross-efficiency matrix. Thus, it may prompt a more extensive use of this definitely useful but so far less often employed cross efficiency formulation due to its complicated estimation process. In addition, this equivalence gives rise to alternative interpretations for both the TB cross efficiency scores and the VEA efficiency scores.

The results of this paper can be used in several cases of performance assessment, in which peer appraisal seems appropriate. These include, but are not limited to, faculty member or institution appraisal in higher education, assessment of player performance in sports, evaluation of alternative portfolios and investment projects, but also cases of participatory decision-making such as budget allocation and R&D project selection and group decision-making in international organizations (e.g., EU, NATO, the IMF, or the World Bank) and national government bodies such as parliamentary committees and municipal councils.

CHAPTER 6

On VEA, production trade-offs and weights restrictions

6.1. Introduction

Production trade-offs, their dual weights restrictions, and Value Efficiency Analysis (VEA) are alternative ways of incorporating preference information in Data Envelopment Analysis (DEA). In particular, Decision Maker's (DM) preferences are used to restrict the admissible values of the input/output multipliers in a DEA model. Production trade-offs reflect acceptable marginal changes in inputs and/or outputs that modify their target values for each evaluated Decision Making Unit (DMU) in the envelopment form of DEA models (Podinovski, 2004). Their dual counterpart are the well-known weights restrictions, namely additional linear inequalities in the multiplier form of DEA models that restrict the flexibility of input/output weights based on DM's knowledge, value judgements or in general, holding views for their relative importance (see e.g., Allen *et al*, 1997). DEA models including production trade-offs have been used, among others, for the assessment of efficiency in healthcare (Amado and Dyson, 2009), education (Khalili *et al.*, 2010a), electricity distributors (Santos *et al.*, 2011), and farmers (Atici and Podinovski, 2015). On the other hand, in VEA, the performance of each DMU is assessed relative to the Most Preferred Solution (MPS), namely a non-dominated (i.e., efficient) DMU or a combination of DMUs that has the most desirable input/output structure by view of a DM or reflects DM's preferences about input/output mixes (Halme *et al.*, 1999). In such a case, each DMU's input/output weights are restricted to values among only those that are optimal for the MPS in DEA. This in turn results in extending the DEA efficient facets generated by it. Recent applications of VEA include, but are not limited to, the evaluation of hospital departments (Halme and

Korhonen, 2000), academic institutions (Korhonen *et al.*, 2001), retail stores (Korhonen *et al.,* 2002) as well as bank branches (Halme *et al.,* 2014).

Several studies in the literature have examined the effect of including production trade-offs or their dual weight restrictions in DEA models. For example, Podinovski (2005) demonstrated the effects of using additional restrictions such as weight bounds in the evaluation results of DEA models, Asmild *et al.* (2006) investigated the potential relation of DEA models with trade-offs to models assessing economic (i.e., cost, revenue, or profit) efficiency, and Podinovski (2007a) developed a procedure for obtaining efficient targets in DEA models with production trade-offs. Also, Podinovski and Forsund (2010) assessed the effects of introducing production trade-offs in the indication of returns-to-scale (i.e., whether a DMU exhibits constant, increasing or decreasing returns to scale) and the scale elasticity estimates of DMUs, while Podinovski and Bouzdine-Chameeva (2013) developed linear programs for testing whether the use of a particular set of production trade-offs in DEA models results in violating production assumptions. On the other hand, the similarities between VEA and various forms of weight restrictions have been noted in the literature, but not yet thoroughly examined. For instance, Sarrico and Dyson (2004, p. 18) considered VEA as ‘’another alternative to incorporating the decision maker’s preferences into the assessment of DMUs’’, while Kao and Hung (2005, p. 1197) noted that VEA is ‘’essentially an approach of weight restrictions’’. Angulo-Meza and Estellita-Lins (2002, p. 225) viewed VEA and weights restrictions as methodologies incorporating “information provided by a decision maker or expert into the model”, while Adler *et al.* (2002) referred to VEA as one of the methods that use “preference information to further refine the discriminatory power of DEA models”.[88] Nevertheless, none of these studies have explicitly related VEA to DEA models including weights restrictions, as well as their dual production trade-offs. Such explicit relationships, if any, have, to the best of our knowledge, not yet been investigated.

The purpose of this chapter is to explore the relation between VEA and DEA models including production trade-offs and their dual weights restrictions in a detailed manner. More specifically, we show that, under constant returns to scale, the VEA

---

[88] This is also one of the main reasons motivating the incorporation of weights restrictions.

model is equivalent to a DEA model including production trade-offs, for which the trade-off coefficient vectors are given by the negative of the input and output quantities of the DMUs chosen as the MPS in VEA. We also show that, regardless of the nature of returns to scale, the VEA model is equivalent to a DEA model with production trade-offs, for which the trade-offs coefficient vectors are given by the deviations of every DMU's input and output quantities from those of the MPS. These production trade-offs result in extending certain facets of the DEA frontier and in both cases are dual to Type II Assurance Region (AR-II) weight restrictions (see Thompson *et al.*, 1990). Considering the above trade-offs for only inputs or outputs we can prove a similar equivalence between pure input or output VEA and DEA models. The dual form of these trade-offs, which refer only to inputs or outputs, are type I Assurance Region (AR-I) weight restrictions (Thompson *et al.*, 1986).

The rest of the chapter unfolds as follows: In the second section we discuss VEA and DEA models with production trade-offs. The chapters' main results are presented in the third section, while an empirical application follows in the fourth section. Concluding remarks follow in the last section.

6.2. Materials and methods

Production trade-offs are the dual form of weights restrictions that are usually appended in the multiplier form of DEA models. They refer to marginal changes between inputs and/or outputs that take place at some point at the conventional DEA frontier and enlarge the feasible space with additional input/output possibilities (Podinovski, 2004). These changes represent perceptions regarding the normative substitution rates between inputs or transformation rates between outputs, or simply judgements about the relative importance of different inputs and outputs. They are considered as acceptable by all evaluated DMUs, in the sense that it is unanimously agreed that they result in feasible (technologically possible) input/output combinations. Then, one may argue that the targets identified for inefficient DMUs on the enlarged parts of the DEA frontier are in principle technologically realistic or feasible (Podinovski, 2007b).

Let us consider a set of K DMUs ($k = 1, \dots, o, \dots, K$) using the same technology and producing a set of J ($j = 1, \dots, J$) outputs utilizing I ($i = 1, \dots, I$) inputs. Assume further that there exists a number of R ($r = 1, \dots, R$) trade-off relations among inputs and/or outputs, which may be represented as:

$$P^r = [p_1^r, \dots, p_i^r, \dots, p_I^r]^T, \;\; r = 1, \dots, R$$
$$Q^r = \left[q_1^r, \dots, q_j^r, \dots, q_J^r\right]^T, \;\; r = 1, \dots, R \tag{6.1}$$

Each of the trade-offs in (6.1) refers to an agreed postulate among DMUs that by changing the level of each of a DMU's inputs by the trade-off coefficient p_i^r and each of its outputs by the trade-off coefficient q_j^r results in a new unobserved input/output combination that is feasible. Thus, the vectors P^r and Q^r modify respectively the target values of inputs and outputs in the envelopment form of a DEA model, which in turn results in enlarging the DEA efficient frontier with additional linear segments, i.e., facets.

The multiplier and envelopment form of an input-oriented, constant returns to scale (CRS) DEA model including trade-offs as in (6.1) is given as (Podinovski, 2004):[89]

$$\begin{array}{llll}
\max\limits_{u_j^o, v_i^o} & \sum\limits_{j=1}^{J} u_j^o y_j^o & & \\
s.t. & \sum\limits_{j=1}^{J} u_j^o y_j^k - \sum\limits_{i=1}^{I} v_i^o x_i^k \le 0 & k = 1, \dots, K \\
& \sum\limits_{j=1}^{J} u_j^o q_j^r - \sum\limits_{i=1}^{I} v_i^o p_i^r \le 0 & r = 1, \dots, R \\
& \sum\limits_{i=1}^{I} v_i^o x_i^o = 1 & \\
& u_j^o \ge 0 & j = 1, \dots, J \\
& v_i^o \ge 0 & i = 1, \dots, I
\end{array}
\qquad
\begin{array}{lll}
\min\limits_{\theta_{TO}^o, \lambda_k^o} \theta_{TO}^o & & \\
s.t. \; \sum\limits_{k=1}^{K} \lambda_k^o y_j^k + \sum\limits_{r=1}^{R} \pi_r^o q_j^r \ge y_j^o & j = 1, \dots, J \\
\sum\limits_{h=1}^{K} \lambda_k^o x_i^k + \sum\limits_{r=1}^{R} \pi_r^o p_i^r \le \theta_{TO}^o x_i^o & i = 1, \dots, I \\
\lambda_k^o \ge 0 & k = 1, \dots, K \\
\pi_r^o \ge 0 & r = 1, \dots, R \\
\theta_{TO}^o \; free &
\end{array} \tag{6.2}$$

where *x* and *y* are respectively the quantities of inputs and outputs, *v* and *u* are their input and output weights, θ is the efficiency score, λ are the intensity variables, and π are the proportions by which each of the trade-offs is applied to modify the input and output targets. On the other hand, the variable returns to scale (VRS) counterpart of (6.2) is given as:

[89] We focus on the input-oriented model, but our results can be straightforwardly extended to the output-oriented model.

$$
\begin{aligned}
&\max_{u_j^o, v_i^o, u^o} \sum_{j=1}^{J} u_j^o y_j^o + u^o \\
&s.t. \quad \sum_{j=1}^{J} u_j^o y_j^k - \sum_{i=1}^{I} v_i^o x_i^k + u^o \le 0 \quad && k = 1, \dots, K \\
&\sum_{j=1}^{J} u_j^o q_j^r - \sum_{i=1}^{I} v_i^o p_i^r \le 0 && r = 1, \dots, R \\
&\sum_{i=1}^{I} v_i^o x_i^o = 1 \\
&u_j^o \ge 0 && j = 1, \dots, J \\
&v_i^o \ge 0 && i = 1, \dots, I \\
&u^o \ free
\end{aligned}
\qquad
\begin{aligned}
&\min_{\theta_{TO}^o, \lambda_k^o} \theta_{TO}^o \\
&s.t. \quad \sum_{k=1}^{K} \lambda_k^o y_j^k + \sum_{r=1}^{R} \pi_r^o q_j^r \ge y_j^o \quad && j = 1, \dots, J \\
&\sum_{k=1}^{K} \lambda_k^o x_i^k + \sum_{r=1}^{R} \pi_r^o p_i^r \le \theta_{TO}^o x_i^o && i = 1, \dots, I \\
&\sum_{k=1}^{K} \lambda_k^o = 1 \\
&\lambda_k^o \ge 0 && k = 1, \dots, K \\
&\pi_r^o \ge 0 && r = 1, \dots, R \\
&\theta_{TO}^o \ free
\end{aligned}
\tag{6.3}
$$

in which the free variable u^o is dual to the convexity constraint in the envelopment form of (6.3).

From (6.2) and (6.3) we can see that incorporation of trade-offs such as in (6.1) into the envelopment form of the DEA model is equivalent to imposing the following set of homogeneous weight restrictions in its multiplier form:[90]

$$
\sum_{j=1}^{J} u_j^o q_j^r - \sum_{i=1}^{I} v_i^o p_i^r \le 0, \quad r = 1, \dots, R \tag{6.4}
$$

The weight restrictions in (6.4) concern value judgments regarding (i) only inputs, if $q_j^r = 0$ for $j = 1, \dots, J$, (ii) only outputs, if $p_i^r = 0$ for $i = 1, \dots, I$, or (iii) both inputs and outputs, if $q_j^r \neq 0$ for at least one $j = 1, \dots, J$ and $p_i^r \neq 0$ for at least one $i = 1, \dots, I$.[91] In the former two cases they are referred to as AR-I (Thompson *et al.*, 1986), while in the latter case as AR-II weight restrictions (Thompson *et al.*, 1990).

On the other hand, in VEA, a DM expresses his/her preferences over the desirable input/output structure or mix by choosing a DMU or a combination of DMUs as the MPS (Halme *et al.*, 1999). This might be a more appealing way of expressing preferences, as DMs are usually more keen to choose desirable values for the inputs and outputs rather that weight bounds (Korhonen *et al.*, 2002). The VEA frontier is

---

[90] Podinovski (2004) has shown that non-homogeneous linear weight restrictions, i.e., those for which the right-hand side of (6.4) is non-zero, can also be represented in the form of production trade-offs in the envelopment form of the DEA model

[91] Cases (i) and (ii) may also refer to a subset of inputs and outputs if respectively $p_i^r = 0$ for some *i* in Case (i) and $q_j^r = 0$ for some *j* in Case (ii).

then constructed as the lower envelope of the extended DEA efficient facets intercepting at the MPS. As the facets of the DEA efficient frontier are generated by extreme-efficient DMUs, the MPS will in essence be either a single extreme-efficient DMU or a combination of extreme-efficient DMUs that are *jointly* efficient, in the sence that they generate at least one common facet. VEA then extends only these common facets among the DMUs comprising the MPS.

The input-oriented CRS VEA model in its multiplier and envelopment form is given as:

$$
\begin{array}{llll}
\max\limits_{u_j^o, v_i^o} & \sum\limits_{j=1}^{J} u_j^o y_j^o & & \\
s.t. & \sum\limits_{j=1}^{J} u_j^o y_j^k - \sum\limits_{i=1}^{I} v_i^o x_i^k \leq 0 & k = 1, \dots, K, \;\; k \neq r \\
& \sum\limits_{j=1}^{J} u_j^o y_j^k - \sum\limits_{i=1}^{I} v_i^o x_i^k = 0 & k = r = 1, \dots, R \\
& \sum\limits_{i=1}^{I} v_i^o x_i^o = 1 & \\
& u_j^o \geq 0 & j = 1, \dots, J \\
& v_i^o \geq 0 & i = 1, \dots, I
\end{array}
\qquad
\begin{array}{lll}
\min\limits_{\theta_{VEA}^o, \lambda_k^o} & \theta_{VEA}^o & \\
s.t. & \sum\limits_{k=1}^{K} \lambda_k^o y_j^k \geq y_j^o & j = 1, \dots, J \\
& \sum\limits_{k=1}^{K} \lambda_k^o x_i^k \leq \theta_{VEA}^o x_i^o & i = 1, \dots, I \\
& \lambda_k^o \geq 0 & k = 1, \dots, K, \;\; k \neq r \\
& \lambda_k^o \;\; free & k = r = 1, \dots, R \\
& \theta_{VEA}^o \, free &
\end{array}
\tag{6.5}
$$

where the set $\mathcal{R}$ $(r = 1, \dots, R)$ contains the DMUs comprising the MPS. On the other hand, the input-oriented VRS VEA model in its multiplier and envelopment form is given as:

$$
\begin{array}{lll}
\max\limits_{u_j^o, v_i^o, u^o} & \sum\limits_{j=1}^{J} u_j^o y_j^o + u^o & \\
s.t. & \sum\limits_{j=1}^{J} u_j^o y_j^k - \sum\limits_{i=1}^{I} v_i^o x_i^k + u^o \leq 0 & k = 1, \dots, K, \;\; k \neq r \\
& \sum\limits_{j=1}^{J} u_j^o y_j^k - \sum\limits_{i=1}^{I} v_i^o x_i^k + u^o = 0 & k = r = 1, \dots, R \\
& \sum\limits_{i=1}^{I} v_i^o x_i^o = 1 & \\
& u_j^o \geq 0 & j = 1, \dots, J \\
& v_i^o \geq 0 & i = 1, \dots, I \\
& u^o \, free &
\end{array}
\qquad
\begin{array}{lll}
\min\limits_{\theta_{VEA}^o, \lambda_k^o} & \theta_{VEA}^o & \\
s.t. & \sum\limits_{k=1}^{K} \lambda_k^o y_j^k \geq y_j^o & j = 1, \dots, J \\
& \sum\limits_{k=1}^{K} \lambda_k^o x_i^k \leq \theta_{VEA}^o x_i^o & i = 1, \dots, I \\
& \sum\limits_{k=1}^{K} \lambda_k^o = 1 & \\
& \lambda_k^o \geq 0 & k = 1, \dots, K, \;\; k \neq r \\
& \lambda_k^o \;\; free & k = r = 1, \dots, R \\
& \theta_{VEA}^o \, free &
\end{array}
\tag{6.6}
$$

where the free variable u^o and the convexity constraint for the intensity variables are added in the multiplier and envelopment form, respectively. In the envelopment form of (6.5) and (6.6), the non-negativity restrictions are removed from the intensity variables of the DMUs comprising the MPS (Halme *et al.*, 1999). This in turn implies that the inequalities referring to these DMUs should change into strict equalities in the multiplier form of the model, essentially restricting each evaluated DMU to choose input and output weights only among those that are optimal (in the conventional DEA model) for the DMU or the combination of DMUs chosen as the MPS.

6.3. Main results

6.3.1. Production trade-offs dual to AR-II type of weight restrictions

To relate the VEA models in (6.5) and (6.6) to the DEA models with production-trade-offs and their dual weight restrictions in (6.2) and (6.3), notice that each of the side equality restrictions in (6.5) and (6.6) can be broken up into the following equivalent pair of inequalities: $\sum_{j=1}^{J} u_j^o y_j^k - \sum_{i=1}^{I} v_i^o x_i^k \leq 0$ and $\sum_{j=1}^{J} u_j^o y_j^k - \sum_{i=1}^{I} v_i^o x_i^k \geq 0$ for (6.5) and $\sum_{j=1}^{J} u_j^o y_j^k - \sum_{i=1}^{I} v_i^o x_i^k + u^o \leq 0$ and $\sum_{j=1}^{J} u_j^o y_j^k - \sum_{i=1}^{I} v_i^o x_i^k + u^o \geq 0$ for (6.6). Based on these, (6.5) and (6.6) may be rewritten as:

$$
\begin{array}{lll}
\max\limits_{u_j^o, v_i^o} & \sum\limits_{j=1}^{J} u_j^o y_j^o & \\
s.t. & \sum\limits_{j=1}^{J} u_j^o y_j^k - \sum\limits_{i=1}^{I} v_i^o x_i^k \leq 0 & k = 1, \dots, K \\
& \sum\limits_{j=1}^{J} u_j^o (-y_j^r) - \sum\limits_{i=1}^{I} v_i^o (-x_i^r) \leq 0 & r = 1, \dots, R \\
& \sum\limits_{i=1}^{I} v_i^o x_i^o = 1 & \\
& u_j^o \geq 0 & j = 1, \dots, J \\
& v_i^o \geq 0 & i = 1, \dots, I
\end{array}
\qquad
\begin{array}{lll}
\min\limits_{\theta_{VEA}^o, \lambda_k^o} & \theta_{VEA}^o & \\
s.t. & \sum\limits_{k=1}^{K} \lambda_k^o y_j^k + \sum\limits_{r=1}^{R} \pi_r^o (-y_j^r) \geq y_j^o & j = 1, \dots, J \\
& \sum\limits_{k=1}^{K} \lambda_k^o x_i^k + \sum\limits_{r=1}^{R} \pi_r^o (-x_i^r) \leq \theta_{VEA}^o x_i^o & i = 1, \dots, I \\
& \lambda_k^o \geq 0 & k = 1, \dots, K \\
& \pi_r^o \geq 0 & r = 1, \dots, R \\
& \theta_{VEA}^o \ free &
\end{array}
\tag{6.7}
$$

and as:

$$\begin{aligned}
&\max_{u_j^o, v_i^o, u^o} \sum_{j=1}^{J} u_j^o y_j^o + u^o \\
&s.t.\ \sum_{j=1}^{J} u_j^o y_j^k - \sum_{i=1}^{I} v_i^o x_i^k + u^o \le 0 && k = 1, \dots, K \\
&\sum_{j=1}^{J} u_j^o(-y_j^r) - \sum_{i=1}^{I} v_i^o(-x_i^r) - u^o \le 0 && r = 1, \dots, R \\
&\sum_{i=1}^{I} v_i^o x_i^o = 1 \\
&u_j^o \ge 0 && j = 1, \dots, J \\
&v_i^o \ge 0 && i = 1, \dots, I \\
&u^o\ free
\end{aligned}
\qquad
\begin{aligned}
&\min_{\theta_{VEA}^o, \lambda_k^o} \theta_{VEA}^o \\
&s.t.\ \sum_{k=1}^{K} \lambda_k^o y_j^k + \sum_{r=1}^{R} \pi_r^o(-y_j^r) \ge y_j^o && j = 1, \dots, J \\
&\sum_{k=1}^{K} \lambda_k^o x_i^k + \sum_{r=1}^{R} \pi_r^o(-x_i^r) \ge \theta_{VEA}^o x_i^o && i = 1, \dots, I \\
&\sum_{k=1}^{K} \lambda_k^o - \sum_{r=1}^{R} \pi_r^o = 1 \\
&\lambda_k^o \ge 0 && k = 1, \dots, K \\
&\pi_r^o \ge 0 && r = 1, \dots, R \\
&\theta_{VEA}\ free
\end{aligned}
\tag{6.8}$$

Then one can easily verify that (6.5) and (6.7) and (6.6) and (6.8) are equivalent to each other.

The additional restrictions in the multiplier form of (6.7) and (6.8) result in new terms in the left-hand sides of the inequalities in the envelopment form. These terms contain the negative of the input and output quantities of the DMUs constituting the MPS. We may thus view (6.7) and (6.8) as models with $(K + R)$ DMUs, where inputs and outputs take negative values for the DMUs in the set of the MPS (the R additional ones) and positive values for the sample DMUs. By using Emrouznejad *et al.* (2010) data transformations, (6.7) and (6.8) may be seen as semi-oriented DEA models with $(K + R)$ DMUs. Specifically, we may redefine the input and output variables in (6.7) and (6.8) as:

$$x_{1i}^k = \begin{cases} x_i^k, & k = 1, \dots, K \\ 0, & r = 1, \dots, R \end{cases} \quad and\ x_{2i}^k = \begin{cases} 0, & k = 1, \dots, K \\ -x_i^r, & r = 1, \dots, R \end{cases} \quad i = 1, .., I$$

and:

$$y_{1j}^k = \begin{cases} y_j^k, & k = 1, \dots, K \\ 0, & r = 1, \dots, R \end{cases} \quad and\ y_{2j}^k = \begin{cases} 0, & k = 1, \dots, K \\ -y_j^r, & r = 1, \dots, R \end{cases} \quad j = 1, .., J$$

Then, (6.7) and (6.8) are respectively be written as:

$$\begin{aligned}
&\max_{u_j^o, v_i^o, u^o} \sum_{j=1}^{J} u_{1j}^o y_{1j}^o - \sum_{j=1}^{J} u_{2j}^o y_{2j}^o \\
&s.t. \ \sum_{j=1}^{J} u_{1j}^o y_{1j}^k - \sum_{j=1}^{J} u_{2j}^o y_{2j}^k - \\
&\quad -\sum_{i=1}^{I} v_{1i}^o x_{1i}^k + \sum_{i=1}^{I} v_{2i}^o x_{2i}^k \le 0 \qquad k = 1, \dots, K+R \\
&\quad \sum_{i=1}^{I} v_{1i}^o x_{1i}^k - \sum_{i=1}^{I} v_{2i}^o x_{2i}^k = 1 \\
&\quad u_{1j}^o \ge 0 \qquad j = 1, \dots, J \\
&\quad u_{2j}^o \ge 0 \qquad j = 1, \dots, J \\
&\quad v_{2i}^o \ge 0 \qquad i = 1, \dots, I \\
&\quad v_{2i}^o \ge 0 \qquad i = 1, \dots, I
\end{aligned}
\qquad
\begin{aligned}
&\min_{\theta_{VEA}^o, \lambda_k^o} \theta_{VEA}^o \\
&s.t. \ \sum_{k=1}^{K+R} \lambda_k^o y_{1j}^k \ge y_{1j}^o \qquad j = 1, \dots, J \\
&\quad \sum_{k=1}^{K+R} \lambda_k^o y_{2j}^k \le y_{2j}^o \qquad j = 1, \dots, J \\
&\quad \sum_{k=1}^{K+R} \lambda_k^o x_{1i}^k \le \theta_{VEA}^o x_{1i}^o \quad i = 1, \dots, I \\
&\quad \sum_{k=1}^{K+R} \lambda_k^o x_{2i}^k \ge \theta_{VEA}^o x_{2i}^o \quad i = 1, \dots, I \\
&\quad \lambda_k^o \ge 0 \qquad k = 1, \dots, K+R \\
&\quad \theta_{VEA}^o \ free
\end{aligned}
\tag{6.9}$$

and

$$\begin{aligned}
&\max_{u_j^o, v_i^o, u^o} \sum_{j=1}^{J} u_{1j}^o y_{1j}^o - \sum_{j=1}^{J} u_{2j}^o y_{2j}^o + u^o \\
&s.t. \ \sum_{j=1}^{J} u_{1j}^o y_{1j}^k - \sum_{j=1}^{J} u_{2j}^o y_{2j}^k - \\
&\quad -\sum_{i=1}^{I} v_{1i}^o x_{1i}^k + \sum_{i=1}^{I} v_{2i}^o x_{2i}^k + u^o \le 0 \qquad k = 1, \dots, K+R \\
&\quad \sum_{i=1}^{I} v_{1i}^o x_{1i}^k - \sum_{i=1}^{I} v_{2i}^o x_{2i}^k = 1 \\
&\quad u_{1j}^o \ge 0 \qquad j = 1, \dots, J \\
&\quad u_{2j}^o \ge 0 \qquad j = 1, \dots, J \\
&\quad v_{2i}^o \ge 0 \qquad i = 1, \dots, I \\
&\quad v_{2i}^o \ge 0 \qquad i = 1, \dots, I \\
&\quad u^o \ free
\end{aligned}
\qquad
\begin{aligned}
&\min_{\theta_{VEA}^o, \lambda_k^o} \theta_{VEA}^o \\
&s.t. \ \sum_{k=1}^{K+R} \lambda_k^o y_{1j}^k \ge y_{1j}^o \qquad j = 1, \dots, J \\
&\quad \sum_{k=1}^{K+R} \lambda_k^o y_{2j}^k \le y_{2j}^o \qquad j = 1, \dots, J \\
&\quad \sum_{k=1}^{K+R} \lambda_k^o x_{1i}^k \le \theta_{VEA}^o x_{1i}^o \quad i = 1, \dots, I \\
&\quad \sum_{k=1}^{K+R} \lambda_k^o x_{2i}^k \ge \theta_{VEA}^o x_{2i}^o \quad i = 1, \dots, I \\
&\quad \sum_{k=1}^{K+R} \lambda_k^o = 1 \\
&\quad \lambda_k^o \ge 0 \qquad k = 1, \dots, K+R \\
&\quad \theta_{VEA}^o \ free
\end{aligned}
\tag{6.10}$$

where k is used to index all DMUs, i.e., $k = 1, \dots, (K + R)$.

We can now provide sufficient conditions under which the DEA model including production trade-offs or their dual weight restrictions is equivalent to the VEA model. Under CRS, a comparison of (6.2) and (6.7) shows that the two models are equivalent to each other if the number of trade-offs in the former is equal to the number of DMUs constituting the MPS in the latter and the trade-off coefficient vectors are given as:

$$P^r = [-x_1^r, \dots, -x_i^r, \dots, -x_I^r]^T, \;\; r = 1, \dots, R$$
$$Q^r = \left[-y_1^r, \dots, -y_j^r, \dots, -y_J^r\right]^T, \;\; r = 1, \dots, R \tag{6.11}$$

where $\left(x_i^r, y_j^r\right)$ correspond to the inputs and outputs of each of the DMUs ($r=1,\dots,R$) constituting the MPS. The trade-offs in (6.11) are dual to the following set of AR-II type weight restrictions:

$$\sum_{j=1}^{J} u_j^o\left(-y_j^r\right) - \sum_{i=1}^{I} v_i^o\left(-x_i^r\right) \le 0, \;\; r = 1, \dots, R \tag{6.12}$$

which are essentially the same as the second set of restrictions in the multiplier form of (6.7). Thus, we have:

Proposition 6.1: Under constant returns to scale, the VEA model is equivalent to a DEA model including production trade-offs, for which the trade-off coefficient vectors contain the negative of the input and output quantities of the DMUs constituting the MPS in VEA.

When VRS is assumed, substituting (6.11) or its dual (6.12) into (6.3) will not result in a model equivalent to (6.8), since the convexity constraints in the envelopment form of (6.3) and (6.8) are different from each other.

However, we can show that the VEA model is related to the DEA model including another form of trade-offs:

Proposition 6.2: Regardless of the nature of the returns to scale, the VEA model is equivalent to a DEA model including production trade-offs, for which the trade-off coefficient vectors contain the deviations of each DMU's input and output quantities from those of each of the DMUs constituting the MPS.

To show this, consider the following trade-offs:

$$P_r^k = \left[\left(x_1^k - x_1^r\right), \dots, \left(x_I^k - x_I^r\right)\right]^T, \;\; k = 1, \dots, K, \;\; r = 1, \dots, R$$
$$Q_r^k = \left[\left(y_1^k - y_1^r\right), \dots, \left(y_J^k - y_J^r\right)\right]^T, \;\; k = 1, \dots, K, \;\; r = 1, \dots, R \tag{6.13}$$

which are dual to the following set of AR-II type of weight restrictions:

$$\sum_{j=1}^{J} u_j^o\left(y_j^k - y_j^r\right) - \sum_{i=1}^{I} v_i^o\left(x_i^k - x_i^r\right) \leq 0, \;\; k = 1, \dots, K, \;\; r = 1, \dots, R \qquad (6.14)$$

Let's assume, initially, that $r = 1$, i.e., the MPS is a single DMU. Then, (6.13) consists of K trade-off coefficient vectors given as the deviations of each DMU's *(k=1,…,K)* input and output quantities of from those of the MPS. That is, $p_i^k = \left(x_i^k - x_i^r\right),\; i = 1, \dots, I,\; k = 1, \dots, K$ and $q_j^k = \left(y_j^k - y_j^r\right),\; j = 1, \dots, J,\; k = 1, \dots, K$. In such a case, the envelopment form of the VRS DEA model in (6.3) is given as:

$$\begin{aligned}
&\min_{\theta_{TO}^o, \lambda_k^o, \pi_k^o} \;\; \theta_{TO}^o \\
&s.t. \;\; \sum_{k=1}^{K} \lambda_k^o y_j^k + \sum_{k=1}^{K} \pi_k^o\left(y_j^k - y_j^r\right) \geq y_j^o \qquad j = 1, \dots, J \\
&\qquad \sum_{k=1}^{K} \lambda_k^o x_i^k + \sum_{k=1}^{K} \pi_k^o\left(x_i^k - x_i^r\right) \leq \theta_{TO}^o x_i^o \quad i = 1, \dots, I \\
&\qquad \sum_{k=1}^{K} \lambda_k^o = 1 \\
&\qquad \lambda_k^o \geq 0 \qquad\qquad\qquad\qquad\qquad k = 1, \dots, K \\
&\qquad \pi_k^o \;\geq 0 \qquad\qquad\qquad\qquad\qquad k = 1, \dots, K \\
&\qquad \theta_{TO}^o \;\;\; free
\end{aligned} \qquad (6.15)$$

or equivalently as:

$$\begin{aligned}
&\min_{\theta_{TO}^o, \delta_k^o, \gamma^o} \;\; \theta_{TO}^o \\
&s.t. \;\; \sum_{k=1}^{K} \delta_k^o y_j^k + \gamma^o\left(-y_j^r\right) \geq y_j^o \qquad j = 1, \dots, J \\
&\qquad \sum_{k=1}^{K} \delta_k^o x_i^k + \gamma^o(-x_i^r) \leq \theta_{TO}^o x_i^o \quad i = 1, \dots, I \\
&\qquad \sum_{k=1}^{K} \delta_k^o - \gamma^o = 1 \\
&\qquad \delta_k^o \geq 0 \qquad\qquad\qquad\qquad k = 1, \dots, K \\
&\qquad \gamma^o \geq 0 \\
&\qquad \theta_{TO}^o \;\;\; free
\end{aligned} \qquad (6.16)$$

where $\delta_k^o = (\lambda_k^o + \pi_k^o) \geq 0$ and $\sum_{k=1}^{K} \pi_k^o = \gamma^o \geq 0$. Then (6.16) is equivalent to the envelopment form in (6.8) if the *r*th DMU is chosen as the MPS. If $r > 1$, namely that the MPS is a combination of several DMUs, then (6.13) consists of $K \times R$ trade-off coefficient vectors given as the deviations of each DMU's *(k=1,…,K)* input and output quantities from those of each DMU *(r=1,…,R)* comprising the MPS. As a result, the second term in the left hand side of the first two inequality restrictions in (6.15) reflect summations over both *k* (*k=1,…,K)* and *r (r=1,…,R)* and π_k^o should be changed to π_{kr}^o. Moreover, by defining $\delta_k^o = (\lambda_k^o + \sum_{r=1}^R \pi_{kr}^o) \geq 0$ and $\gamma_r^o = \sum_{k=1}^K \pi_k^o \geq 0$ we may obtain a model similar to (16) in which the second term in the left hand side of the first two inequality restrictions reflect summations over *r (r=1,…,R)* and the third restriction is stated as $\sum_{k=1}^K \delta_k^o - \sum_{r=1}^R \gamma_r^o = 1$. This model is equivalent to the envelopment form in (6.8) if the set of *R (r=1,…,R)* DMUs comprise the MPS. In a similar fashion, if increasing (decreasing) returns to scale are assumed, then the equality sign in the third restriction of the envelopment forms in (6.3) and (6.8) and in (6.15) and (6.16) is simply changed to a less-than-or-equal (greater-than-or-equal) sign, while if constant returns to scale are assumed, the third restriction in the envelopment forms in (6.3) and (6.8) and in (6.15) and (6.16) should be dropped.

From the above, it is also clear that, under CRS, the DEA model with *R* trade-off coefficient vectors, given as the negative of the input and output quantities of the DMUs chosen as the MPS in VEA, is equivalent to the DEA model with $(K \times R)$ trade-off coefficient vectors, given as the deviations of each DMU's (*k=1,…,K*) input and output quantities from those of each of the DMUs chosen as the MPS in VEA. This is evident as long as the trade-off coefficient vectors in (6.2) are given as in either (6.11) or (6.13). Let's assume, initially, that $r = 1$, i.e., the MPS is a single DMU. Then, the envelopment form of (6.2) when the trade-off coefficient vectors are given by (6.13) is:

$$
\begin{aligned}
&\min_{\theta_{TO}^o, \lambda_k^o, \pi_k^o} \ \theta_{TO}^o \\
&s.t. \ \sum_{k=1}^K \lambda_k^o y_j^k + \sum_{k=1}^K \pi_k^o \left(y_j^k - y_j^r\right) \geq y_j^o \qquad j = 1, \dots, J \\
&\qquad \sum_{k=1}^K \lambda_k^o x_i^k + \sum_{k=1}^K \pi_k^o \left(x_i^k - x_i^r\right) \leq \theta_{TO}^o x_i^o \quad i = 1, \dots, I \\
&\qquad \lambda_k^o \geq 0 \qquad k = 1, \dots, K \\
&\qquad \pi_k^o \geq 0 \qquad k = 1, \dots, K \\
&\qquad \theta_{TO}^o \quad free
\end{aligned}
\tag{6.18}
$$

while when the trade-off coefficient vectors are given by (6.11), it is as follows:

$$\min_{\theta_{TO}^o,\zeta_k^o,\gamma^o} \theta_{TO}^o$$
$$\begin{aligned} s.t. \ & \sum_{k=1}^{K} \zeta_k^o y_j^k + \beta^o\left(-y_j^r\right) \geq y_j^o \qquad && j = 1, \dots, J \\ & \sum_{k=1}^{K} \zeta_k^o x_i^k + \beta^o\left(-x_i^r\right) \leq \theta_{TO}^o x_i^o && i = 1, \dots, I \\ & \zeta_k^o \geq 0 && k = 1, \dots, K \\ & \beta^o \ \geq 0 && \\ & \theta_{TO}^o \quad free && \end{aligned} \tag{6.19}$$

If $\zeta_k^o = (\lambda_k^o + \pi_k^o) \geq 0$ and $\sum_{k=1}^{K} \pi_k^o = \beta^o \geq 0$, then (6.18) is equivalent to (6.19). If $r > 1$, then (6.13) consists of $(K \times R)$ trade-off coefficient vectors given as $p_i^{kr} = \left(x_i^k - x_i^r\right),\ i = 1, \dots, I,\ k = 1, \dots, K, r = 1, \dots, R$ and $q_j^{kr} = \left(y_j^k - y_j^r\right),\ j = 1, \dots, J,$ $h = 1, \dots, K, r = 1, \dots, R$. Thus, the second terms in the left hand side of the first two inequality restrictions in (6.18) reflect summations over both *k* (*k=1,…,K)* and *r (r=1,…,R)* and π_k^o should be changed to π_{kr}^o. Furthermore, (6.11) consists of *R* trade-off coefficient vectors given by the negative of the input and output quantities of the DMUs chosen as the MPS in VEA. Thus, the second terms in the left hand side of the first two inequality restrictions in (6.19) reflect summations over *r (r=1,…,R)* and β^o should be changed to β_r^o. Then, by defining $\zeta_k^o = (\lambda_k^o + \sum_{r=1}^{R} \pi_{kr}^o) \geq 0$ and $\beta_r^o = \sum_{k=1}^{K} \pi_k^o \geq 0$, (6.18) is equivalent to (6.19). Consequently, a CRS DEA model augmented with the trade-offs as in (6.11) and a CRS DEA model augmented with the trade-offs as in (6.13) are equivalent to each other.

The above results indicate that, under certain circumstances, the DM preferences underlying the evaluation of DMUs in the VEA model may be seen as a particular form of trade-offs or AR-II type of weight restrictions and vice versa. This provides an alternative interpretation of the efficiency scores obtained from both the VEA model and the DEA model including production trade-offs.

The production trade-offs in (6.11) and (6.13) and their dual weight restrictions in (6.12) and (6.14) enlarge the DEA efficient frontier by extending certain of its existing facets, in particular, those associated with the DMU or the combination of DMUs comprising the MPS, instead of introducing new linear segments.[92] The

[92] Weight restrictions that result in extending facets of the DEA frontier are discussed in Portela and Thanassoulis (2006), but are not related to VEA.

implications of this are: (i) (6.11) and (6.13) do not introduce additional information in the envelopment form of the DEA model other than that already implicit in the data, namely the rates of substitution between inputs, the rates of transformation among outputs, and the marginal products between inputs and outputs that are reflected in each of the extended facets, and (ii) the efficiency scores from the multiplier form of the DEA models with (6.12) and (6.14) do not underestimate the true efficiency of the evaluated DMUs, as may occur in several other cases where additional restrictions of the general form in (6.4) are imposed in DEA models (see Tracy and Chen, 2005; Khalili *et al.*, 2010b). This is because each facet of the DEA frontier enlarged with (6.12) or (6.14) is already tangent to the conventional DEA frontier at some point.

6.3.2. Production trade-offs dual to AR-I type of weight restrictions

In the previous section we considered production trade-offs related to both inputs and outputs, which are dual to AR-II type of weight restrictions. In this section we consider weight restrictions of the AR-I type, and we restrict our attention to pure input or output models, i.e., models that contain respectively no inputs and outputs.

Consider first the DEA model without inputs, which is equivalent to a DEA model with a single or multiple constant (unitary) inputs (Lovell and Pastor, 1999).[93] The latter is known as the Benefit-of-the-Doubt model (BoD) and its multiplier and envelopment form are given as (Cherchye *et al.*, 2007a):

$$
\begin{array}{llll}
\max\limits_{u_j^o, v_i^o} & \sum\limits_{j=1}^{J} u_j^o y_j^o & & \\
s.t. & \sum\limits_{j=1}^{J} u_j^o y_j^k \leq 1 & k = 1, \dots, K & \\
& u_j^o \geq 0 & j = 1, \dots, J &
\end{array}
\qquad
\begin{array}{lll}
\min\limits_{\lambda_k^o} & \sum\limits_{k=1}^{K} \lambda_k^o & \\
s.t. & \sum\limits_{k=1}^{K} \lambda_k^o y_j^k \geq y_j^o & j = 1, \dots, J \\
& \lambda_k^o \geq 0 & k = 1, \dots, K \\
& \theta^o \ free &
\end{array}
\qquad (6.20)
$$

The model in (6.20) is obtained from (6.2) by dropping the terms associated with the (input and output) trade-offs or their dual weight restrictions, and by considering that

[93] Note that when we consider only outputs it makes no sense to have an input-oriented model. Also, as Lovell and Pastor (1999) have shown, a pure-output CRS output-oriented DEA model rates all DMUs as infinitely inefficient, while an input-oriented VRS DEA model with a single constant input rates all DMUs as efficient.

$i = 1,\ x^k = 1, k = 1, \dots, K$ which implies that $v^o = 1$.[94] The BoD model has recently been adapted to a VEA framework (see the third chapter in this Thesis) and its multiplier and envelopment form are given as:

$$
\begin{array}{llll}
\max\limits_{u_j^o} & \sum\limits_{j=1}^{J} u_j^o y_j^o & & \\
s.t. & \sum\limits_{j=1}^{J} u_j^o y_j^k \leq 1 & k = 1, \dots, K,\ k \neq r & \\
& \sum\limits_{j=1}^{J} u_j^o y_j^k = 1 & k = r = 1, \dots, R & \\
& u_j^o \geq 0 & j = 1, \dots, J &
\end{array}
\qquad
\begin{array}{lll}
\min\limits_{\lambda_k^o} & \sum\limits_{k=1}^{K} \lambda_k^o & \\
s.t. & \sum\limits_{k=1}^{K} \lambda_k^o y_j^k \geq y_j^o & j = 1, \dots, J \\
& \lambda_k^o \geq 0 & k = 1, \dots, K,\ k \neq r \\
& \lambda_k^o\ free & k = r = 1, \dots, R
\end{array}
\tag{6.21}
$$

where $r = 1, \dots, R$ refers to the DMUs comprising the MPS.

For $i = 1$ and $x^k = 1, k = 1, \dots, K,$ $(x^k - x^r) = 0, k = 1, \dots, K,\ r = 1, \dots, R,$ and thus the vector P_r^k in (6.13) is a scalar with a value equal to zero. Consequently, the associated weight restrictions in (6.14) consider only outputs, i.e., are of the AR-I type, as the second component in each of the relations in (6.14) is equal to zero. In a similar fashion, the vector P_r in (6.11) is a scalar with a value equal to -1 and the second component in each of the associated weight restrictions in (6.12) is also equal to -1, namely (6.12) considers only outputs. Thus, we can show the following:

PROPOSITION 6.3: The VEA BoD model is equivalent to a BoD model including production trade-offs, for which the trade-off coefficient vectors contain either (i) the negative of the output quantities of the DMUs constituting the MPS, or (ii) the deviations of each DMU's output quantities from those of each of the DMUs constituting the MPS.

Next, consider the DEA model without outputs, which is equivalent to a DEA model with a single or multiple constant (unitary) outputs (Lovell and Pastor, 1999).[95]

---

[94] Variants of (6.20) including weight restrictions have been employed in, among others, the construction of composite indicators of environmental performance (Zanella *et al.*, 2013), the re-estimation of the Technology Achievement Index (Cherchye *et al.*, 2008), and the aggregation of several measures of money into a synthetic indicator (Sahoo and Acharya, 2010).

[95] Note that when we consider only inputs it makes no sense to have an output-oriented model. Also, as Lovell and Pastor (1999) have shown, a pure-input CRS input-oriented DEA model rates all DMUs as

The latter is known as the Inverted BoD model and its multiplier and envelopment form are given as (Färe and Karagiannis, 2014): [96]

$$\begin{array}{llllll} \min_{v_i^o} & \sum_{i=1}^{I} v_i^o x_i^o & & \max_{\lambda_k^o} & \sum_{k=1}^{K} \lambda_k^o & \\ s.t. & \sum_{i=1}^{I} v_i^o x_i^k \geq 1 & k = 1,\dots,K & s.t. & \sum_{k=1}^{K} \lambda_k^o x_i^k \leq x_i^o & i = 1,\dots,I \\ & v_i^o \geq 0 & i = 1,\dots,I & & \lambda_k^o \geq 0 & k = 1,\dots,K \end{array} \quad (6.22)$$

The model in (6.22) is obtained from the output-oriented counterpart of (6.2) by dropping the terms associated with the (input and output) trade-offs or their dual weight restrictions, and by assuming that $j = 1$, $y^k = 1, k = 1,\dots,K$, which implies that $u^o = 1$. Compared to the BoD model, the Inverted BoD model provides a pessimistic perspective of performance evaluation (Karagiannis, 2021). Consider now a set of R DMUs reflecting the most desirable input bundle from DM's point of view. Then, the multiplier and envelopment form of the Inverted VEA BoD model will be given as:

$$\begin{array}{llllll} \min_{v_i^o} & \sum_{i=1}^{I} v_i^o x_i^o & & \max_{\lambda_k^o} & \sum_{k=1}^{K} \lambda_k^o & \\ s.t. & \sum_{i=1}^{I} v_i^o x_i^k \geq 1 & k = 1,\dots,K, k \neq r & s.t. & \sum_{k=1}^{K} \lambda_k^o x_i^k \leq x_i^o & i = 1,\dots,I \\ & \sum_{i=1}^{I} v_i^o x_i^k = 1 & k = r = 1,\dots,R & & \lambda_k^o \geq 0 & k = 1,\dots,K, k \neq r \\ & v_i^o \geq 0 & i = 1,\dots,I & & \lambda_k^o \; free & k = r = 1,\dots,R \end{array} \quad (6.23)$$

Then for $j = 1$ and $y^k = 1, k = 1,\dots,K$, we have that $\left(y^k - y^r\right) = 0, k = 1,\dots,K, r = 1,\dots,R$ and thus the vector Q_r^k in (6.13) is a scalar than takes the value of zero. Thus, the weights restrictions dual to the production trade-offs in (6.13) are AR-I, as the first component in each of the relations in (6.14) is equal to zero. Similarly, the vector Q_r in (6.11) is a scalar with a value equal to -1 and the same holds for the

---

infinitely inefficient, while an output-oriented VRS DEA model with a single constant output rates all DMUs as efficient

[96] Variants of (6.22) including weight restrictions have been used by, among others, Zhou *et al.* (2007) to construct a sustainable energy index, and Rogge (2012) to re-estimate the Environmental Performance Index.

first component in each of the associated weight restrictions in (6.12). Thus, we can show that:

PROPOSITION 6.4: The Inverted BoD VEA model is equivalent to an Inverted BoD model including production trade-offs, for which the trade-off coefficient vectors contain either (i) the negative of the input quantities of the DMUs constituting the MPS, or (ii) the deviations of each DMU's input quantities from those of each of the DMUs constituting the MPS.

6.4. An empirical application

To illustrate the usefulness of our findings, we consider the case of a DM evaluating alternatives in a technology selection problem, using the dataset of 27 industrial robots in Khouja (1995) and Baker and Talluri (1997).[97] For the purposes of this chapter, we may consider the DM assessing these 27 DMUs as either a potential buyer, i.e., the manager of an industrial plant, or a technology manufacturer, namely the owner of a company producing one of the assessed DMUs.

Data for the 27 DMUs are given in columns 2 to 5 of Table 6.1. Four among the most important performance features of industrial robots are considered, which are (i) the robots' cost (in 10.000$), (ii) repeatability, namely a measure of the distance (in mm) covered by the robot in repeated trials, (iii) the robot's payload capacity (in kg) and (iv) its minimum possible velocity (in m/s). For the former two features lower values indicate better performance, and hence they are treated as inputs, while larger values are more preferable for capacity and velocity and these are treated as outputs (Khouja, 1995). Efficiency estimates based on the input-oriented CRS and VRS DEA models are given in columns 6 and 9 of Table 6.1. From these, we can see that nine DMUs are efficient with CRS, while other two DMUs are added to the list of efficient DMUs in the VRS model. The assumption of VRS results in a noticeable increase in average efficiency (0.801 compared to 0.725 in the CRS model).

[97] The applications of DEA and other multi-criteria decision-making methods in technology selection are nowadays voluminous and include, but are not limited to, the selection of flexible manufacturing systems, industrial robots, and dispatching rules. A review of such applications is a task out of the scope of this chapter, and the interested reader is referred to Hamzeh and Xu (2019), for a recent review.

Table 6.1:Data and efficiency estimates for the illustrative example.

DMU	inputs		outputs		CRS models			VRS models		
	Cost (in 10.000$)	Repeatability (in mm)	Load capacity (in kg)	Velocity (in m/s)	DEA	VEA (MPS: #19)	VEA (MPS: #20)	DEA	VEA (MPS: #19)	VEA (MPS: #20)
1	7.200	0.150	60.000	1.350	1.000	0.871	0.479	1.000	1.000	0.512
2	4.800	0.050	6.000	1.100	0.904	0.511	0.417	0.907	0.868	0.437
3	5.000	1.270	45.000	1.270	0.529	0.465	0.529	0.667	0.507	0.568
4	7.200	0.030	1.500	0.660	1.000	0.195	0.167	1.000	1.000	0.196
5	9.600	0.250	50.000	0.050	0.592	0.392	0.108	0.594	0.594	0.177
6	1.070	0.100	1.000	0.300	0.482	0.414	0.482	0.865	0.865	0.865
7	1.760	0.100	5.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
8	3.200	0.100	15.000	1.000	0.783	0.783	0.618	0.783	0.783	0.618
9	6.720	0.200	10.000	1.100	0.378	0.362	0.306	0.383	0.376	0.315
10	2.400	0.050	6.000	1.000	1.000	0.891	0.758	1.000	1.000	0.759
11	2.880	0.500	30.000	0.900	0.671	0.671	0.669	0.677	0.677	0.677
12	6.900	1.000	13.600	0.150	0.102	0.099	0.069	0.142	0.142	0.142
13	3.200	0.050	10.000	1.200	1.000	0.874	0.701	1.000	0.990	0.744
14	4.000	0.050	30.000	1.200	1.000	1.000	0.658	1.000	1.000	0.691
15	3.680	1.000	47.000	1.000	0.613	0.561	0.613	0.624	0.607	0.623
16	6.880	1.000	80.000	1.000	0.604	0.592	0.437	0.604	0.604	0.441
17	8.000	2.000	15.000	2.000	0.405	0.272	0.405	1.000	0.362	0.525
18	6.300	0.200	10.000	1.000	0.365	0.355	0.299	0.367	0.367	0.299
19	0.940	0.050	10.000	0.300	1.000	1.000	0.733	1.000	1.000	1.000
20	0.160	2.000	1.500	0.800	1.000	0.169	1.000	1.000	1.000	1.000
21	2.810	2.000	27.000	1.700	0.852	0.397	0.852	1.000	0.774	1.000
22	3.800	0.050	0.900	1.000	0.829	0.509	0.476	0.913	0.906	0.477
23	1.250	0.100	2.500	0.500	0.694	0.648	0.694	0.923	0.923	0.923
24	1.370	0.100	2.500	0.500	0.636	0.606	0.636	0.847	0.847	0.847
25	3.630	0.200	10.000	1.000	0.553	0.553	0.511	0.556	0.556	0.511
26	5.300	1.270	70.000	1.250	0.581	0.581	0.577	0.771	0.582	0.613
27	4.000	2.030	205.000	0.750	1.000	1.000	1.000	1.000	1.000	1.000
average	4.224	0.589	28.315	0.929	0.725	0.584	0.563	0.801	0.753	0.628

Note: The data are taken from Baker and Talluri (1997).

For our purposes let’s assume that DMU #19 is chosen as the MPS. The DM in this case may be the manager of a manufacturing plant that operates using this particular robot, or be a potential buyer for which this robot has an attractive combination of low cost and low repeatability. The CRS and VRS VEA efficiency scores when DMU #19 is chosen as the MPS are given in columns 7 and 10 of Table 6.1. In the CRS case, five DMUs drop from the list of efficient DMUs compared to DEA, while when VRS is assumed the efficient DMUs are reduced to eight, compared to 11 in DEA. By Proposition 2, the same efficiency scores would result from respectively a CRS and a VRS DEA model including the following set of weight restrictions:

$$
\begin{array}{rrrrll}
50.000u_1^k & +1.050u_2^k & -6.260v_1^k & -0.100v_2^k & \leq 0 & \text{(DMU \#1)} \\
-4.000u_1^k & +0.800u_2^k & -3.860v_1^k & & \leq 0 & \text{(DMU \#2)} \\
35.000u_1^k & +0.970u_2^k & -4.060v_1^k & -1.220v_2^k & \leq 0 & \text{(DMU \#3)} \\
-8.500u_1^k & +0.360u_2^k & -6.260v_1^k & +0.025v_2^k & \leq 0 & \text{(DMU \#4)} \\
40.000u_1^k & -0.250u_2^k & -8.660v_1^k & -0.200v_2^k & \leq 0 & \text{(DMU \#5)} \\
-9.000u_1^k & & -0.130v_1^k & -0.050v_2^k & \leq 0 & \text{(DMU \#6)} \\
-5.000u_1^k & +0.700u_2^k & -0.820v_1^k & -0.050v_2^k & \leq 0 & \text{(DMU \#7)} \\
5.000u_1^k & +0.700u_2^k & -2.260v_1^k & -0.050v_2^k & \leq 0 & \text{(DMU \#8)} \\
 & +0.800u_2^k & -5.780v_1^k & -0.150v_2^k & \leq 0 & \text{(DMU \#9)} \\
-4.000u_1^k & +0.700u_2^k & -1.460v_1^k & & \leq 0 & \text{(DMU \#10)} \\
20.000u_1^k & +0.600u_2^k & -1.940v_1^k & -0.450v_2^k & \leq 0 & \text{(DMU \#11)} \\
3.600u_1^k & -0.150u_2^k & -5.960v_1^k & -0.950v_2^k & \leq 0 & \text{(DMU \#12)} \\
 & 0.900u_2^k & -2.260v_1^k & & \leq 0 & \text{(DMU \#13)} \\
20.000u_1^k & +0.900u_2^k & -3.060v_1^k & & \leq 0 & \text{(DMU \#14)} \\
37.000u_1^k & +0.700u_2^k & -2.740v_1^k & -0.950v_2^k & \leq 0 & \text{(DMU \#15)} \\
70.000u_1^k & +0.700u_2^k & -5.940v_1^k & -0.950v_2^k & \leq 0 & \text{(DMU \#16)} \\
5.000u_1^k & +1.700u_2^k & -7.060v_1^k & -1.950v_2^k & \leq 0 & \text{(DMU \#17)} \\
 & 0.700u_2^k & -5.360v_1^k & -0.150v_2^k & \leq 0 & \text{(DMU \#18)} \\
-8.500u_1^k & +0.500u_2^k & +0.780v_1^k & -1.950v_2^k & \leq 0 & \text{(DMU \#20)} \\
17.000u_1^k & +1.400u_2^k & -1.870v_1^k & -1.950v_2^k & \leq 0 & \text{(DMU \#21)} \\
-9.100u_1^k & +0.700u_2^k & -2.860v_1^k & & \leq 0 & \text{(DMU \#22)} \\
-7.500u_1^k & +0.200u_2^k & -0.310v_1^k & -0.050v_2^k & \leq 0 & \text{(DMU \#23)} \\
-7.500u_1^k & +0.200u_2^k & -0.430v_1^k & -0.050v_2^k & \leq 0 & \text{(DMU \#24)} \\
 & 0.700u_2^k & -2.690v_1^k & -0.150v_2^k & \leq 0 & \text{(DMU \#25)} \\
60.000u_1^k & +0.950u_2^k & -4.360v_1^k & -1.220v_2^k & \leq 0 & \text{(DMU \#26)} \\
195.000u_1^k & +0.450u_2^k & -3.060v_1^k & -1.980v_2^k & \leq 0 & \text{(DMU \#27)}
\end{array}
$$

The figures attached to the input and output weights in each of the above restrictions are equal to the deviations of the input and output quantities corresponding to the evaluated DMU listed in parentheses, from those of DMU #19. This set of restrictions forces the marginal rates of substitution and transformation for the evaluated DMUs to take values only within the range of marginal rates prevailing on the efficient frontier in the neighborhood of DMU #19. Note that by Proposition 1 the CRS VEA efficiency scores could also be obtained through a DEA model including the following weight restriction:

$$-10.000u_1^k \quad -0.300u_2^k \quad +0.940v_1^k \quad +0.050v_2^k \quad \leq 0$$

the coefficients of which are the negative of the input and output quantities of DMU #19.

Let us now assume that the following trade-off is included in the envelopment form of the CRS DEA model in (6.1):

$$P = [-0.160, -2.000]^T$$
$$Q = [-1.500, -0.800]^T$$

This trade-off implies that, if the DM is willing to accept a decrease in load capacity by 1.500 kg and in velocity by 0.800 m/s, then the robot's cost and repeatability could be decreased by at most 1600$ and 2.000 mm respectively. The above trade-off coefficient vectors are equal to the negative of the input and output quantities of the DEA-efficient DMU #20. Thus, by Proposition 1, a CRS DEA model augmented with this trade-off is equivalent to a CRS VEA model in which DMU #20 is chosen as the MPS. The resulting efficiency scores when either the above trade-off is included in the CRS DEA model, or DMU #20 is the MPS in the CRS VEA model are given in column 8 of Table 6.1. Compared to the DEA results, we see that only three DMUs remain efficient in the VEA model, while average efficiency decreases to 0563.

The efficiency scores from a VRS VEA model in which DMU #20 is used as the MPS are given in column 11 of Table 6.1. By Proposition 2, these scores can be obtained via a DEA model including production trade-offs the coefficient vectors of which contain the deviations of each DMU's input and output quantities from those of DMU #20.

6.5. Concluding remarks

In this chapter, we have examined the links between DEA models with weights restrictions or their dual production trade-offs and VEA and we showed that VEA may be viewed as a class of DEA models with particular trade-offs. More specifically, we showed that, irrespective of the nature of the returns to scale, the VEA model is equivalent to the DEA model including production trade-offs, for which the trade-off coefficient vectors are given by the deviations of the input and output quantities of each sample DMU from those of DMUs chosen as the MPS. In addition, with CRS, the VEA model is equivalent to the DEA model with trade-off coefficient vectors given by the negative of the input and output quantities of the DMUs chosen as the MPS in VEA. These trade-offs are dual to AR-II weight restrictions. In addition, we showed that, when we are considering these particular trade-offs only for the inputs or the outputs, a similar equivalence results between the pure output or input VEA models and their DEA counterparts including trade-offs. In these cases, the dual forms of the trade-offs are AR-I weight restrictions.

The results in this chapter indicate that the DM preferences about the most preferred input/output structure as reflected in the MPS in the VEA model may be seen as a particular form of trade-offs or their dual AR-II or AR-I type of weight restrictions and vice versa. This provides an alternative interpretation of the efficiency scores obtained from both the VEA model and its equivalent DEA model including production trade-offs. In particular, the VEA efficiency scores can also be interpreted as including restrictions in the acceptable values of the marginal rates of substitution and transformation, while it may be said that the efficiency scores obtained from the DEA model including production trade-offs reflect the DM's judgements about the most preferred input/output structure.

CHAPTER 7

On Value Efficiency Analysis and Cone-Ratio DEA models

7.1. Introduction

Weight restrictions, their dual production trade-offs, Cone ratio Data Envelopment Analysis (CR-DEA) and Value Efficiency Analysis (VEA) are alternative and seemingly unrelated ways for incorporating Decision Maker's (DM) preferences in DEA models. In the literature there is a long interest about their inter-relations, from which several useful results have emerged. *First*, Charnes *et al.* (1990) demonstrated that CR-DEA models including absolute or relative bounds on the input/output multipliers are equivalent to DEA models with Assurance Region type I or II weight restrictions. *Second*, Olesen and Petersen (2003) inferred that, in the case of a single model DMU, a CR-DEA model, in which the set of feasible weights consists of all the weight vectors that are optimal in DEA for the *model* DMU provides the same efficiency scores with a VEA model in which the *model* DMU is chosen as the Most Preferred Solution (MPS).[98] In addition, they stated that in the case of multiple *model* DMUs the efficiency scores of a CR-DEA model, in which the set of feasible weights is the union of the sets of the weight vectors that are optimal in DEA for each of the model DMUs, are given by the maximum among the scores of a series of VEA models, each of which uses one among the *model* DMUs as the MPS. *Third,* Podinovski (2004;2005) showed that DEA models with non-homogeneous weight restrictions are equivalent to DEA models including a particular form of production trade-offs. *Fourth*, Ravanos and Karagiannis (2022a) have illustrated that the choice of the MPS in VEA

[98] *Model* DMUs are DEA-efficient DMUs viewed as exceptional performers by the DM (Charnes *et al.*, 1990).

models can be viewed as a particular form of trade-offs or AR-II type of weight restrictions and *vice versa*.

In this paper, we elaborate more on the relations between VEA and CR-DEA models following two distinct routes: *first,* for the case of multiple *model* DMUs, we consider in addition the case of CR-DEA models, in which the feasible weight vectors are given by the *intersection* of the sets of the weight vectors that are optimal in DEA for each of the *model* DMUs. In this case, all the *model* DMUs will be rendered efficient when each DMU is evaluated, whereas in the case of Olesen and Petersen (2003), where the set of feasible weights is the union of the sets of the weight vectors that are optimal in DEA for each of the *model* DMUs, at least one of the *model* DMUs will be efficient. The proposed formulation can accommodate cases where the DM wishes to compare each evaluated DMU with all chosen *model* DMUs, which in some empirical applications may be a valuable option. *Second*, we extend the comparison between VEA and CR-DEA models by considering a formulation of the CR-DEA model in which the set of feasible weight vectors contains only those vectors with strictly positive components, each of which is optimal in DEA for the (single or the multiple) *model* DMUs. This model essentially extends each Fully Dimensional Efficient Facet (FDEF) jointly generated by all the *model* DMUs.

Our theoretical results indicate that, in the case of multiple model DMUs, the efficiency score of a CR-DEA model, in which the feasible weight vectors are given as the intersection of the sets of weight vectors that are optimal in DEA for each of the *model* DMUs, is equal to that of a VEA model in which the chosen *model* DMUs are also considered as the MPS. The former is also lower than or equal to the minimum among the efficiency scores obtained from a number of VEA models, each of which uses a single *model* DMU as the MPS. On the other hand, we verify that the efficiency score of a CR-DEA model, in which the feasible weight vectors are given as the union of the sets of weight vectors that are optimal in DEA for each of the *model* DMUs, is greater than or equal to that of a VEA model in which the MPS comprises of all the *model* DMUs. We also show that the efficiency score of a CR-DEA model, in which the set of feasible weight vectors contains only those vectors with strictly positive components, each of which is optimal in DEA for the (single or multiple) *model* DMUs is lower than or equal to that of a VEA model in which the chosen *model* DMUs are also considered as the MPS.

An immediate implication of our theoretical results is that the efficiency scores of the alternative CR-DEA models can be estimated or approximated by means of either a single VEA model in which the chosen *model* DMUs are also considered as the MPS or a series of VEA models that use different *model* DMUs as the MPS. We expect that this will facilitate the empirical applications of the CR-DEA models, which are appropriate for several study cases. For example, in assessing the performance in the banking industry where certain banks or bank branches are viewed as excellent performers or "global leaders". Agreement on the set of *model* banks may be unanimous or not, as DMs might have diverging views on what constitutes good performance, or the chosen banks are viewed as different types of examples to follow. Another case could concern the assessment of higher education institutions in terms of their research quality and teaching excellence, in which the set of *model* institutions may be unanimously viewed as top-performing or excel in different evaluation dimensions. The use of CR-DEA models in such cases has remained up to date limited as their estimation required identifying all the efficient facets generated by each of the *model* DMUs and the weight vectors normal to each of them (Olesen and Petersen, 2003; Portela and Thanassoulis, 2006). This can only be done via complex non-linear programs or additional software that can prove to be quite complicated and time consuming (Zhu *et al.*, 2022). In contrast, VEA models --which can be used to estimate or approximate CR-DEA scores--involve only changing some linear inequalities in the multiplier form of a DEA model to equalities.

We illustrate the usefulness of our findings using data from Japanese regional banks. Using various sets of *model* DMUs we illustrate how the CR-DEA efficiency scores can be obtained or approximated by VEA. We also discuss the intuition behind the choice of different sets of *model* DMUs, which is useful for empirical applications. The rest of the paper unfolds as follows: A literature review follows in the next section, while in the third section we present the CR-DEA and VEA models. The papers' main results are presented in the fourth section, while the empirical application is discussed in the fifth section. Finally, concluding remarks follow in the last section.

7.2. CR-DEA and VEA: A brief overview

CR-DEA models were developed by Charnes *et al.* (1989; 1990) to cope with unsatisfactory results from conventional DEA, which identified "notoriously inefficient [DMUs] as efficient" (Charnes *et al.*, 1990, p. 74).[99] In the multiplier form of these models, the range of optimal input/output weight vectors is restricted to cones smaller than the non-negative orthant. These cones can reflect preferences over the relative importance of inputs and/or outputs, information about the variation of their prices, or alternatively they can be defined based on the weight vectors that are optimal for a certain set of *model* DMUs that DMs view as exceptional performers (Charnes *et al.*, 1990). In both cases, this information is used to define transformation matrices that modify the input/output quantities of the evaluated DMUs in the envelopment form of these models.

In CR-DEA models using information about price variation or views over input/output importance, separate cones may be used to restrict the input and the output weight sub-vectors, or a cone can be defined for the input/output weight vector. When separate cones are used, the input and the output weight sub-vector of each DMU are allowed to vary independently from each other in the multiplier form of the model and separate transformation matrices are used to modify input and output quantities in its envelopment form (Portela and Thanassoulis, 2006). Charnes *et al.* (1990) have shown that in this case the CR-DEA model is equivalent to DEA models including AR-I type of weight restrictions (Thompson *et al.*, 1986). Empirical applications of such CR-DEA models include the assessment of primary care physicians (Chilingerian and Sherman, 1997), textile factories (Zhu, 1996), and manufacturing technologies (Talluri and Yoon, 2000). Recently, Ding *et al.* (2015) introduced such restrictions in a network DEA model with shared resources. When a single cone is used to restrict the input/output weight vector, there are also restrictions linking input and output weights that correspond to AR-II type restrictions (Thompson *et al.*, 1990). Thompson *et al.* (1994) used information on input and output prices and costs to determine upper and lower bounds for ratios of input, output and input and output multipliers in an evaluation of oil companies. Subsequent applications of such CR-DEA models include the

[99] From now on, when a quotation is used, the words in brackets and the underlying are our own additions.

evaluation of coal mines (Thompson *et al*, 1995) and US banks (Thompson *et al.*, 1996).

When preference information concerns *model* DMUs, which is the focus of this paper, the cone of feasible weights contains only weight vectors which are optimal for the chosen *model* DMUs in the DEA model (Olesen and Petersen, 2003). In this case, "the [weights] are restricted to lie in hyperplanes" (Charnes *et al.*, 1991b, p. 2070), which implies that the resulting efficient frontier is defined by extending the DEA facets generated by the *model* DMUs. Charnes *et al.* (1990, p. 79) used the optimal weight vectors ("efficient basic duals") obtained from the linear DEA model for each of the *model* DMUs to define the cone of feasible weights. Notice that each such weight vector is optimal for one of the *model* DMUs but not necessarily for the others as well, and that it may contain zero components. It is thus hard to tell whether Charnes *et al.* (1990) intention was to include only weight vectors that are optimal for *all* the chosen *model* DMUs or to account for every weight vector that is optimal for at least one of the *model* DMUs. Clarifying this issue has important implications for empirical applications. The same holds for whether optimal weight vectors containing zero components should be included in the cone or not.[100] In addition, Charnes *et al.* (1990) defined the cone of feasible weights in such a way that the input and the output weight sub-vectors were allowed to vary independently from each other within separately defined cones. Empirical applications of this form of CR-DEA models include performance assessment in the banking sector (see Charnes *et al.*, 1990 and Brockett *et al.*, 1997). Later, Tone (1997) noted that for each *model* DMU there exist multiple optimal weight vectors and developed three linear programs to choose one among those vectors for each *model* DMU to use in defining the cone of feasible weights.

The process followed by Charnes *et al.* (1990) to define the cone of feasible weights was thoroughly challenged by Olesen and Petersen (2003), in that i) each *model* DMU generates more than one efficient facets and thus there exist multiple vectors of optimal weights for it, all of which should be considered in the cone of feasible weights, and ii) allowing the input and the output weight sub-vectors to vary independently from

[100] Charnes et al. (1990) used weight vectors with strictly positive components to define the cone of feasible weights in their application. On the other hand, Brockett et al. (1997) included weight vectors containing zero components as well.

each other may result in the evaluated DMUs adopting optimal input weights referring to one *model* DMU and optimal output weights associated with another, and thus a single cone should be defined using the input/output weight vectors instead of separate cones for input and output weight sub-vectors. They noted (p. 357) that the efficient frontier of the CR-DEA model can be estimated by extending facets with well-defined rates of input substitution (MRSs) and output transformation (MRTs) and "other" facets (i.e., the cone of feasible weights can be defined using both weight vectors with strictly positive components and vectors having zero components). Thus, to properly incorporate preferences regarding a set of *model* DMUs in CR-DEA, one needs to identify all the facets of the DEA frontier generated by the *model* DMUs and the weight vectors that are normal to each of them (Portela and Thanassoulis, 2006). This requires identifying all the facets of the DEA frontier (Thanassoulis, *et al.* 2008).

Olesen and Petersen (1996; 2015) proposed the use of a mixed integer model to obtain CR-DEA efficiency scores when information on *model* DMUs is available. The cone of feasible weights in this model is defined using only weight vectors that contain strictly positive components. The resulting efficient frontier is defined by extending facets with well-defined MRSs and MRTs generated by the *model* DMUs. However, optimal weight vectors for *model* DMUs frequently contain zero components as well. Thus, this model is considered hereafter as a distinct variant of CR-DEA.

On the other hand, VEA (Halme *et al.,* 1999) accommodates DM preferences over the most favorable input/output structure by means of an implicitly known value function (i.e., an indifference curve), which is maximized at a point on the strongly efficient DEA frontier that constitutes the MPS. The MPS essentially corresponds to either an extreme-efficient DMU or a combination of extreme-efficient DMUs that are *jointly* efficient, and is chosen by the DM by means of various criteria (see Korhonen *et al.* (2002) for an early exploration and the fourth chapter in this Thesis for a recent review). In the multiplier form of the VEA model, each DMU is implicitly compared to the MPS by restricting the feasible weight vectors to those that are optimal in the DEA model for *all* the DMUs that constitute the MPS. This model appeared for the first time in Oral and Yolalan (1990), who used it to compare the performance of banks in terms of efficiency and profitability to that of a particular efficient bank ("global leader"). This bank was chosen at a previous step. Subsequent studies explored the potential of obtaining VEA scores which are better approximations of the scores that could be obtained if an explicit functional form was available for the DM's value function, and their interpretation in terms of value differences between the assessed DMU and the MPS (see Joro *et al.*, 2003; Korhonen and Syrjanen, 2005). Recent theoretical advancements include non-convex (Halme *et al.*, 2014) and non-radial (Gerami *et al.*, 2022) VEA models. Regarding empirical work, VEA has been applied for the evaluation of hospital departments (Halme and Korhonen, 2000), higher education institutions (Korhonen *et al.*, 2001), local governments (Marshall and Shortle, 2005), banks (Halme *et al.*, 2014; Eskelinen *et al.*, 2014), and the construction of composite indicators (see the third chapter in this Thesis)

7.3. Materials and methods

7.3.1. Preliminaries

Let us consider a set of *K* DMUs *(k=1,…,K)* that use the same technology to produce a set of J *(j=1,…,J)* different outputs utilizing I *(i=1,…,I)* different inputs. The fractional programming form of an input-oriented variable-returns-to-scale (VRS) DEA model for the o^{th} DMU is given as (Banker *et al.*, 1984):[101]

$$
\begin{aligned}
&\max_{\xi_j^o,\omega_i^o,\zeta^k} \left(\sum_{j=1}^{J} \xi_j^o y_j^o + \zeta^o\right) \Bigg/ \sum_{i=1}^{I} \omega_i^o x_i^o \\
&s.t. \quad \left(\sum_{j=1}^{J} \xi_j^o y_j^k + \zeta^o\right) \Bigg/ \sum_{i=1}^{I} \omega_i^o x_i^k \le 1 \quad \forall\, k \\
&\qquad \xi_j^o \ge 0 \qquad \forall\, j \\
&\qquad \omega_i^o \ge 0 \qquad \forall\, i \\
&\qquad \zeta^o \ \ free
\end{aligned}
\tag{7.1}
$$

where x and *y* are respectively the quantities of inputs and outputs, *ω* and *ξ* are their weights, and *ζ* is a free variable to be estimated. The constant-returns-to-scale (CRS) counterpart of (7.1) is obtained by removing the free variable *ζ* (see Charnes *et al.*, 1978).

Denote $\mathbb{E}$ as the set containing the *extreme efficient* DMUs in model (1), namely those residing at a point of the convex DEA efficient frontier where more than one

[101] We limit our discussion to input-oriented DEA models. The extension of the results to output oriented models is straightforward.

facets intercept. The polyhedral cone containing the feasible input/output weight vectors in (7.1) is given as (Räty, 2002; Olesen and Petersen, 2015, p. 155):

$$\mathcal{F} = \left\{ \left(\xi_j^k, -\omega_i^k, \zeta^k\right) \in \mathbb{R}_+^J \times \mathbb{R}_-^I \times \mathbb{R} : \sum_{j=1}^{J} \xi_j^k y_j^k - \sum_{i=1}^{I} \omega_i^k x_i^k + \zeta^k \leq 0, \left(\xi_j^k, \omega_i^k\right) \neq (0,0),\ \ k \in \mathbb{E} \right\} \tag{7.2}$$

The cone in (7.2) is determined from the halfspace constraints in (7.1) and is represented as a collection, i.e., a set, of vectors $(\xi_j^k, -\omega_i^k, \zeta^k)$.[102],[103] It is expressed in terms of the extreme efficient DMUs only, since these generate the facets of the DEA efficient frontier and they are used to "evaluate all of the points that represent the performances of the DMUs that are to be evaluated" (Cooper *et al.*, 2007b, p. 444). In essence, this means that any weight vector that is feasible for a particular DMU will also be feasible for at least one extreme efficient DMU. For this reason, all similar cones from now on will be expressed in terms of the extreme efficient DMUs.

For efficient DMUs --and in rare occasions also for some inefficient DMUs-- there exist more than one weight vectors that are optimal in (7.1). The cone containing all the weight vectors in $\mathcal{F}$ which are optimal for a DMU $k \in \mathbb{E}$, (namely, render it efficient) is given as (Olesen and Petersen, 2015, p. 157):

$$\mathcal{F}^k = \left\{ \left(\xi_j^k, -\omega_i^k, \zeta^k\right) \in \mathcal{F} : \sum_{j=1}^{J} \xi_j^k y_j^k - \sum_{i=1}^{I} \omega_i^k x_i^k + \zeta^k = 0 \right\} \tag{7.3}$$

[102] If a weight vector $(\xi_j^k, \omega_i^k, \zeta^k)$ is optimal for some DMU in (1) then $\left(a\xi_j^k, a\omega_i^k, \alpha\zeta^k\right),\ \alpha > 0$ will also be optimal for this DMU. Thus, in the input/output weights space, $\mathcal{F}$ is a polyhedral cone spanned by the weight vectors $(\xi_j^k, \omega_i^k, \zeta^k)$ and containing all their multiples $\left(a\xi_j^k, a\omega_i^k, \alpha\zeta^k\right),\ \alpha > 0$. Notice that α could also be equal to $1/\xi_j^k$ for some *j*, or to $1/\omega_i^k$ for some *i*, in the sense that one of the positive input or output weights is used as a numeraire.

[103] In Olesen and Petersen (2015), the set $\mathbb{E}$ contains the strongly efficient DMUs instead of the extreme efficient DMUs, while Räty (2002) uses the extreme efficient DMUs to define $\mathcal{F}$. Strongly efficient DMUs are further classified into extreme efficient and non-extreme efficient DMUs. The latter reside in the interior of a facet of the strongly efficient frontier (Charnes et al., 1991a). Non-extreme efficient DMUs can be expressed as linear combinations of the extreme efficient DMUs generating the facet in which they reside. In this sense, the weight vectors that are optimal for a non-extreme efficient DMU are also optimal for all the extreme efficient DMUs generating the facet in which it resides, and they are already included in $\mathcal{F}$.

Each weight vector contained in $\mathcal{F}^k$ is a generating normal vector for a supporting hyperplane of the DEA efficient frontier with DMU k located on it.[104] It may contain only strictly positive input and output weight components or may be associated with at least one zero input or output weight. In the former case, the vector is normal to a facet of the strongly DEA efficient frontier in which the MRSs and MRTs are well-defined and can thus be given an interpretation in terms of relative prices, while ζ^k can be interpreted as a measure of local scale elasticity (Olesen and Petersen, 2015). Such facets are jointly generated --provided that certain regularity conditions are met-- by a unique combination of $(I + J - 1)$ extreme efficient DMUs in CRS DEA models and one of $(I + J)$ extreme efficient DMUs when VRS is assumed (Olesen and Petersen, 2003). In the latter case, the weight vector $(\xi_j^k, -\omega_i^k, \zeta^k)$ is normal to a facet of the weakly DEA efficient frontier, in the sense that MRSs and MRTs are not well-defined. Following Olesen and Petersen (1996; 2003; 2015), we term the former facets as Fully Dimensional Efficient Facets (FDEFs) and the latter as non-Fully Dimensional Efficient Facets (non FDEFs).[105] It is possible that $\mathcal{F}^k$ contains only weight sets of the latter form, namely that the DMU k does not contribute to generating an FDEF.

Let us consider a subset $\mathcal{R} \subseteq \mathbb{E}$ of extreme efficient DMUs. The cone:

$$\mathcal{F}_{\mathbb{I}}^{\mathcal{R}} = \bigcap\nolimits_{k \in \mathcal{R}} \mathcal{F}^k = \left\{ (\xi_j^k, -\omega_i^k, \zeta^k) \in \mathcal{F}\colon \sum_{j=1}^{J} \xi_j^k y_j^k - \sum_{i=1}^{I} \omega_i^k x_i^k + \zeta^k = 0\,, \forall\, k \in \mathcal{R} \right\} \quad (7.4)$$

contains all the weight vectors, each of which is optimal for *all* the DMUs in $\mathcal{R}$, i.e., renders *all* of them efficient (Olesen and Petersen, 2015), where $\mathbb{I}$ refers to *intersection*. Thus, each weight vector in $\mathcal{F}_{\mathbb{I}}^{\mathcal{R}}$ is a generating normal vector for a supporting hyperplane of the (weakly or strongly) efficient frontier with each and every member of $\mathcal{R}$ located on it (Banker *et al.*, 1984). $\mathcal{F}_{\mathbb{I}}^{\mathcal{R}} \neq \emptyset$ implies the existence of at least one

---

[104] In a similar fashion with $\mathcal{F}$, in the input/output weights space, $\mathcal{F}_k$ is a polyhedral cone spanned by the weight vectors $(\xi_j^k, \omega_i^k, \zeta^k)$ and containing all their multiples. See Olesen and Petersen (1996) for the CRS counterparts of (2) and (3).

[105] See Olesen and Petersen (1996) for a formal definition of FDEFs and non FDEFs and Olesen and Petersen (2015) for an interesting discussion on alternative definitions of efficient facets. Linear models that test for the existence of at least one FDEF in the empirical CRS or VRS DEA frontier are provided in Olesen and Petersen (2015). Given a set of $\mathcal{R}$ extreme-efficient DMUs, one may investigate whether these jointly generate at least one FDEF by identifying all the FDEFs of the empirical DEA frontier.

weight vector $(\xi_j^k, -\omega_i^k, \zeta^k)$ which renders all DMUs in set $\mathcal{R}$ efficient, namely that the DMUs in $\mathcal{R}$ *jointly* generate at least one facet of the DEA efficient frontier. If $\mathcal{F}_{\mathbb{I}}^{\mathcal{R}} = \emptyset$ such a facet does not exist. However, there might be facets generated by a subset of the DMUs in $\mathcal{R}$. Furthermore, if the number of DMUs in $\mathcal{R}$ equal to $I + J - 1$ (in CRS models) or to $I + J$ (in VRS models), then $\mathcal{F}_{\mathbb{I}}^{\mathcal{R}} \neq \emptyset$ only when the DMUs in $\mathcal{R}$ jointly generate an FDEF, and it will contain only *one* vector with strictly positive input/output weights, namely the one that is normal to the FDEF generated by the DMUs in $\mathcal{R}$. Lastly, $\mathcal{F}_{\mathbb{I}}^{\mathcal{R}} = \emptyset$ when the number of DMUs in $\mathcal{R}$ is greater than $I + J - 1$ (in CRS models) or $I + J$ (in VRS models), since DEA facets cannot be spanned by more than $(I + J - 1)$ (in CRS models) and $(I + J)$ DMUs (in VRS models). It is also evident from (7.4) that when the set $\mathcal{R}$ contains only one DMU, then $\mathcal{F}_{\mathbb{I}}^{k} \equiv \mathcal{F}^{k}$.

We may also define the following set:

$$\mathcal{F}_{\mathbb{U}}^{\mathcal{R}} = \bigcup_{k \in \mathcal{R}} \mathcal{F}^{k} = \left\{ (\xi_j^k, -\omega_i^k, \zeta^k) \in \mathcal{F} : \sum_{j=1}^{J} \xi_j^k y_j^k - \sum_{i=1}^{I} \omega_i^k x_i^k + \zeta^k = 0 \; for\ at\ least\ one\ k \in \mathcal{R} \right\} \quad (7.5)$$

which contains each weight vector among those in $\mathcal{F}$ that is optimal in the DEA model (1) for *at least one* of the extreme efficient DMUs in set $\mathcal{R}$, and $\mathbb{U}$ refers to *union*. Thus, a weight vector $(\xi_j^k, -\omega_i^k, \zeta^k) \in \mathcal{F}_{\mathbb{U}}^{\mathcal{R}}$ will be optimal for one --or more-- among the DMUs in $\mathcal{R}$, but may or may not be optimal for the others. By comparing (7.4) and (7.5) we see that: i) $\mathcal{F}_{\mathbb{I}}^{\mathcal{R}}$ is a subset of $\mathcal{F}_{\mathbb{U}}^{\mathcal{R}}$, ii) $\mathcal{F}_{\mathbb{U}}^{\mathcal{R}} \neq \emptyset$ even if there does not exist a facet of the DEA efficient frontier generated by all the DMUs in $\mathcal{R}$, and iii) when $\mathcal{R}$ contains only one DMU, then $\mathcal{F}_{\mathbb{U}}^{k} \equiv \mathcal{F}^{k} \equiv \mathcal{F}_{\mathbb{I}}^{k}$. Furthermore, it is evident from (8) that if $\mathcal{R} \equiv \mathbb{E}$, then $\mathcal{F}_{\mathbb{U}}^{\mathbb{E}} \equiv \mathcal{F}$. However, this is a sufficient but not a necessary condition for $\mathcal{F}_{\mathbb{U}}^{\mathcal{R}} \equiv \mathcal{F}$.

The model in (7.1) can be converted to the following linear model:

$$
\begin{aligned}
\max_{u_j^o, v_i^o, z^k} \quad & \sum_{j=1}^{J} u_j^o y_j^o + u^k \\
s.t. \quad & \sum_{j=1}^{J} u_j^o y_j^k - \sum_{i=1}^{I} v_i^o x_i^k + u^k \leq 0 \ \forall\, k \\
& \sum_{i=1}^{I} v_i^o x_i^o = 1 \\
& u_j^o \geq 0 \qquad\qquad \forall\, j \\
& v_i^o \geq 0 \qquad\qquad \forall\, i \\
& z^k \ \ free
\end{aligned}
\tag{7.6}
$$

where $\xi_j^o = u_j^o/\beta^o$, $\omega_i^o = v_i^o/\beta^o$, $\zeta^k = u^k/\beta^o$ and $\beta^o = (\sum_{i=1}^{I} \omega_i^o x_i^o)^{-1}$. Then the set containing the weight vectors which are optimal for a DMU $k \in \mathbb{E}$ in (7.6) is given as:

$$
\tilde{\mathcal{F}}^k = \left\{ (u_j^k, -v_i^k, u^k) = \beta^o(\xi_j^k, \omega_i^k, \zeta^k) \middle| (\xi_j^k, -\omega_i^k, \zeta^k) \in \mathcal{F}^k,\ \beta^o = \left(\sum_{i=1}^{I} \omega_i^k x_i^k\right)^{-1} \right\} \tag{7.7}
$$

which contains the multiples of the weight vectors in $\mathcal{F}^k$ that satisfy the normalizing equality in the second constraint in (7.6) when k is the evaluated DMU.

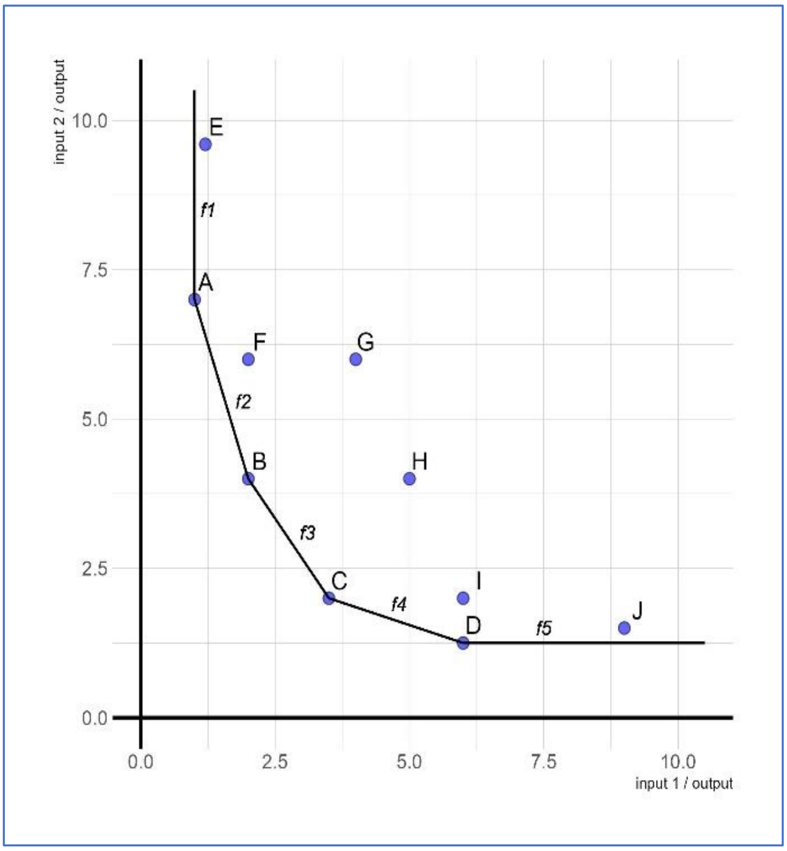
To demonstrate the alternative sets of weight vectors discussed above, let's consider an illustrative example involving 10 DMUs, each using two inputs to produce a single output. Data for these DMUs are given in the upper panel of Table 7.1, while the middle and the lower panel contain the DMUs' input-oriented CRS DEA efficiency scores, the *normalized* optimal vectors of input/output weights $H^o = \left(\frac{v_1^o}{u^o}, \frac{v_2^o}{u^o}, 1\right) = \left(\frac{\omega_1^o}{\xi^o}, \frac{\omega_2^o}{\xi^o}, 1\right)$ and the optimal values of the intensity variables.[106] The DEA efficient frontier is given in Figure 7.1, where we see that DMUs *A*, *B*, *C*, and *D* are extreme-efficient, while the remaining DMUs are inefficient. As such, the set $\mathcal{R}$ may contain only one of the extreme-efficient DMUs *A*, *B*, *C*, and *D*, or any combination of more than two of them. In the latter case, $\mathcal{R}$ will be equal to either $\{A,B\}, \{A,C\}, \{A,D\}, \{B,C\}, \{B,D\}, \{C,D\}, \{A,B,C\}, \{A,B,D\}, \{A,C,D\}, \{B,C,D\}$, or

[106] Note that, since the DMUs in our example produce the same quantity of a single output, the CRS efficiency scores displayed in Table 1 are equivalent to those of a VRS DEA model (see Lovell and Pastor, 1999).

Table 7.1: Data and DEA optimal solutions for the illustrative example

DMU	*A*	*B*	*C*	*D*	*E*	*F*	*G*	*H*	*I*	*J*
x_1	1	2	3.5	6	1.2	2	4	5	6	9
x_2	7	4	2	1.25	9.6	6	6	4	2	1.5
y	1	1	1	1	1	1	1	1	1	1
optimal solution (1) from the DEA model										
ω_1^*/ξ^*	0.300	0.200	0.200	0.098	1.000	0.300	0.200	0.200	0.098	0.000
ω_2^*/ξ^*	0.100	0.150	0.150	0.328	0.000	0.100	0.150	0.150	0.328	0.800
θ^*	1.000	1.000	1.000	1.000	0.833	0.833	0.588	0.625	0.803	0.833
λ^*	λ_A^*=1.000	λ_B^*=1.000	λ_C^*=1.000	λ_D^*=1.000	λ_A^*=1.000	λ_A^*=0.333	λ_B^*=0.765	λ_B^*=0.250	λ_C^*=0.474	λ_E^*=1.000
						λ_B^*=0.667	λ_C^*=0.235	λ_C^*=0.750	λ_D^*=0.526	
facet	*f2*	*f3*	*f3*	*f4*	*f1*	*f2*	*f3*	*f3*	*f4*	*f5*
optimal solution (2) from the DEA model										
ω_1^*/ξ^*	1.000	0.300	0.098	0.000						
ω_2^*/ξ^*	0.000	0.100	0.328	0.800						
θ^*	1.000	1.000	1.000	1.000						
λ^*	λ_A^*=1.000	λ_B^*=1.000	λ_C^*=1.000	λ_D^*=1.000						
facet	*f1*	*f2*	*f4*	*f5*						

Figure 7.1: The DEA efficient frontier for the Illustrative example



$\{A, B, C, D\}$. From Table 7.1 we see that there are two input/output weight vectors which are optimal for each of the extreme-efficient DMUs. For DMU A these are $H_1^A = (1,0,1) = H_1$ and $H_2^A = (0.3, 0.1, 1)$, which are the vectors normal to the facets *f1* and *f2* respectively. For DMU B, the optimal input/output weights vectors are$H_1^B = (0.3, 0.1, 1) = H_2^A = H_2$ and $H_2^B = (0.2, 0.15,1)$. The latter is normal to facet *f3*. The optimal weight vectors for DMU C are $H_1^C = (0.2, 0.15,1) = H_2^B = H_3$ and $H_2^C = (0.098, 0.328,1)$, of which the latter is normal to facet *f4*. Lastly, the two weight vectors that render DMU D efficient are $H_1^D = (0.098, 0.328,1) = H_2^C = H_4$ and $H_2^D = (0, 0.8,1) = H_5$, which are normal to facets *f4* and *f5*, respectively. Thus, the set $\mathcal{F}$ is given as:

$$\mathcal{F} = \{H_1, H_2, H_3, H_4, H_5\}$$

Among the five weight vectors in $\mathcal{F}$, H_2, H_3 and H_4 are FDEFs, while H_1 and H_5 are non FDEFs. Furthermore, the sets $\mathcal{F}^k,\ k = A, B, C\ or\ D$ containing the optimal vectors of weights for each of the extreme-efficient DMUs are defined as:

$$\mathcal{F}^A = \{H_1, H_2\},\ \ \mathcal{F}^B = \{\, H_2, H_3\},\ \ \mathcal{F}^C = \{H_3, H_4\},\ \ \mathcal{F}^D = \{H_4, H_5\}$$

On the other hand, there is only one optimal weight vector for each of the inefficient DMUs (see Table 1).

The set $\mathcal{F}_{\mathbb{I}}^{\mathcal{R}}$ is equal to $\mathcal{F}_{\mathbb{I}}^{A} \equiv \mathcal{F}^A$, $\mathcal{F}_{\mathbb{I}}^{B} \equiv \mathcal{F}^B$, $\mathcal{F}_{\mathbb{I}}^{C} \equiv \mathcal{F}^C$, and $\mathcal{F}_{\mathbb{I}}^{D} \equiv \mathcal{F}^D$ when $\mathcal{R}$ contains respectively only one of the DMUs *A*, *B*, *C*, and *D*. The same holds for $\mathcal{F}_{\mathbb{U}}^{\mathcal{R}}$, i.e., $\mathcal{F}_{\mathbb{U}}^{A} \equiv \mathcal{F}^A \equiv \mathcal{F}_{\mathbb{I}}^{A}$, $\mathcal{F}_{\mathbb{U}}^{C} \equiv \mathcal{F}^B \equiv \mathcal{F}_{\mathbb{I}}^{B}$, $\mathcal{F}_{\mathbb{U}}^{C} \equiv \mathcal{F}^C \equiv \mathcal{F}_{\mathbb{I}}^{C}$, and $\mathcal{F}_{\mathbb{U}}^{D} \equiv \mathcal{F}^D \equiv \mathcal{F}_{\mathbb{I}}^{D}$. On the other hand, for $\mathcal{R}$ containing different combinations of the DMUs *A*, *B*, *C*, and *D*, set $\mathcal{F}_{\mathbb{I}}^{\mathcal{R}}$ is given as:

$$\mathcal{F}_{\mathbb{I}}^{AB} = \{\, H_2\},\ \ \mathcal{F}_{\mathbb{I}}^{AC} = \{\emptyset\},\ \ \mathcal{F}_{\mathbb{I}}^{AD} = \{\emptyset\},\ \ \mathcal{F}_{\mathbb{I}}^{BC} = \{\Omega_3\},\ \ \mathcal{F}_{\mathbb{I}}^{BD} = \{\emptyset\},\ \ \mathcal{F}_{\mathbb{I}}^{CD} = \{H_4\},$$
$$\mathcal{F}_{\mathbb{I}}^{ABC} = \{\emptyset\},\ \ \mathcal{F}_{\mathbb{I}}^{ABD} = \{\emptyset\},\ \ \mathcal{F}_{\mathbb{I}}^{ACD} = \{\emptyset\},\ \ \mathcal{F}_{\mathbb{I}}^{BCD} = \{\emptyset\},\ \ \mathcal{F}_{\mathbb{I}}^{ABCD} = \{\emptyset\}$$

and $\mathcal{F}_{\mathbb{U}}^{\mathcal{R}}$ is as follows:

$$\mathcal{F}_{\mathbb{U}}^{AB} = \{\, H_1, H_2, H_3\}, \mathcal{F}_{\mathbb{U}}^{AC} = \{H_1, H_2, H_3, H_4\}, \mathcal{F}_{\mathbb{U}}^{AD} = \{H_1, H_2, H_4, H_5\}, \mathcal{F}_{\mathbb{U}}^{BC} = \{H_2, H_3, H_4\},$$
$$\mathcal{F}_{\mathbb{U}}^{BD} = \{H_2, H_3, H_4, H_5\} \equiv\ \mathcal{F}_{\mathbb{U}}^{BCD},\ \ \mathcal{F}_{\mathbb{U}}^{CD} = \{H_3, H_4, H_5\},\ \ \mathcal{F}_{\mathbb{U}}^{ABC} = \{H_1, H_2, H_3, H_4\},$$
$$\mathcal{F}_{\mathbb{U}}^{ABD} = \{H_1, H_2, H_3, H_4, H_5\} \equiv \mathcal{F}_{\mathbb{U}}^{ACD} \equiv\ \mathcal{F}_{\mathbb{U}}^{ABCD} \equiv \mathcal{F}$$

where the capital letters in the superscripts correspond to the DMUs in the set $\mathcal{R}$.[107]

From these we see that: i) each of the sets $\mathcal{F}_{\mathbb{I}}^{\mathcal{R}}$ is a subset of the corresponding set $\mathcal{F}_{\mathbb{U}}^{\mathcal{R}}$, ii) $\mathcal{F}_{\mathbb{I}}^{AB}$,$\mathcal{F}_{\mathbb{I}}^{BC}$, and $\mathcal{F}_{\mathbb{I}}^{CD}$ are non-empty sets, since the DMUs in $\mathcal{R}$ jointly generate a DEA facet, while they contain only *one* vector of strictly positive input/output weights as in this case the number of DMUs in $\mathcal{R}$ is equal to $I + J - 1 = 2$, iii) the sets $\mathcal{F}_{\mathbb{I}}^{AC}$, $\mathcal{F}_{\mathbb{I}}^{AD}$, and $\mathcal{F}_{\mathbb{I}}^{BD}$ are empty as there is not a facet jointly generated by the DMUs in $\mathcal{R}$, while the corresponding $\mathcal{F}_{\mathbb{U}}^{\mathcal{R}}$ sets are non-empty, iv) $\mathcal{F}_{\mathbb{I}}^{ABC} \equiv \mathcal{F}_{\mathbb{I}}^{ABD} \equiv \mathcal{F}_{\mathbb{I}}^{ACD} \equiv \mathcal{F}_{\mathbb{I}}^{BCD} \equiv \mathcal{F}_{\mathbb{I}}^{ABCD} \equiv \{\emptyset\}$ since in this case he number of DMUs in $\mathcal{R}$ is equal to $3 > I + J - 1$, while the corresponding $\mathcal{F}_{\mathbb{U}}^{\mathcal{R}}$ sets are non-empty, and v) $\mathcal{F}_{\mathbb{U}}^{ABCD}$ is nonempty and is equivalent to $\mathcal{F}$ as in this case $\mathcal{R} \equiv \mathbb{E}$. Note that this also holds for sets $\mathcal{F}_{\mathbb{U}}^{ABD}$ and $\mathcal{F}_{\mathbb{U}}^{ACD}$, for which $\mathcal{R} \subset \mathbb{E}$.

[107] For instance, a superscript ABC refers to $\mathcal{R} = \{A, B, C\}$. Other capital letter superscripts appearing in the illustrative example are elaborated in a similar fashion.

7.3.2. *CR-DEA including preferences on model DMUs*

The fractional programming form of an input-oriented VRS CR-DEA model is given as:

$$
\begin{aligned}
\max_{\xi_j^o,\omega_i^o,\zeta^k} \quad & \left(\sum_{j=1}^{J} \xi_j^o y_j^o + \zeta^k\right) \Big/ \sum_{i=1}^{I} \omega_i^o x_i^o \\
s.t. \quad & \left(\sum_{j=1}^{J} \xi_j^o y_j^k + \zeta^k\right) \Big/ \sum_{i=1}^{I} \omega_i^o x_i^k \le 1 \quad \forall\, k \\
& \left(\xi_j^k, -\omega_i^k, \zeta^k\right) \in W \qquad \forall\, i,j
\end{aligned}
\tag{7.8}
$$

where $W \subseteq \mathbb{R}_+^J \times \mathbb{R}_-^I \times \mathbb{R}$ is a cone smaller than the non-negative orthant, which is essentially a collection of vectors that contains a *subset* of the weight vectors $\left(\xi_j^k, -\omega_i^k, \zeta^k\right) \in \mathcal{F}$ and their multiples. The CRS counterpart of (7.8) is obtained by removing the free variable ζ. The last constraint in (7.8) restricts the choice of optimal weights only among the weight vectors contained in W and their multiples, instead of the larger set $\mathcal{F}$, which is the case for the DEA model in (7.1). The model in (7.8) can be converted to the following linear program:

$$
\begin{aligned}
\max_{u_j^o,v_i^o z^k} \quad & \sum_{j=1}^{J} u_j^o y_j^o + u^k \\
s.t. \quad & \sum_{j=1}^{J} u_j^o y_j^k - \sum_{i=1}^{I} v_i^o x_i^k + u^k \le 0 \quad \forall\, k \\
& \sum_{i=1}^{I} v_i^o x_i^o = 1 \\
& \left(u_j^o, -v_i^o, z^k\right) \in \widetilde{W^o} \qquad \forall\, i,j
\end{aligned}
\tag{7.9}
$$

where $\widetilde{W^o}$ contains the multiples of the weight vectors in W that satisfy the normalizing equality for the evaluated DMU o in the second constraint in (7.8). The CRS counterpart of (7.9) is provided in Olesen and Petersen (2003) and can be obtained by removing the free variable u.

Let us assume that the DM wishes to compare each evaluated DMU with some chosen *model* efficient DMUs having exceptional performance and let the set $\mathcal{R}_{CR} \subseteq \mathbb{E}$ contain these *model* DMUs. In this case, W contains only weight vectors $\left(\xi_j^k, -\omega_i^k, \zeta^k\right)$

which are optimal for the *model* DMUs in the DEA model in (7.1) (Olesen and Petersen, 2003). Charnes *et al.* (1990) employed the CRS counterpart of (7.9) and used *one* among the --possibly multiple-- optimal vectors $(u_j^o, -v_i^o)$ for each of the *model* DMUs to define $\widetilde{W^o}$, namely the one resulting from the DEA model. In terms of the fractional model in (7.8), this means that *W* was defined by considering, for each DMU $k \in \mathcal{R}_{CR}$, one among the vectors $(\xi_j^k, -\omega_i^k)$ contained in $\mathcal{F}^k$, namely the one resulting from the CRS counterpart of (7.1). Each such weight vector is optimal for one *model* DMU but it is not necessarily optimal for other *model* DMUs as well, while it may or may not contain strictly positive components. This implies that the weight vectors contained in *W* are drawn from $\mathcal{F}_{\mathbb{U}}^{\mathcal{R}}$, namely the *union* of the sets containing the weight vectors that are optimal in the DEA model for each of the *model* DMUs. The cone in Charnes *et al.* (1990) was also specified as restricting the input and the output weights separately, namely $\widetilde{W^o}$ was in the form $\begin{bmatrix} U & 0 \\ 0 & V \end{bmatrix}$, where the cones $U \subseteq \mathbb{R}_+^J$, $V \subseteq \mathbb{R}_+^I$ were respectively defined using the optimal input and output weight sub-vectors.

Olesen and Petersen (2003, p. 329) argued that the intention of Charnes *et al.* (1990) was to generate (i.e., extend) "all facets containing" certain excellent DMUs. In this sense, they argue that *W* should be defined as a single cone using input/output weight vectors --instead of separate cones for input and output weight sub-vectors. As such, the boundary of the feasible set in the input/output space for models (7.8) and (7.9) will be a piecewise linear frontier defined by the extended efficient facets generated by the weight vectors contained in *W*. The authors note (p. 328) that *W* should account for all the multiple optimal weight vectors for each of the *model* DMUs, and that it should be "suitable for defining the boundary of (…) a frontier with well defined (i.e., with strictly positive and finite) rates of substitution and transformation", i.e., that it should contain vectors with strictly positive input and output weight components. However, they also point (p. 357) that vectors having zero components could also be included in *W*.[108] These ultimately imply that cone of feasible weight vectors in the CR-DEA model in (7.8) should be defined as the *union* of the sets containing the weight

[108] See their footnote 7 in which they note that weights normal to facets that are not of full dimension could also be included in *W*.

vectors that are optimal in the DEA model (7.1) for each of the *model* DMUs, namely $W \equiv \mathcal{F}_{\mathbb{U}}^{\mathcal{R}}$.

Alternatively, one could consider in W only those weight sets which are *jointly* optimal for all the chosen *model* DMUs, in which case W would be equal to the *intersection* of the sets containing the weight vectors that are optimal in the DEA model (7.1) for each of the *model* DMUs, namely $W \equiv \mathcal{F}_{\mathbb{I}}^{\mathcal{R}}$. We thus may define two variants of the CR-DEA model in (7.8). In the first, termed CR(𝕀)-DEA, the cone of feasible weight vectors weights is specified as the *intersection* of the sets containing the weight vectors that are optimal in DEA for each of the *model* DMUs, i.e., as:

$$\mathcal{F}_{CR(\mathbb{I})}^{\mathcal{R}} \equiv \mathcal{F}_{\mathbb{I}}^{\mathcal{R}} = \cap_{k \in \mathcal{R}} \mathcal{F}^{k} \tag{7.10}$$

The efficient frontier in CR(𝕀)-DEA will be the lower envelope of the extended DEA efficient facets (both FDEFs and non FDEFs) that are *jointly* generated by all the *model* DMUs. In the second variant, referred to as CR(𝕌)-DEA, the cone of feasible weights contains every weight vector that is optimal in the DEA model for at least one among the *model* DMUs, namely:

$$\mathcal{F}_{CR(\mathbb{U})}^{\mathcal{R}} \equiv \mathcal{F}_{\mathbb{U}}^{\mathcal{R}} = \cup_{k \in \mathcal{R}} \mathcal{F}^{k} \tag{7.11}$$

in which case the efficient frontier will be the lower envelope of the extended efficient facets (both FDEFs and non FDEFs) generated by at least one among the *model* DMUs.

The distinction between these two variants of model (7.8) is made, to the best of our knowledge, for the first time. Their main difference is that in CR(𝕀)-DEA, all the *model* DMUs will be rendered efficient when each DMU is evaluated, while in CR(𝕌)-DEA at least one among the *model* DMUs will be efficient, but not necessarily all of them. Thus, the *model* DMUs should jointly generate at least one efficient facet in order for CR(𝕀)-DEA to have feasible solutions (as otherwise $\mathcal{F}_{\mathbb{I}}^{\mathcal{R}} = \emptyset$), while this is not a necessity for CR(𝕌)-DEA. Irrespective of the choice of *model* DMUs, CR(𝕌)-DEA results in efficiency scores that are greater than or equal to those obtained from CR(𝕀)-DEA, since it holds that $\mathcal{F}_{CR(\mathbb{I})}^{\mathcal{R}} \subseteq \mathcal{F}_{CR(\mathbb{U})}^{\mathcal{R}}$. More specifically, each evaluated DMU in the CR(𝕀)-DEA model can choose only among weight vectors that are optimal --in the DEA model-- for *all* the *model* DMUs. In the CR(𝕌)-DEA model, each evaluated DMU can choose among both the weight vectors that are feasible in the

CR($\mathbb{I}$)-DEA model *as well as* those that are optimal for one *model* DMU but not necessarily for the others. Thus, it may receive a larger efficiency score in the CR($\mathbb{U}$)-DEA model compared to its corresponding CR($\mathbb{I}$)-DEA score. Furthermore, if the number of DMUs in $\mathcal{R}_{CR}$ is greater than $I+J-1$ (in CRS models) or $I+J$ (in VRS models) then $\mathcal{F}_{\mathbb{I}}^{\mathcal{R}}=\emptyset$ and $\mathcal{F}_{\mathbb{U}}^{\mathcal{R}}\neq\emptyset$, and in this case the CR($\mathbb{U}$)-DEA model will result in feasible solutions while the CR($\mathbb{I}$)-DEA model will not. Lastly, when there a single *model* DMU is considered, the two variants provide the same efficiency scores, as $\mathcal{F}_{CR(\mathbb{U})}^{k}\equiv\mathcal{F}_{CR(\mathbb{I})}^{k}\equiv\mathcal{F}^{k}$.

The distinction between CR($\mathbb{U}$)-DEA and CR($\mathbb{I}$)-DEA is also important for possible empirical applications of CR-DEA models. For a particular choice of *model* DMUs, CR($\mathbb{I}$)-DEA is a more restricted model than CR($\mathbb{U}$)-DEA. Thus, the former may be used when the DM opts for a more thorough performance assessment in which each DMU should be compared to all *model* DMUs, while the latter could be used in settings where comparison to one of the *model* DMUs is viewed as adequate enough. Alternatively, when more than one DMs (e.g., a council of stakeholders or an expert panel) choose the *model* DMUs, CR($\mathbb{I}$)-DEA could be used when there is unanimity in the choice of *model* DMUs, while CR($\mathbb{U}$)-DEA might be more preferrable when expert opinions on *model* DMUs diverge. Furthermore, CR($\mathbb{U}$)-DEA might be a more suitable modelling option compared to CR($\mathbb{I}$)-DEA when each of the chosen *model* DMUs represents a different type of "good performance". This could be the case, for instance, in an evaluation of public and private education institutions or an assessment of rural and urban bank branches. The use of CR($\mathbb{U}$)-DEA in such cases would not force the evaluated DMUs belonging to one type to be compared with the *model* DMU of the other type.

Regardless of the number of *model* DMUs, estimating CR($\mathbb{U}$)-DEA and CR($\mathbb{I}$)-DEA models requires to define cone *W*, namely to identify the weight vectors contained in $\mathcal{F}_{\mathbb{I}}^{\mathcal{R}}$ *or* $\mathcal{F}_{\mathbb{U}}^{\mathcal{R}}$, in a prior step. In particular, Portela and Thanassoulis (2006) note that once these weight vectors are identified, it is rather straightforward to obtain the efficiency scores of CR-DEA models by estimating, for each evaluated DMU, its efficiency score $\theta_k^o=\left(\sum_{j=1}^{J}\xi_j^k y_j^o+\zeta^k\right)/\sum_{i=1}^{I}\omega_i^k x_i^o$ using each of the identified weight vectors $\left(\xi_j^k,\omega_i^k,\zeta^k\right)\in W$, and then choosing the maximum among these efficiency scores. However, they argue that identifying the weight vectors is a difficult process, as it requires to identify all the facets of the DEA frontier and the weight vectors that are normal to each. This often involves the estimation of non-linear programs, such as those outlined in Olesen and Petersen (2003), Fukuyama and Sekitani (2012), and Davtalab-Olyaie *et al.* (2014). An alternative advocated by Olesen and Petersen (2003) and Thanassoulis *et al.* (2008) is to use the software program *Qhull*. It is however argued that this program is not developed exclusively for DEA and it is not always easy to use in DEA applications (see Aparicio *et al.*, 2007; Zhu *et al.*, 2022).

Another facet-extending variant of CR-DEA was presented in Olesen and Petersen (2015, pp. 167-168). This variant extends only FDEFs that are *jointly* generated by the *model* DMUs (hereafter referred to as EXFA-CR-DEA), provided that at least one such FDEF exists.[109] Assuming VRS, it is given as the following mixed-integer linear model:

$$
\begin{array}{lll}
\max\limits_{v_i^o, u_j^o, b_k^o, s_k^o, z^k} & \sum\limits_{j=1}^{J} u_j^o\, y_j^o + u^k & \\
s.t. & \sum\limits_{j=1}^{J} u_j^o\, y_j^k - \sum\limits_{i=1}^{I} v_i^o\, x_i^k + s_k^o + u^k = 0 & \forall\, k \in \mathbb{E} \\
& \sum\limits_{i=1}^{I} v_i^o\, x_i^o = 1 & \\
& s_k^o - M b_k^o \leq 0 & \forall\, k \in \mathbb{E} \\
& \sum\limits_{k \in \mathbb{E}} b_k^o \leq E - (I + J - 1) & \\
& b_k^o = \{0,1\} & \forall\, k \in \mathbb{E} \setminus \mathcal{R}_{CR} \\
& b_k^o = 0 & \forall\, k \in \mathcal{R}_{CR} \\
& s_k^o \geq 0 & \forall\, k \in \mathbb{E} \\
& v_i^o \geq \varepsilon & \forall\, i \\
& u_j^o \geq \varepsilon & \forall\, j \\
& u^k \;\; free &
\end{array}
\qquad (7.12)
$$

where E refers to the number of extreme efficient DMUs, b_k^o and s_k^o are additional variables to be estimated, and ε is a non-Archimedean number. The CRS counterpart

[109] This CR-DEA variant is related to the Extended Facet (EXFA) efficiency model developed by Olesen and Petersen (1996; 2015), for evaluating efficiency relative to a technology spanned by FDEFs and the extensions of these facets; see the relations (6.27) and (6.40), as footnote 21 in Olesen and Petersen (2015, pp. 167-170).

of (7.12) is provided in Olesen and Petersen (2015, p. 168) and can be obtained by removing the free variable u^k and changing the right-hand side in the fourth constraint to $E-(I+J-1)$.

In (7.12), the first set of constraints considers only the extreme efficient DMUs, as only these may generate an FDEF. Each binary variable corresponds to one extreme efficient DMU and is associated with a slack parameter s_k^o. When $b_k^o = 0$ and thus $s_k^o = 0$, the extreme efficient DMU generates the FDEF at which the evaluated DMU is radially projected. The fourth constraint in (7.12) is used to ensure than each DMU is evaluated against an FDEF. The binary variables corresponding to each of the *model* DMUs are set as equal to zero. This guarantees that all *model* DMUs should jointly generate the extended FDEF. As is evident, if the *model* DMUs do not generate at least on FDEF of the DEA frontier, (7.12) will not have a feasible solution.

When (7.12) reaches an optimal solution --provided that such a solution exists-- the constraint $\sum_{k\in\mathbb{E}} b_k^o \leq E-(I+J)$ will be satisfied as a strict equality and exactly $I+J$ relations in the first set of constraints will be strict equalities, which means that the evaluated DMU will be projected on an FDEF of the DEA efficient frontier or on its extension (Olesen and Petersen, 2003). Due to the sixth set of constraints in (7.12), this FDEF will be necessarily generated by all the *model* DMUs. The optimal weight vector for the evaluated DMU will be the multiple $\left(u_j^k, v_i^k, u^k\right) = \beta^k\left(\xi_j^k, \omega_i^k, \zeta^k\right), \beta^k = \left(\sum_{i=1}^{I} \omega_i^k x_i^o\right)^{-1}$ of the weight vector $(\xi_j^k, \omega_i^k, \zeta^k)$ that is normal to the (extended) FDEF at which the DMU is projected. This vector will contain only strictly positive input and output weight components, and hence each DMU will be assessed by means of well-defined MRSs and MRTs. Thus, in terms of the fractional weights $(\xi_j^k, \omega_i^k, \zeta^k)$, the cone of feasible weight vectors in (7.12) contains only the weight vectors with strictly positive components (i.e., those normal to FDEFs) that are jointly optimal in the DEA model (7.1) for all the *model* DMUs. It is thus a subset of $\mathcal{F}_{CR(\mathbb{I})}^{\mathcal{R}} \equiv \mathcal{F}_{\mathbb{I}}^{\mathcal{R}}$ and may be written as:

$$\mathcal{F}_{EXFA-CR}^{\mathcal{R}} = \left\{\left(\xi_j^k, \omega_i^k, \zeta^k\right) \in \mathcal{F}_{\mathbb{I}}^{\mathcal{R}}: \xi_j^k > 0,\ \ \forall\, j,\ \ \omega_i^k > 0, \forall\, i\right\} \tag{7.13}$$

which can be an empty set even if $\mathcal{F}_{\mathbb{I}}^{\mathcal{R}} \neq \emptyset$. This occurs when the chosen *model* DMUs jointly generate efficient facets of the DEA frontier but not one of full dimension. In this case, the CR($\mathbb{I}$)-DEA and CR($\mathbb{U}$)-DEA variants of (7.8) have feasible solutions, but

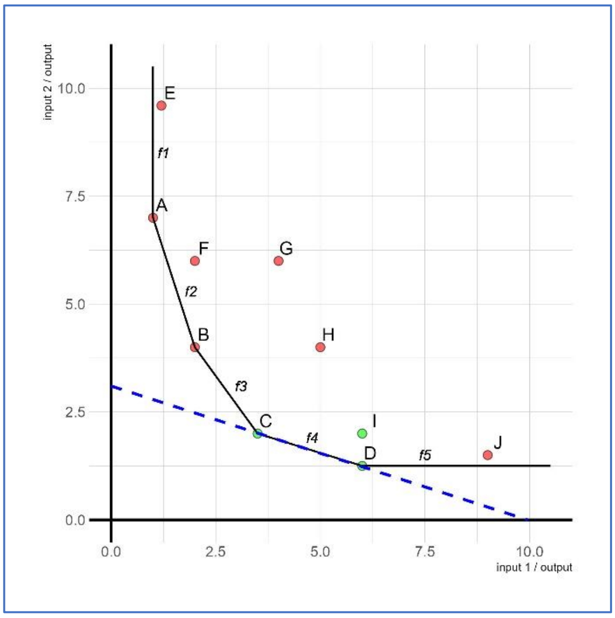
(7.12) does not. Following the suggestions in Portela and Thanassoulis (2006), the efficiency scores of the model in (7.12) could also be obtained in a similar fashion to those of the CR($\mathbb{I}$)-DEA and CR($\mathbb{U}$)-DEA models. One would first need to identify the weight vectors contained in $\mathcal{F}_{EXFA-CR}^{\mathcal{R}}$ and then estimate, for each evaluated DMU, its efficiency score using each of the weight vectors $\left(\xi_j^k, \omega_i^k, \zeta^k\right) \in \mathcal{F}_{FDEF-CR}^{\mathcal{R}}$. The maximum among these scores would be the EXFA-CR-DEA efficiency score.

In terms of our example in Figure 7.1, the three cone-ratio variants discussed above can be demonstrated as follows: when the DM chooses one model DMU among those extreme-efficient i.e., k is equal to either $A, B, C, or\ D$, then $\mathcal{F}_{CR(\mathbb{I})}^{k} \equiv \mathcal{F}_{CR(\mathbb{U})}^{k} \equiv \mathcal{F}^{k}$. If on the other hand there are more than one model DMUs, i.e., $\mathcal{R}_{CR}$ is equal to either $\{A,B\}, \{A,C\}, \{A,D\}, \{B,C\}, \{B,D\}, \{C,D\}, \{A,B,C\}, \{A,B,D\}, \{A,C,D\}, \{B,C,D\},$ or $\{A,B,C,D\}$, then the cones of feasible weight vectors for the CR($\mathbb{I}$)-DEA and the CR($\mathbb{U}$)-DEA models are given as:

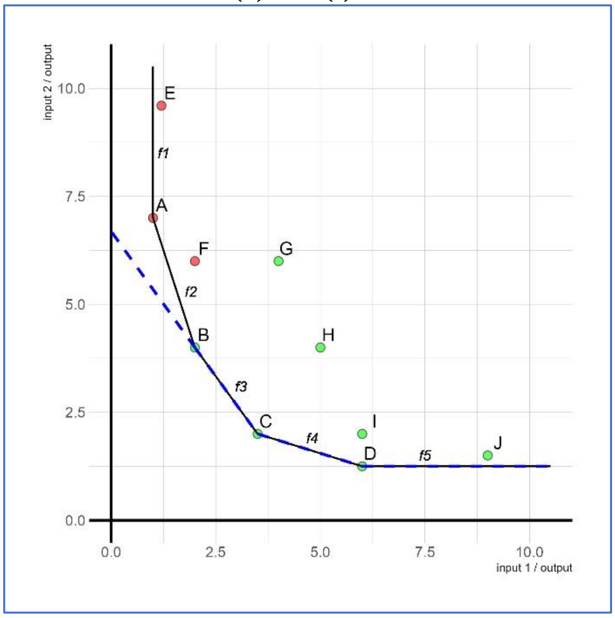
$$
\begin{aligned}
&\mathcal{F}_{CR(\mathbb{I})}^{AB} \equiv \mathcal{F}_{\mathbb{I}}^{AB} = \{H_2\} \subset \mathcal{F}_{CR(\mathbb{U})}^{AB} \equiv \mathcal{F}_{\mathbb{U}}^{AB} = \{H_1, H_2, H_3\}\\
&\mathcal{F}_{CR(\mathbb{I})}^{AC} \equiv \mathcal{F}_{\mathbb{I}}^{AC} = \emptyset \subset \mathcal{F}_{CR(\mathbb{U})}^{AC} \equiv \mathcal{F}_{\mathbb{U}}^{AC} = \{H_1, H_2, H_3, H_4\}\\
&\mathcal{F}_{CR(\mathbb{I})}^{AD} \equiv \mathcal{F}_{\mathbb{I}}^{AD} = \emptyset \subset \mathcal{F}_{CR(\mathbb{U})}^{AD} \equiv \mathcal{F}_{\mathbb{U}}^{AD} = \{H_1, H_2, H_4, H_5\}\\
&\mathcal{F}_{CR(\mathbb{I})}^{BC} \equiv \mathcal{F}_{\mathbb{I}}^{BC} = \{H_3\} \subset \mathcal{F}_{CR(\mathbb{U})}^{BC} \equiv \mathcal{F}_{\mathbb{U}}^{BC} = \{H_2, H_3, H_4\}\\
&\mathcal{F}_{CR(\mathbb{I})}^{BD} \equiv \mathcal{F}_{\mathbb{I}}^{BD} = \emptyset \subset \mathcal{F}_{CR(\mathbb{U})}^{BD} \equiv \mathcal{F}_{\mathbb{U}}^{BD} = \{H_2, H_3, H_4, H_5\}\\
&\mathcal{F}_{CR(\mathbb{I})}^{CD} \equiv \mathcal{F}_{\mathbb{I}}^{CD} = \{H_4\} \subset \mathcal{F}_{CR(\mathbb{U})}^{CD} \equiv \mathcal{F}_{\mathbb{U}}^{CD} = \{H_3, H_4, H_5\}\\
&\mathcal{F}_{CR(\mathbb{I})}^{ABC} \equiv \mathcal{F}_{\mathbb{I}}^{ABC} = \emptyset \subset \mathcal{F}_{CR(\mathbb{U})}^{ABC} \equiv \mathcal{F}_{\mathbb{U}}^{ABC} = \{H_1, H_2, H_3, H_4\}\\
&\mathcal{F}_{CR(\mathbb{I})}^{ABD} \equiv \mathcal{F}_{\mathbb{I}}^{ABD} = \emptyset \subset \mathcal{F}_{CR(\mathbb{U})}^{ABD} \equiv \mathcal{F}_{\mathbb{U}}^{ABD} = \{H_1, H_2, H_3, H_4, H_5\} \equiv \mathcal{F}\\
&\mathcal{F}_{CR(\mathbb{I})}^{ACD} \equiv \mathcal{F}_{\mathbb{I}}^{ACD} = \emptyset \subset \mathcal{F}_{CR(\mathbb{U})}^{ACD} \equiv \mathcal{F}_{\mathbb{U}}^{ACD} = \{H_1, H_2, H_3, H_4, H_5\} \equiv \mathcal{F}\\
&\mathcal{F}_{CR(\mathbb{I})}^{BCD} \equiv \mathcal{F}_{\mathbb{I}}^{BCD} = \emptyset \subset \mathcal{F}_{CR(\mathbb{U})}^{BCD} \equiv \mathcal{F}_{\mathbb{U}}^{BCD} = \{H_2, H_3, H_4, H_5\}\\
&\mathcal{F}_{CR(\mathbb{I})}^{ABCD} \equiv \mathcal{F}_{\mathbb{I}}^{ABCD} = \emptyset \subset \mathcal{F}_{CR(\mathbb{U})}^{ABCD} \equiv \mathcal{F}_{\mathbb{U}}^{ABCD} = \{H_1, H_2, H_3, H_4, H_5\} \equiv \mathcal{F}
\end{aligned}
$$

from which we can see that $\mathcal{F}_{CR(\mathbb{I})}^{\mathcal{R}}$ is always a subset of $\mathcal{F}_{CR(\mathbb{U})}^{\mathcal{R}}$ and thus the CR($\mathbb{U}$)-DEA efficiency scores will be greater than or equal to those of the CR($\mathbb{I}$) DEA model. For example, when $\mathcal{R}_{CR} = \{C, D\}$, then in the CR($\mathbb{I}$)-DEA model each DMU is evaluated based only on the weight vector H_4 that is normal to facet *f4*. The frontier against which the DMUs are evaluated is thus the extended facet *f4*, which is portrayed as the orange dashed line in Figure 7.2a. In that, DMUs for which at least one optimal weight vector is Ω_4, i.e., DMUs *C*, *D*, and *I*, are indicated with a green color. The remaining DMUs appear as red-colored points, meaning that they will be assigned an efficiency score lower than their corresponding DEA one. On the other hand, in the

Figure 7.2: Efficient frontiers for CR(I)-DEA and CR(U)-DEA models (model DMUs: C and D)



(a) CR(I)-DEA



(b) CR(U)-DEA

corresponding CR(𝕌)-DEA model, each DMU can choose among vectors H_3, H_4 *and* H_5 the one that maximizes its efficiency score. Thus, each DMU can secure at least an efficiency score equal to that obtained from the CR(𝕀)-DEA model, while it may also attain a larger score using one of the vectors H_3 *and* H_5. The frontier against which the DMUs are evaluated is the lower envelope of the extended facets *f3*, *f4*, and *f5* which is portrayed as the orange dashed line in Figure 7.2b. In this case only DMUs *A*, *E*, and *F* are indicated with a red color, in the sense that none of their optimal weight vectors are contained in $\mathcal{F}_{CR(\mathbb{U})}^{CD}$. Also, the CR($\mathbb{U}$)-DEA model results in feasible solutions irrespective of the choice of model DMUs, while this is not the case for the CR($\mathbb{I}$)-DEA model when $\mathcal{R}_{CR}$ contains DMUs that do not jointly generate a facet of the DEA frontier. This is the case, for example, when $\mathcal{R}_{CR} = \{A, D\}$. The associated CR($\mathbb{U}$)-DEA frontier is given in Figure 7.3 as the lower envelope of the extended facets *f1, f2, f4,* and *f5.* The inefficient DMU *H* will be assigned a score lower to that of the DEA model, since $\mathcal{F}_{CR(\mathbb{U})}^{AD}$ does not include the weight vector Ω_4.

The sets containing the feasible weight vectors in the EXFA-CR-DEA model for alternative choices of the *model* DMUs in our example are given as:

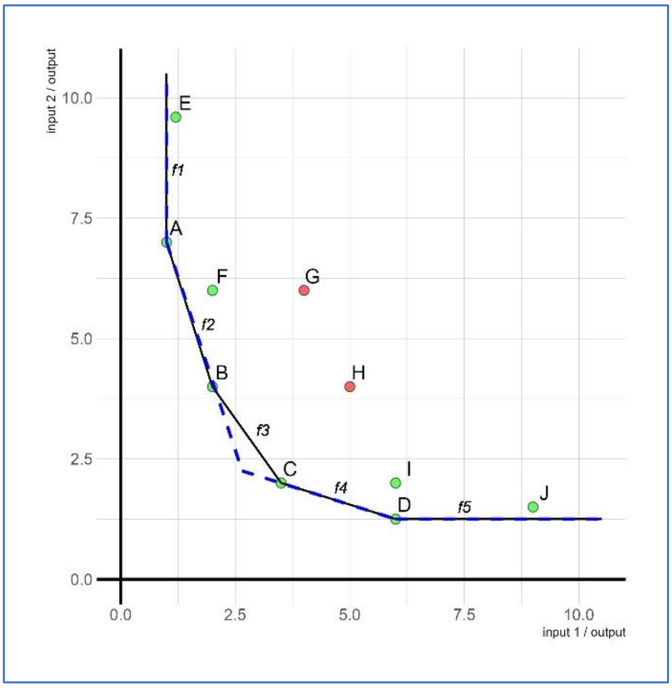
$$\begin{aligned}
&\mathcal{F}_{EXFA-CR}^{A} = \{H_2\},\ \mathcal{F}_{EXFA-CR}^{B} = \{H_2, H_3\},\ \mathcal{F}_{EXFA-CR}^{C} = \{H_3, H_4\},\ \mathcal{F}_{EXFA-CR}^{D} = \{H_4\},\\
&\mathcal{F}_{EXFA-CR}^{AB} = \{H_2\}, \mathcal{F}_{EXFA-CR}^{AC} = \emptyset, \mathcal{F}_{EXFA-CR}^{AD} = \emptyset, \mathcal{F}_{EXFA-CR}^{BC} = \{\Omega_3\}, \mathcal{F}_{EXFA-CR}^{BD} = \emptyset,\\
&\mathcal{F}_{EXFA-CR}^{CD} = \{H_4\}, \mathcal{F}_{EXFA-CR}^{ABC} \equiv \mathcal{F}_{EXFA-CR}^{ABD} \equiv \mathcal{F}_{EXFA-CR}^{ACD} \equiv \mathcal{F}_{EXFA-CR}^{BCD} \equiv \mathcal{F}_{EXFA-CR}^{ABCD} = \emptyset
\end{aligned}$$

When non-empty, each of these sets consists only of weight vectors normal to FDEFs and is a subset of the corresponding $\mathcal{F}_{\mathbb{I}}^{\mathcal{R}}$ set. For example, when DMU A is the *model* DMU, $\mathcal{F}_{CR(\mathbb{I})}^{A}$ contains both vectors H_1 and H_2, while $\mathcal{F}_{EXFA-CR}^{A}$ contains only the latter, since the former is not of full dimension. The efficient frontiers defined by the two models are given in Figures 7.4a and 7.4b respectively. The CR($\mathbb{I}$)-DEA frontier -- which in this case coincides with the CR($\mathbb{U}$)-DEA frontier since there is one *model* DMU-- is the lower envelope of the extended facets *f1* and *f2*. In this case, the inefficient DMU *E* for which Ω_1 is the only optimal weight vector in DEA, receives a score equal to its DEA one. The EXFA-CR-DEA frontier consists only of the extended facet *f2*, and DMU *E* is now marked with red color since it is assigned an efficiency score lower than that of the DEA model.

7.3.3. Value Efficiency Analysis

In VEA, a DM expresses his/her preferences over the desirable input/output structure of DMUs by choosing a non-dominated (i.e., efficient) point on the DEA frontier that constitutes the MPS (Halme *et al.*, 1999). This point is assumed to maximize the DM's implicitly known value function and will in essence be either a single extreme-efficient t DMU or a combination of extreme-efficient DMUs that are *jointly* efficient, in the sence that sense that they *jointly* generate at least one facet of the DEA efficient frontier.

Figure 7.3: Efficient frontier for the CR(𝕌)-DEA model (model DMUs: A and D)



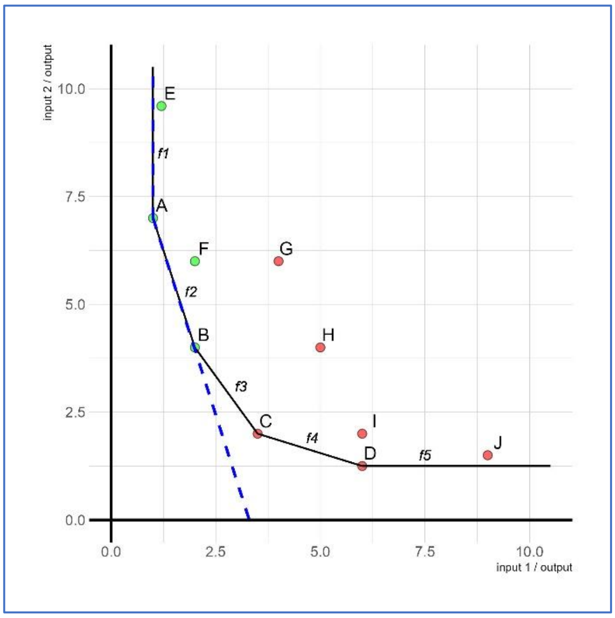
Let set $\mathcal{R}_V \subseteq \mathbb{E}$ contain the DMUs comprising the MPS.[110] The fractional programming form of an input-oriented CRS VEA model for the o^{th} DMU is given as:

$$
\begin{aligned}
&\max_{\xi_j^o,\omega_i^o,\zeta^k} \left(\sum_{j=1}^{J} \xi_j^o y_j^o + \zeta^k\right) \Big/ \sum_{i=1}^{I} \omega_i^o x_i^o \\
&s.t. \quad \left(\sum_{j=1}^{J} \xi_j^o y_j^k + \zeta^k\right) \Big/ \sum_{i=1}^{I} \omega_i^o x_i^k \le 1 \quad \forall\, k \notin \mathcal{R}_V \\
&\qquad \left(\sum_{j=1}^{J} \xi_j^o y_j^k + \zeta^k\right) \Big/ \sum_{i=1}^{I} \omega_i^o x_i^k = 1 \quad \forall\, k \in \mathcal{R}_V \\
&\qquad \xi_j^o \ge 0 \qquad \forall\, j \\
&\qquad \omega_i^o \ge 0 \qquad \forall\, i \\
&\qquad \zeta^k \ free
\end{aligned}
\tag{7.14}
$$

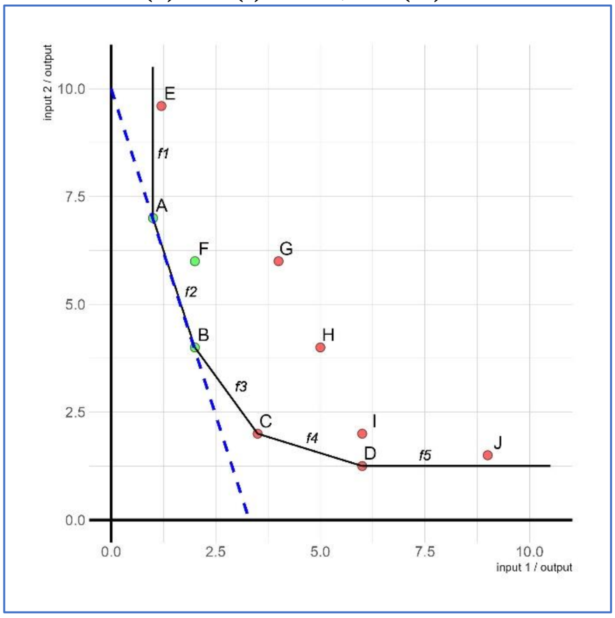
which differs from the DEA model in (7.1) in that the inequality constraints associated with the DMUs constituting the MPS are turned to strict equalities. This means that, when each DMU is evaluated, the optimal weight vector $\left(\xi_j^o, -\omega_i^o, \zeta^k\right)$ resulting from

[110] As in most cases not all the DMUs in $\mathbb{E}$ are jointly efficient with each other, it frequently holds that $\mathcal{R}^V \subset \mathbb{E}$. However, in rare cases there is a facet jointly generated by all the DMUs in $\mathbb{E}$, in which case $\mathcal{R}^V \equiv \mathbb{E}$ can be a valid choice.

Figure 7.4: Efficient frontier for different Cone-ratio DEA models (model DMU: A)



(a) CR(I)-DEA, CR(U)-DEA



(b) EXFA-CR-DEA

(7.14) should be such that *all* the DMUs constituting the MPS are rendered efficient. Thus, the polyhedral cone of feasible weight vectors in (7.14) contains only those which are optimal in model (7.1) for *all* the DMUs comprising the MPS, namely:

$$\mathcal{F}_{V}^{\mathcal{R}} \equiv \mathcal{F}_{\mathbb{I}}^{\mathcal{R}} \tag{7.15}$$

This follows from the fact that the second set of constraints in (7.14) is also the set of constraints that define the set $\mathcal{F}_{\mathbb{I}}^{\mathcal{R}}$ in (7.4). In the case where a single DMU $k \in \mathbb{E}$ is the MPS, $\mathcal{F}_V^k$ coincides with $\mathcal{F}^k$.

The weight vectors $\left(\xi_j^k, -\omega_i^k, \zeta^k\right)$ contained in $\mathcal{F}_V^{\mathcal{R}}$ are normal to the DEA facets (both FDEFs and non FDEFs) intercepting at the MPS, and the VEA frontier is constructed as the intersection (i.e., the lower envelope) of these extended facets. As $\mathcal{F}_V^{\mathcal{R}}$ may contain more than one weight vectors, (7.14) gives to each evaluated DMU the benefit-of-the doubt to choose among them the one that maximizes its efficiency score. Essentially, the VEA efficiency scores can be obtained in a similar process as that outlined by Portela and Thanassoulis for the CR-DEA models: *First*, estimate (7.1) and identify all the DMUs in $\mathbb{E}$ and the weight vectors that are optimal for each of them (i.e., obtain set $\mathcal{F}^k \ \forall \ k \in \mathbb{E}$). *Second*, define the set $\mathcal{R}$ containing the DMUs that comprise the MPS and obtain the intersection $\mathcal{F}_{\mathbb{I}}^{\mathcal{R}} \equiv \mathcal{F}_V^{\mathcal{R}}$ of their optimal weight vectors. Notice that the DMUs comprising the MPS should be chosen such that $\mathcal{F}_{\mathbb{I}}^{\mathcal{R}} \equiv \mathcal{F}_V^{\mathcal{R}} \neq \emptyset$. *Third*, for each evaluated DMU, estimate its efficiency score $\theta_k^o = \left(\sum_{j=1}^{J} \xi_j^k y_j^o + \zeta^k\right) / \sum_{i=1}^{I} \omega_i^k x_i^o$ using each of the weight vectors $\left(\xi_j^k, \omega_i^k, \zeta^k\right) \in \mathcal{F}_{\mathbb{I}}^{\mathcal{R}}$. The maximum among those efficiency scores will be the VEA score. If the vector of optimal weights for the evaluated DMU in the DEA model in (7.1) is contained in (7.15), then the DMU's VEA efficiency score is equal to its corresponding DEA one. Otherwise, the VEA score will be lower than that of the DEA model. Also, when the number of DMUs in $\mathcal{R}_V$ is equal to $I + J - 1$ (in CRS models) or to $I + J$ (in VRS models), then $\mathcal{F}_V^{\mathcal{R}}$ will be a non-empty set and (7.14) will have a feasible solution only if the chosen DMUs in $\mathcal{R}_V$ jointly generate an FDEF. In this case, all DMUs will be evaluated based on a common vector of strictly positive weights, namely the vector normal to the FDEF generated by the DMUs in $\mathcal{R}_V$.

The model in (7.14) can be converted to the following linear model:

$$
\begin{aligned}
\max_{u_j^o, v_i^o} \quad & \sum_{j=1}^{J} u_j^o y_j^o + u^k \\
s.t. \quad & \sum_{j=1}^{J} u_j^o y_j^k - \sum_{i=1}^{I} v_i^o x_i^k + u^k \le 0 \quad \forall\, k \notin \mathcal{R}_V \\
& \sum_{j=1}^{J} u_j^o y_j^k - \sum_{i=1}^{I} v_i^o x_i^k + u^k = 0 \quad \forall\, k \in \mathcal{R}_V \\
& \sum_{i=1}^{I} v_i^o x_i^o = 1 \\
& u_j^o \ge 0 \qquad\qquad\qquad\qquad \forall\, j \\
& v_i^o \ge 0 \qquad\qquad\qquad\qquad \forall\, i \\
& u^k \; free
\end{aligned}
\tag{7.16}
$$

Then, the polyhedral cone containing the feasible weight vectors in (7.16) for the evaluated DMU is given as:

$$
\tilde{\mathcal{F}}_V^{o,\mathcal{R}} = \left\{ \left(u_j^k, v_i^k, u^k\right) = \beta^o\left(\xi_j^k, \omega_i^k, \zeta^k\right) : \left(\xi_j^k, \omega_i^k, \zeta^k\right) \in \mathcal{F}_V^{\mathcal{R}},\ \beta^o = \left(\sum_{i=1}^{I} \omega_i^k x_i^o\right)^{-1} \right\} \tag{7.17}
$$

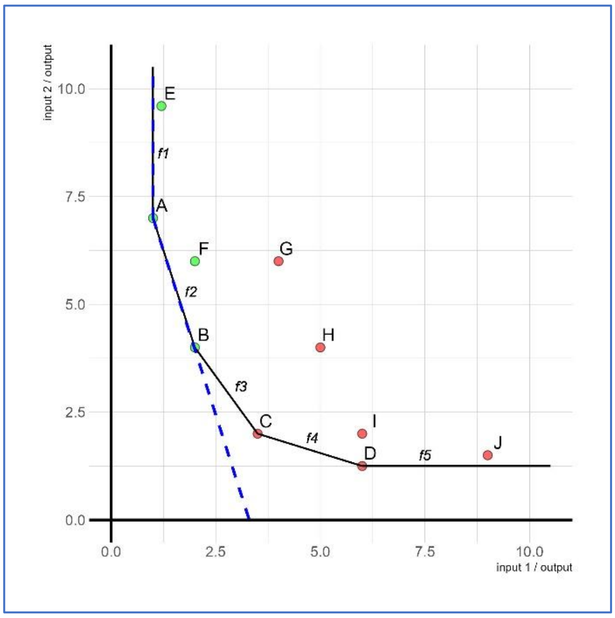
where the first superscript in $\tilde{\mathcal{F}}_V^{o,\mathcal{R}}$ refers to the evaluated DMU and the second to the set of the DMUs comprising the MPS. The set $\tilde{\mathcal{F}}_V^{o,\mathcal{R}}$ contains the multiples of the weight sets in $\mathcal{F}_V^{\mathcal{R}}$ that satisfy the normalizing equality in the second constraint in (7.16) for the evaluated DMU.

Using the example in Table 7.1 to demonstrate the VEA model, notice that the MPS can be one of the DMUs *A*, *B*, *C*, and *D*, be a combination of DMUs A and B, a combination of DMUs B and C, or a combination of DMUs C and D, since only these pairs of DMUs jointly generate a DEA facet. Let's assume, without loss of generality, that the DM chooses DMU *A* as the MPS. Then, each evaluated DMU in the VEA model can choose only among the weight vectors H_1 and H_2 which render DMU *A* efficient. Thus, we have:

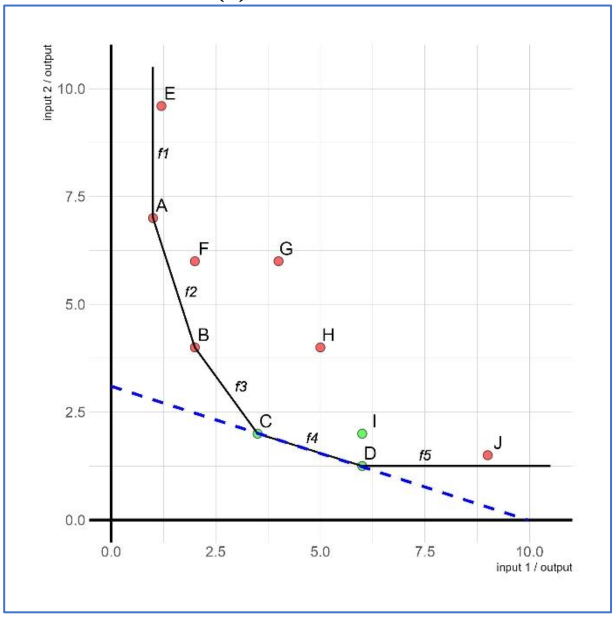
$$
\mathcal{F}_V^A \equiv \mathcal{F}_{\mathbb{I}}^A \equiv \mathcal{F}^A = \{H_1, H_2\}
$$

and the VEA frontier is formed by extending facets *f1* and *f2* towards the axes (see Figure 7.5a). The VEA scores for the efficient DMU *B* and the inefficient DMU *E* will be equivalent to their corresponding DEA ones, since at least one among the weight

Figure 7.5: VEA efficient frontier for different MPS specifications



(c) MPS: DMU A



(d) MPS: DMUs C and D

vectors that are optimal for these DMUs in DEA are contained in $\mathcal{F}_V^A$. On the other hand, this does not occur for the remaining DMUs, which are indicated by a red color in Figure 5a. The VEA scores for these DMUs will be lower than those of the DEA model, meaning that DMUs *C* and *D* drop from the efficiency list.

In a similar manner we can define:

$$\mathcal{F}_V^B \equiv \mathcal{F}_{\mathbb{I}}^B \equiv \mathcal{F}^B = \{H_2, H_3\},\ \ \mathcal{F}_V^C \equiv \mathcal{F}_{\mathbb{I}}^C \equiv \mathcal{F}^C = \{H_3, H_4\},\ \ \mathcal{F}_V^D \equiv \mathcal{F}_{\mathbb{I}}^D \equiv \mathcal{F}^D = \{H_4, H_5\}$$
$$\mathcal{F}_V^{AB} \equiv \mathcal{F}_{\mathbb{I}}^{AB} = \{H_2\},\ \ \mathcal{F}_V^{BC} \equiv \mathcal{F}_{\mathbb{I}}^{BC} = \{H_3\},\ \ \mathcal{F}_V^{CD} \equiv \mathcal{F}_{\mathbb{I}}^{CD} = \{H_4\}$$

for the remaining options the DM has for choosing the MPS. From that we can see that $\mathcal{F}_V^{\mathcal{R}}$ contains only a single vector of strictly positive weights, i.e., H_2 or H_3, when the MPS is a combination of two efficient DMUs, since in this case the number of DMUs in $\mathcal{R}_V$ is equal to $I + J - 1 = 2$. The case where the DM chooses DMUs *C* and *D* as the MPS is portrayed in Figure 7.5b, in which where we see that, apart from DMUs *C* and *D*, only DMU *I* has an optimal weight vector contained in $\mathcal{F}_V^{CD}$.

7.4. Main results

We may now relate the efficiency scores provided by the CR($\mathbb{I}$)-DEA, the CR($\mathbb{U}$)-DEA, and the EXFA-CR-DEA model to those of the VEA model, when the set of the model DMUs in CR-DEA coincides with that of the DMUs comprising the MPS in VEA, i.e., when $\mathcal{R}_{CR} \equiv \mathcal{R}_V \equiv \mathcal{R} \subseteq \mathbb{E}$. In such a case, the four models have the same objective function but different feasible regions, which are however subsets of one another. We consider between two cases, namely the case of a single *model* DMU and the case of multiple *model* DMUs, and we demonstrate these relations through our illustrative example in Table 1.

7.4.1. Single model DMU

In this case, it is evident from relations (7.3), (7.4) and (7.5) that $\mathcal{F}_{\mathbb{U}}^k \equiv \mathcal{F}^k \equiv \mathcal{F}_{\mathbb{I}}^k$ and thus, the CR($\mathbb{I}$)-DEA and CR($\mathbb{U}$)-DEA models provide the same efficiency scores. Let us assume, for example, that DMU *A* is chosen as the *model* DMU in CR-DEA and as the MPS in VEA. Then, the sets of feasible weight vectors in the three variants of the CR-DEA model and the VEA model are related as follows:

$$\mathcal{F}_{EXFA-CR}^A = \{H_2\} \subset \mathcal{F}_{CR(\mathbb{I})}^A \equiv \mathcal{F}_{\mathbb{I}}^A \equiv \mathcal{F}^A \equiv \mathcal{F}_V^A \equiv \mathcal{F}_{CR(\mathbb{U})}^A \equiv \mathcal{F}_{\mathbb{U}}^A = \{H_1, H_2\}$$

from which we deduce that the same efficiency scores will be obtained from the CR-DEA (CR($\mathbb{I}$)-DEA or CR($\mathbb{U}$)-DEA) and VEA models. These scores will be greater than or equal to those obtained from the EXFA-CR-DEA model, since in that all DMUs are evaluated based only on the weight vector H_2, while they may or may not secure a larger efficiency score based on the weight vector H_3. The same is the case when DMU D is the *model* DMU or the MPS: $\mathcal{F}_{EXFA-CR}^D$ contains only the weight vector H_4, while the sets of feasible weight vectors for the other three models contain also H_5. On the other hand, when $\mathcal{R} = \{B\}$, the sets of feasible weight vectors in the three variants of the CR-DEA model and the VEA model are related as follows:

$$\mathcal{F}^{B}_{EXFA-CR} = \{ H_2, H_3\} \equiv \mathcal{F}^{B}_{CR(\mathbb{I})} \equiv \mathcal{F}^{B}_{\mathbb{I}} \equiv \mathcal{F}^{B} \equiv \mathcal{F}^{B}_{V} \equiv \mathcal{F}^{B}_{CR(\mathbb{U})} \equiv \mathcal{F}^{B}_{\mathbb{U}}$$

In this case we see that the set of feasible weight vectors for all four models coincide with each other, and the same efficiency scores will be obtained from each model. The same is true when DMU C is the model DMU or the MPS. Thus, when a single DMU $k \in \mathbb{E}$ is chosen as the *model* DMU or as the MPS, the relation between the sets of feasible weights in the CR-DEA models and the VEA model is given as:

$$\mathcal{F}^{k}_{EXFA-CR} \subseteq \mathcal{F}^{k}_{CR(\mathbb{I})} \equiv \mathcal{F}^{k}_{\mathbb{I}} \equiv \mathcal{F}^{k} \equiv \mathcal{F}^{k}_{V} \equiv \mathcal{F}^{k}_{CR(\mathbb{U})} \equiv \mathcal{F}^{k}_{\mathbb{U}} \tag{7.18}$$

and the relation between the efficiency scores obtained, for each evaluated DMU, from these models is given as:

$$\theta^{o,k}_{EXFA-CR} \leq \theta^{o,k}_{CR(\mathbb{I})} = \theta^{o,k}_{V} = \theta^{o,k}_{CR(\mathbb{U})} \tag{7.19}$$

where θ corresponds to the relevant efficiency scores. Thus, we have:

Proposition *7.1 (Olesen and Petersen, 2003):* Given a single *model* DMU or MPS, a VEA model provides equal efficiency scores to a CR-DEA model in which the set of feasible weight vectors is determined by those that are optimal in DEA for the *model* DMU

which confirms the first inference in Olesen and Petersen (2003) about the relation of VEA and CR-DEA models. Relation (7.18) is also deduced by comparing relations (7.10), (7.11), (7.13) and (7.15) when only one DMU is included in $\mathcal{R}$.

Notice that in the general case given by relation (7.18) the *model* DMU may not generate at least one FDEF of the DEA frontier. In this case $\mathcal{F}^{k}_{EXFA-CR} = \emptyset$ and the EXFA-CR-DEA model will not provide a feasible value for $\theta^{o,k}_{EXFA-CR}$. However, in the case where the *model* DMU generates at least one FDEF of the DEA frontier, then we have:

Proposition *7.2:* Given a single *model* DMU that generates at least one FDEF, the efficiency scores of a VEA model in which this DMU is the MPS are greater than or equal to the scores of a CR-DEA model in which the set of feasible weight vectors is given by those vectors with strictly positive input/output weight components, that are optimal in DEA for the *model* DMU.

This establishes the relation between the scores of the VEA and the EXFA-CR-DEA models for the case of a single *model* DMU.

7.4.2. Multiple model DMUs

In this case one needs to distinguish between CR-DEA models, in which the set of feasible weight vectors is given by the intersection (i.e., CR($\mathbb{I}$)-DEA) or the union (CR($\mathbb{U}$)-DEA) of the sets of weights that are optimal in DEA for each of the *model* DMUs. This is because in this case it holds that $\mathcal{F}^{\mathcal{R}}_{CR(\mathbb{I})} \subseteq \mathcal{F}^{\mathcal{R}}_{CR(\mathbb{U})}$ and thus the CR($\mathbb{U}$)-DEA model results in efficiency scores that are greater than or equal to those obtained from CR($\mathbb{I}$)-DEA. In particular, from the relations in (7.10), (7.11), (7.13) and (7.15), we obtain the following relation between the sets of feasible weights for the three CR-DEA variants and VEA:

$$\mathcal{F}^{\mathcal{R}}_{EXFA-CR} \subseteq \mathcal{F}^{\mathcal{R}}_{CR(\mathbb{I})} \equiv \mathcal{F}^{\mathcal{R}}_{\mathbb{I}} \equiv \mathcal{F}^{\mathcal{R}}_{V} \subseteq \mathcal{F}^{\mathcal{R}}_{CR(\mathbb{U})} \equiv \mathcal{F}^{\mathcal{R}}_{\mathbb{U}} \tag{7.20}$$

This leads to the following:

$$\theta^{o,\mathcal{R}}_{EXFA-CR} \leq \theta^{o,\mathcal{R}}_{CR(\mathbb{I})} = \theta^{o,\mathcal{R}}_{V} \leq \theta^{o,\mathcal{R}}_{CR(\mathbb{U})} \tag{7.21}$$

which establishes the relations between the efficiency scores obtained, for each evaluated DMU, from the three CR-DEA variants and VEA for the case of multiple *model* DMUs.

In particular, considering first the CR($\mathbb{U}$)-DEA and VEA models, we see from the relation in (7.21) that the VEA efficiency scores are lower than or equal to those provided by the model provides equal efficiency scores to the CR($\mathbb{U}$)-DEA model. To demonstrate this, consider for example, that $\mathcal{R} = \{A, B\}$, then the sets of feasible weight vectors for the evaluated DMUs in the CR($\mathbb{U}$)-DEA and the VEA models are related as follows:

$$\mathcal{F}^{AB}_{V} = \{H_2\} \subset \mathcal{F}^{AB}_{CR(\mathbb{U})} \equiv \mathcal{F}^{AB}_{\mathbb{U}} = \{H_1, H_2, H_3\}$$

in which case the DMUs may secure a larger efficiency score in the CR($\mathbb{U}$)-DEA model based on one of the weight vectors H_1 and H_3 compared to H_2, which is the only choice in the VEA model. The same is true when $\mathcal{R} = \{B, C\}$ and $\mathcal{R} = \{C, D\}$, since these sets contain jointly efficient DMUs. A similar relation cannot be provided when $\mathcal{R}$ is equal to either $\{A, C\}, \{A, D\}$, or $\{B, D\}$, since the DMUs in these sets do not jointly generate a DEA facet. In these cases, only the CR($\mathbb{U}$)-DEA model provides feasible solutions. Thus, we have:

PROPOSITION 7.3: Given a set of *model* DMUs that are jointly efficient, a VEA model in which these DMUs comprise the MPS provides a lower bound for the efficiency score of a CR-DEA model in which the set of feasible weights is given as the *union* of the sets containing the weight vectors that are optimal in the DEA model for each of the *model* DMUs

We now show that the inference in Olesen and Petersen (2003) about the relation of CR-DEA and VEA models for the case of multiple *model* DMUs, is true. More specifically, when each of the extreme-efficient DMUs *A*, *B*, *C*, and *D* in our illustrative example is the MPS, the sets of feasible weight vectors in the VEA model are given as $\mathcal{F}_V^A = \{H_1, H_2\}$, $\mathcal{F}_V^B = \{H_2, H_3\}$, $\mathcal{F}_V^C = \{H_3, H_4\}$ and $\mathcal{F}_V^D = \{H_4, H_5\}$ respectively. Furthermore, when DMUs *A* and *B* are chosen as *model* DMUs in a CR($\mathbb{U}$)-DEA model, the set of feasible weight vectors is given as $\mathcal{F}_{CR(\mathbb{U})}^{AB} \equiv \mathcal{F}_{\mathbb{U}}^{AB} = \{H_1, H_2, H_3\}$. From these we obtain the following relation:

$$\mathcal{F}_{CR(\mathbb{U})}^{AB} \equiv \mathcal{F}_{\mathbb{U}}^{AB} = \{H_1, H_2, H_3\} = \mathcal{F}_V^A \cup \mathcal{F}_V^B$$

namely that the set of feasible weight vectors for the CR($\mathbb{U}$)-DEA model is the *union* of sets containing the feasible weight vectors in the two VEA models in which DMUs *A* and *B* are respectively the MPS. A similar relation can be deduced when any other combination of the DMUs *A*, *B*, *C*, and *D* is chosen as *model* DMUs in a CR($\mathbb{U}$)-DEA model. Thus, we have:

$$\mathcal{F}_{CR(\mathbb{U})}^{\mathcal{R}} = \bigcup_{k \in \mathcal{R}} \mathcal{F}_V^k \tag{7.22}$$

namely that the set of feasible weight vectors for the CR($\mathbb{U}$)-DEA model given a set $\mathcal{R}$ of *model* DMUs is the *union* of sets containing the feasible optimal weight vectors in every VEA model in which *one* of the DMUs in set $\mathcal{R}$ is used as the MPS. This means that the efficiency score of the CR($\mathbb{U}$)-DEA model can be obtained as follows:

$$\theta_{CR(\mathbb{U})}^{o,\mathcal{R}} = \max_{k \in \mathcal{R}} \left\{ \theta_V^{o,k} \right\} \tag{7.23}$$

Thus, we have:

PROPOSITION 7.4 (Olesen and Petersen, 2003): Given a set $\mathcal{R}$ of *model* DMUs, the efficiency scores of a CR-DEA model in which the set of feasible weights is the *union* of the sets containing the weight vectors that are optimal in the DEA model for each of the *model* DMUs can be obtained, for each evaluated DMU, as the maximum among the efficiency scores obtained from a number of VEA models equal to the number of DMUs in $\mathcal{R}$, each of which uses a single *model* DMU $k \in \mathcal{R}$ as the MPS

The relation in (7.23) can be rewritten, using relation (7.19), as follows:

$$\theta_{CR(\mathbb{U})}^{o,\mathcal{R}} = \max_{k \in \mathcal{R}} \left\{ \theta_{CR(\mathbb{I})}^{o,k} \right\} = \max_{k \in \mathcal{R}} \left\{ \theta_{CR(\mathbb{U})}^{o,k} \right\} \tag{7.24}$$

which demonstrates that, in the case of multiple *model* DMUs, the efficiency scores of the CR($\mathbb{U}$)-DEA can also be obtained as the maximum among the efficiency scores of different CR-DEA models using one of the DMUs in set $\mathcal{R}$ as the model DMU. We thus have:

PROPOSITION 7.5: Given a set $\mathcal{R}$ of *model* DMUs, the efficiency scores of a CR-DEA model in which the set of feasible weight vectors is defined as the *union* of the sets containing the weight vectors that are optimal in the DEA model for each of the *model* DMUs can be obtained as the maximum among the efficiency scores obtained from a number of CR-DEA models equal to the number of DMUs in $\mathcal{R}$, each of which uses a single DMU $k \in \mathcal{R}$ as the *model* DMU.

We now consider the relations between the CR($\mathbb{I}$)-DEA and VEA models. From the relation in (7.21) we see that that the VEA model provides equal efficiency scores to the CR($\mathbb{I}$)-DEA model, provided that the DMUs in set $\mathcal{R}$ jointly generate at least one facet of the DEA frontier. Thus, we have:

PROPOSITION 7.6: Given a set of *model* DMUs that are jointly efficient, a VEA model in which these DMUs comprise the MPS provides equal efficiency scores to a CR-DEA model in which the set of feasible weight vectors is defined as the *intersection* of the sets containing the weight vectors that are optimal in the DEA model for each of the *model* DMUs

We now show that relation similar to that deduced by Olesen and Petersen (2003) for the about the CR($\mathbb{U}$)-DEA and the VEA model in the case of multiple *model* DMUs, can also be obtained for the CR($\mathbb{I}$)-DEA and the VEA models. In particular, when DMUs *A* and *B* are chosen as *model* DMUs in a CR($\mathbb{I}$)-DEA model, we obtain the following relation between the sets of feasible weight vectors in the CR($\mathbb{I}$)-DEA model and in the VEA models in which either DMU *A* or *B* is the MPS:

$$\mathcal{F}_{CR(\mathbb{I})}^{AB} \equiv \mathcal{F}_{\mathbb{I}}^{AB} = \{ H_2\} = \mathcal{F}_V^A \cap \mathcal{F}_V^B$$

namely that the set of feasible weight vectors for the CR($\mathbb{I}$)-DEA model is the *intersection* of the sets containing the feasible weight vectors in the two VEA models in which DMUs *A* and *B* are respectively the MPS. A similar relation can be inferred when $\mathcal{R} = \{B, C\}$ and $\mathcal{R} = \{C, D\}$ in a CR($\mathbb{I}$)-DEA model. Thus, we have the following relation:

$$\mathcal{F}_{CR(\mathbb{I})}^{\mathcal{R}} = \cap_{k\in\mathcal{R}} \mathcal{F}_V^k \tag{7.25}$$

For some for the evaluated DMUs, one of the weight vectors that are optimal (i.e., maximize the DMU's efficiency) within each of the $\mathcal{F}_V^k,\ k \in \mathcal{R}$ will also be included in $\mathcal{F}_{CR(\mathbb{I})}^{\mathcal{R}}$, in which case one can obtain the CR($\mathbb{I}$)-DEA efficiency scores as $\theta_{CR(\mathbb{I})}^{o,\mathcal{R}} = \min\limits_{k\in\mathcal{R}}\{\theta_V^{o,k}\}$. For other DMUs, $\mathcal{F}_{CR(\mathbb{I})}^{\mathcal{R}}$ may not contain any of the weight vectors that are optimal within each of the $\mathcal{F}_V^k,\ k \in \mathcal{R}$. In these cases, $\theta_{CR(\mathbb{I})}^{o,\mathcal{R}} < \min\limits_{k\in\mathcal{R}}\{\theta_V^{o,k}\}$. Thus, we have that:

$$\theta_{CR(\mathbb{I})}^{o,\mathcal{R}} \le \min_{k\in\mathcal{R}}\{\theta_V^{o,k}\} \tag{7.26}$$

which leads to the following:

Proposition 7.7: Given a set $\mathcal{R}$ of *model* DMUs, the efficiency scores of a CR-DEA model in which the set of feasible weight vectors is the *intersection* of the sets containing the weight vectors that are optimal in the DEA model for each of the *model* DMUs are lower than or equal to the minimum among the efficiency scores obtained from a number of VEA models equal to the number of DMUs in $\mathcal{R}$, each of which uses a single *model* DMU $k \in \mathcal{R}$ as the MPS

Thus, in the case of multiple *model* DMUs the minimum among a series of VEA models, each of which uses one of the *model* DMUs as the MPS, provides an upper bound for the efficiency scores of the CR($\mathbb{I}$)-DEA model. Notice that this upper bound can also be obtained using CR-DEA instead of VEA models. In particular, we can use relation (7.19) to rewrite the inequality in (26) as

$$\theta_{CR(\mathbb{I})}^{o,\mathcal{R}} \leq \min_{k \in \mathcal{R}} \left\{ \theta_{CR(\mathbb{I})}^{o,k} \right\} = \min_{k \in \mathcal{R}} \left\{ \theta_{CR(\mathbb{U})}^{o,k} \right\} \tag{7.27}$$

We thus have:

Proposition 7.8: Given a set $\mathcal{R}$ of *model* DMUs, the efficiency scores of a CR-DEA model in which the set of feasible weight vectors is defined as the *intersection* of the sets containing the weight vectors that are optimal in the DEA model for each of the *model* DMUs are lower than or equal to the minimum among the efficiency scores obtained from number of CR-DEA models equal to the number of DMUs in $\mathcal{R}$, each of which uses a single DMU $k \in \mathcal{R}$ as the *model* DMU.

Considering the relation between the efficiency scores obtained from the VEA and the EXFA-CR-DEA model in the case of multiple *model* DMUs, note that for the EXFA-CR-DEA model to result in feasible solutions, the model DMUs need not only to be jointly efficient but also jointly generate at least one FDEF. In this case we have the following from relation (7.21):

Proposition 7.9: Given a set of *model* DMUs that are jointly efficient and generate at least one FDEF of the DEA frontier, a VEA model in which these DMUs comprise the MPS provides an upper bound for the efficiency score of a CR-DEA model in which the set of feasible weight vector contains only vectors with strictly positive input and output weight components, each of which is optimal in the DEA model for all the *model* DMUs

A special case of the above relation arises when the number of DMUs in set $\mathcal{R}$ equals the largest number of DMUs that can jointly generate a facet of the DEA frontier. In our illustrative example this number is equal to two DMUs. If, for example, $\mathcal{R} = \{A, B\}$, then the sets of feasible weight vectors for the evaluated DMUs in the CR-DEA models and the VEA model are related as follows:

$$\mathcal{F}^{AB}_{EXFA-CR} = \{ H_2\} = \mathcal{F}^{AB}_{CR(\mathbb{I})} \equiv \mathcal{F}^{AB}_{\mathbb{I}} \equiv \mathcal{F}^{AB}_{V} \subset \mathcal{F}^{AB}_{CR(\mathbb{U})} \equiv \mathcal{F}^{AB}_{\mathbb{U}} = \{ H_1, H_2, H_3\}$$

in which case the same efficiency scores are obtained from the EXFA-CR-DEA and the VEA models for each evaluated DMU, since both models will evaluate each DMU based on a common vector of strictly positive input and output weights. The same is true when $\mathcal{R} = \{B, C\}$ and $\mathcal{R} = \{C, D\}$, since these sets contain jointly efficient DMUs. A similar relation cannot be provided when $\mathcal{R}$ is equal to either $\{A, C\}$, $\{A, D\}$, or $\{B, D\}$, since the DMUs in these sets do not jointly generate a DEA facet. Thus, when the set $\mathcal{R}$ comprises of exactly $I + J - 1$ (in CRS models) or $I + J$ jointly efficient DMUs (in VRS models), the relation between the sets of feasible weight vectors in the CR-DEA models and the VEA model is given as:

$$\mathcal{F}^{\mathcal{R}}_{EXFA-CR} \equiv \mathcal{F}^{\mathcal{R}}_{CR(\mathbb{I})} \equiv \mathcal{F}^{\mathcal{R}}_{\mathbb{I}} \equiv \mathcal{F}^{\mathcal{R}}_{V} \subset \mathcal{F}^{\mathcal{R}}_{CR(\mathbb{U})} \equiv \mathcal{F}^{\mathcal{R}}_{\mathbb{U}} \tag{7.28}$$

and the relation between the efficiency scores obtained, for each evaluated DMU, from these models is given as:

$$\theta^{o,\mathcal{R}}_{EXFA-CR} = \theta^{o,\mathcal{R}}_{CR(\mathbb{I})} = \theta^{o,\mathcal{R}}_{V} \leq \theta^{o,\mathcal{R}}_{CR(\mathbb{U})} \tag{7.29}$$

respectively. Thus, we have:

COROLLARY: Given a set of $I + J - 1$ $(I + J)$ jointly efficient *model* DMUs, a CRS (VRS) VEA model in which these DMUs comprise the MPS provides equal efficiency scores to a CRS (VRS) CR-DEA model in which the set of feasible weight vectors contains only vectors that (i) are optimal the DEA model for all the *model* DMUs, and (ii) contain strictly positive input and output weight components

Lastly, one can use the results in Propositions 7.6 and 7.7 to obtain a relation concerning the VEA efficiency scores in the case that multiple DMUs comprise the MPS. In particular, combining relations (7.21) and (7.26), we obtain:

$$\theta_V^{o,\mathcal{R}} \leq \min_{k\in\mathcal{R}}\{\theta_V^{o,k}\} \tag{7.30}$$

which results to the following Proposition:

PROPOSITION 7.10: Given a set $\mathcal{R}$ of *jointly efficient* DMUs, the efficiency scores of a VEA model in which these DMUs jointly comprise the MPS are lower than or equal to the minimum among the efficiency scores obtained from a number of VEA models equal to the number of DMUs in $\mathcal{R}$, each of which uses a single DMU $k \in \mathcal{R}$ the MPS.

Lastly, from the relations in (7.21), (7.24), (7.27) and (7.30) we get the following:

$$\begin{aligned}\theta_{FDEF-CR}^{o,\mathcal{R}} \leq \theta_{CR(\mathbb{I})}^{o,\mathcal{R}} = \theta_V^{o,\mathcal{R}} \leq \min_{k\in\mathcal{R}}\{\theta_V^{o,k}\} = \min_{k\in\mathcal{R}}\{\theta_{CR(\mathbb{I})}^{o,k}\} = \\ = \min_{k\in\mathcal{R}}\{\theta_{CR(\mathbb{U})}^{o,k}\} \leq \max_{k\in\mathcal{R}}\{\theta_V^{o,k}\} = \max_{k\in\mathcal{R}}\{\theta_{CR(\mathbb{I})}^{o,k}\} = \max_{k\in\mathcal{R}}\left\{\theta_{CR(\mathbb{U})}^{o,k}\right\} = \theta_{CR(\mathbb{U})}^{o,\mathcal{R}}\end{aligned} \tag{7.31}$$

which provides an overview of the identified relations between the three CR-DEA variants and VEA in the case that multiple DMUs are chosen as *model* DMUs or comprise the MPS.

An immediate implication of our theoretical results is that, given an arbitrary set of *model* DMUs, the efficiency scores of different CR-DEA variants can be estimated or approximated by means of either a single VEA model or a series of VEA models. In particular, the CR($\mathbb{I}$)-DEA efficiency scores can be obtained by estimating a VEA model in which the *model* DMUs jointly comprise the MPS, while the CR($\mathbb{U}$)-DEA scores can be obtained by estimating a series of VEA models, each of which uses a different DMU among those in set $\mathcal{R}$ as the MPS, and then choosing--for each evaluated DMU--the maximum among those VEA efficiency scores. On the other hand, an upper bound for the EXFA-CR-DEA scores can be obtained by means of a VEA model in which the *model* DMUs jointly comprise the MPS, while if the number of DMUs in set $\mathcal{R}$ is equal to $I+J-1$ in CRS models (or $I+J$ in VRS models) then the VEA model will provide efficiency scores that are equal to those of the EXFA-CR-DEA model. The practical usefulness of these results for empirical applications is that the estimation process of CR-DEA efficiency scores is simplified. In contrast to the

CR-DEA models, which would require identifying the facets of the DEA frontier in a prior step, the estimation of VEA models involves simply changing some linear inequalities in the DEA model in (7.1) to equalities.

The discussion in this section concerned *model* DMUs or MPSs chosen among the set of extreme-efficient DMUs. However, the above results also hold if the DM chooses a non-extreme efficient DMU as *model* DMU or as the MPS. The same is true for choosing a weakly efficient DMU, except for the results concerning the EXFA-CR-DEA model. The cone of feasible weights in this model includes only vectors with positive input and output weight components. Thus, this model will not have a feasible solution when a weakly efficient DMU --for which all optimal weight vectors in model (7.1) contain at least one zero input or output weight component-- is included in the set of *model* DMUs.

In case the DM includes a DEA-inefficient DMU in the set of *model* DMUs, it is straightforward from the model in (7.14) that the VEA model will not provide feasible solutions, despite the fact that $\mathcal{F}^k$ for this DEA-inefficient DMU will include at least one weight vector. This is because the model in (7.14) cannot provide feasible solutions if at least one among the DMUs comprising the MPS has an efficiency score lower than unity, as in this case the second set of restrictions would be violated. However, this inconvenience can be circumvented by using as the MPS the chosen DMUs' radial projection $(y_j^{k\prime}, x_i^{k\prime}) = (y_j^k, \theta_{DEA}^k x_i^k)$, where θ_{DEA}^k is the efficiency score obtained by the model in (1) for the DMU. By definition, the set $\mathcal{F}^{k\prime}$ for $(y_j^{k\prime}, x_i^{k\prime})$ contains the same weight vectors as the set $\mathcal{F}^k$. Thus, the results in this paper hold also when a DEA-inefficient DMU is among the set of *model* DMUs, provided that its radial projection is used as the MPS in the VEA models estimated in order to obtain the CR-DEA efficiency scores.

7.5. Empirical application: Japanese regional banks

7.5.1. Preliminaries

In this section we demonstrate our theoretical findings by using a sample of Japanese regional banks. The Japanese banking sector has endured several hurdles in the last decades, including the busting of the real estate bubble in the 1990s, the global financial crisis of 2008 (Hoshi and Kashyap, 2010) and the Great East Japan Earthquake in 2011

(Kourtzidis *et al.,* 2021). As such, it has undergone major changes including continuous government interventions to stabilize it.[111] These changes, along with specificities in the Japanese banking system compared to that of other developed economies have attracted much scientific interest in assessing the performance of Japanese banks. The majority of these studies has used DEA models for evaluating efficiency.[112]

7.5.2. Sample, variables and modelling choices

Japanese regional banks usually operate with the boundaries of a specific prefecture. There are two distinct groups, namely those being members of the Regional Banks Association of Japan (hereafter Regional Banks I) and those comprising the Second Association of Regional Banks (Regional Banks II).[113] The latter were originally joint stock companies ("Sogo Banks"), which were allowed to convert into ordinary regional banks in the end of the 1980s (Fukuyama, 1993). The two groups of banks are nowadays very similar in their operations but differ in various aspects such as their size and the restructuring processes underwent in the past (see, e.g., Drake *et al.*, 2003).

For our purposes, we use a sample of 30 regional banks for the fiscal year 2017.[114] These comprise roughly 30% of the 105 regional banks (64 of Regional Banks I and 41 of Regional Banks II) operating in that year and their selection is based on their relative size expressed as their share of deposits over the total sample deposits. In particular, among the banks for which complete data was available (51 of Regional Banks I and 23 of Regional Banks II) we have selected the 20 largest among the subsample of Regional Banks I and the 10 largest among the Regional Banks II subsample.

The selection of inputs and outputs follows the intermediation approach, in which banks are viewed as intermediates between borrowing and lending entities (Berger & Humphrey, 1992). In particular, we use three inputs, namely the total number of employees and the stocks of fixed assets and deposits, and two outputs, namely the stocks of loans and security investments. All variables except the number

---

[111] See Fukuyama (1993) and Fukuyama and Weber (2002) for an overview of developments in the Japanese baking system since the 1990s reforms.
[112] See Kourtzidis *et al.* (2021) for a recent detailed review of these studies.
[113] A full list of members in those groups is provided in JBA (2019).
[114] The Japanese fiscal year begins on April 1st, end ends the following year on March 31st.

of employees are measured in million Yen and were collected from the banks' balance sheets which are publicly provided by the Japanese Banks' Association. Data for the number of employees are collected from DataStream. Descriptive statistics of the model's variables are given in Table 7.2. In line with the majority of studies using DEA to assess banking performance (see, e.g., Fethi and Pasiouras (2010) for a review) our models are input-oriented. We adopt constant returns-to-scale, which is a reasonable modelling choice as the two groups of regional banks nowadays perform similar operations and operate under the same framework (Kourtzidis *et al.,* 2021).

7.5.3. Empirical results.

The names of the selected banks are given in column (1) of Table 7.3, while column (2) portrays respectively each bank's regional group (I or II). The DEA efficiency scores are given in column (3). The average efficiency score is 0.946, indicating that, on average, a bank could attain its given production of loans and securities with roughly 5% less input usage. This rather low level of inefficiency resonates with (i) latest evidence for improvements in Japanese banking performance in the years after the 2011 Earthquake (Kourtzidis *et al.,* 2021) and (ii) historical evidence that larger Japanese banks --which comprise our sample--perform relatively well (see, e.g., Fukuyama, 1993; Drake and Hall, 2003). Additional room for performance improvement appears to be larger for Regional Banks II, which are on average less efficient compared to those belonging to the first group (average score 0.918 versus 0.961). There are 10 efficient banks, eight from group I and two from group II.

Let us now assume that DMs (e.g., authorities such as the Bank of Japan) wish to compare the performance of each bank in the sample to that of some *model* banks by means of CR-DEA models. These *model* banks will be some among those technically efficient in Table 7.3, which are viewed as excellent performers. For example, given that SMEs financing is a core part of the regional banks' operations and also vital for the local economy (see, e.g., Fukuyama, 1993), excellence could be viewed as having an output mix relying heavily on loans. In this case, Kansai Urban Banking Corporation should be chosen as the *model*, as it has the largest share of loans in its output mix among the efficient banks. Alternatively, excellence might be viewed as receiving good ratings by international credit rating organisations. Chiba bank is an efficient bank that

Table 7.2: Descriptive statistics of model variables

	fixed assets (bn ¥)	employees (thousands)	deposits (bn ¥)	loans (bn ¥)	investment securities (bn ¥)
average	50.011	2.767	4961.282	3755.226	1463.122
minimum	23.871	1.448	1849.584	1385.955	154.340
maximum	96.120	4.543	11565.778	9305.388	3242.629
median	43.393	2.671	4649.862	3371.830	1459.670
standard deviation	20.649	0.903	2263.679	1817.374	848.512

had high credit ratings in the period prior to 2017 (Kourtzidis *et al.*, 2021) and thus could be another choice for the *model* bank.

To estimate the two CR-DEA models (CR(𝕀)-DEA or CR(𝕌)-DEA, as in these cases their scores are equivalent) in which the *model* bank is respectively Kansai Urban Banking Corporation and Chiba Bank, one would need to identify (possibly by means of the *Qhull* sofware) all facets generated by each of these two banks --irrespective of their dimension-- and the weight vectors associated with each facet. Then, for each evaluated DMU, efficiency scores should be estimated using each different weight vector and the maximum among these scores would be the DMUs' CR-DEA score. By virtue of Proposition 7.1, these scores can be obtained by means of VEA models in which the *model* banks are the MPS. The associated VEA efficiency scores are given in columns (4) and (5) of Table 7.3. From that we see that six banks drops from the efficient frontier when Kansai Urban Banking Corporation is used as the *model* bank, all of which are Regional Banks I. Average efficiency score drops to 0.918, while 21 banks in total exhibit efficiency declines compared to the DEA model. The authorities could advice these banks to shift their output mix more towards loan provision in the future. On the other hand, when Chiba Bank is the *model* bank, average efficiency is equal to 0.908 and 20 out of the 30 sample banks have lower efficiency scores compared to their respective DEA ones.

Alternatively, membership in the first or in the second Association of regional banks could be used as a criterion for choosing the *model* banks. Although at present there are no functional differences between banks of the two Associations, the two groups differ in certain aspects. For example, Regional Banks I are larger than those in group II and, as is the case in the resent study, they also tend to perform better on average (see e.g., Barros *et al.*, 2012; Kourtzidis *et al.*, 2021). Based on this, DMs could opt for a *model* bank that is part of group I. *Second*, banks of the Second Regional

Table 7.3: Empirical results from CR-DEA and VEA models for a sample of Japanese banks, 2017

(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)
Bank	group	DEA	VEA	VEA	VEA	VEA	CR(U)-DEA	VEA	min{(8),(9)}	EXFA-CR-DEA
ID of *model*/MPS bank(s)		-	{554}	{134}	{143}	{525}	{143, 525}	{143, 525}	-	{143, 525}
Chiba Bank	I	1.000	0.985	1.000	1.000	0.985	1.000	0.981	0.985	0.981
Shizuoka Bank	I	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
Joyo Bank	I	0.983	0.937	0.983	0.971	0.942	0.971	0.940	0.942	0.940
77 Bank	I	1.000	0.866	1.000	1.000	0.901	1.000	0.901	0.901	0.901
Hiroshima Bank	I	0.981	0.978	0.981	0.960	0.979	0.979	0.960	0.960	0.960
Bank of Kyoto	I	1.000	0.975	1.000	1.000	1.000	1.000	1.000	1.000	1.000
Gunma Bank	I	1.000	0.998	1.000	1.000	1.000	1.000	1.000	1.000	1.000
Hachijuni Bank	I	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
Chugoku Bank	I	1.000	0.936	0.974	1.000	0.982	1.000	0.982	0.982	0.970
Juroku Bank	I	0.883	0.879	0.855	0.883	0.883	0.883	0.883	0.883	0.883
Toho Bank	I	0.865	0.808	0.865	0.863	0.815	0.863	0.815	0.815	0.815
Iyo Bank	I	0.992	0.978	0.961	0.988	0.992	0.992	0.988	0.988	0.988
Ogaki Kyoritsu Bank	I	0.991	0.991	0.911	0.991	0.991	0.991	0.991	0.991	0.975
Yamaguchi Bank	I	0.890	0.890	0.747	0.890	0.890	0.890	0.890	0.890	0.859
Nanto Bank	I	0.895	0.869	0.885	0.895	0.895	0.895	0.895	0.895	0.895
Hyakugo Bank	I	1.000	0.845	0.864	0.950	0.926	0.950	0.926	0.926	0.895
Shiga Bank	I	0.960	0.953	0.952	0.956	0.960	0.960	0.956	0.956	0.956
Daishi Bank	I	0.937	0.898	0.913	0.937	0.935	0.937	0.935	0.935	0.929
Higo Bank	I	0.922	0.904	0.906	0.922	0.922	0.922	0.922	0.922	0.922
Suruga Bank	I	0.909	0.909	0.884	0.853	0.873	0.873	0.853	0.853	0.853
Keiyo Bank	II	0.933	0.932	0.931	0.908	0.933	0.933	0.908	0.908	0.908
Kansai Urban Banking Corporation	II	1.000	1.000	0.950	1.000	1.000	1.000	1.000	1.000	1.000
Bank of Nagoya	II	0.896	0.887	0.871	0.896	0.896	0.896	0.896	0.896	0.896
Minato Bank	II	0.893	0.893	0.793	0.893	0.893	0.893	0.893	0.893	0.877
Aichi Bank	II	0.951	0.843	0.859	0.905	0.897	0.905	0.894	0.897	0.877
Tochigi Bank	II	0.830	0.830	0.775	0.830	0.830	0.830	0.830	0.830	0.825
Yachiyo Bank	II	0.849	0.810	0.769	0.847	0.849	0.849	0.847	0.847	0.829
Towa Bank	II	0.884	0.829	0.791	0.884	0.884	0.884	0.884	0.884	0.857
Ehime Bank	II	0.948	0.924	0.879	0.941	0.948	0.948	0.941	0.941	0.931
Higashi-Nippon Bank	II	1.000	1.000	0.942	1.000	1.000	1.000	1.000	1.000	1.000
efficient banks		10	4	6	9	6	9	6	6	6
average		0.946	0.918	0.908	0.939	0.933	0.942	0.930	0.931	0.924

Association have undergone a different restructuring process in the turn of the century compared to Regional Banks I. Both groups received capital injections but regional banks II had to cope with their bad loans by themselves. On this basis, one could argue that the efficient regional banks II have accumulated significant managerial abilities as a result of enduring more turbulence over the years.

Let us assume that DMs choose one *model* bank from each of the two groups (Hachijuni Bank from group I and Higashi-Nippon Bank from group II) and force each evaluated bank to be compared with at least one of the two *model* banks by means of a CR($\mathbb{U}$)-DEA model. This could be a compromise solution to disagreements among DMs on which of the two should be used as *model* bank. It could also reflect the view that smaller banks (group II) should not be forced to compare their performance to that of a medium-sized bank (the group I *model* bank) and vice versa. The CR($\mathbb{U}$)-DEA efficiency scores can, by virtue of Proposition 7.4, be obtained without the need to identify the cone of feasible weights as the maximum among the efficiency scores of two different VEA models in which Hachijuni Bank and Higashi-Nippon Bank are respectively the MPS. The scores of these two VEA models are given in columns (6) and (7) of Table 7.3, while the CR($\mathbb{U}$)-DEA score (their maximum) is given in column (8). From that we see that only one bank drops from the efficient frontier while the average efficiency score is only slightly below that of the DEA model (0.942 versus 0.946). It is interesting to notice that, for three banks of the regional I group (Iyo Bank, Shiga Bank, and Suruga Bank) the CR($\mathbb{U}$)-DEA efficiency score corresponds to the score of the VEA model in which the MPS is a group II bank (Higashi-Nippon Bank). The reverse holds for Aichi Bank which is part of the regional II group. This could be taken as evidence that the operations of banks from different Regional Associations have over time converged enough, so that banks of the one Association can serve as benchmarks for banks in the other.

The similarities in the operations of regional banks I and II is also indicated by the fact that banks of both Regional Associations appear as peers for inefficient banks in the DEA model. This is the case for Hachijuni Bank and Higashi-Nippon Bank as well. Thus, DMs could have also used a CR($\mathbb{I}$)-DEA model in which the set of *model* DMUs includes both these banks for performance evaluation. By virtue of Proposition 7.6 the scores of such a model are equivalent to those of a VEA model in which Hachijuni Bank and Higashi-Nippon Bank jointly comprise the MPS and are given in column (9) of Table 7.3. We see that average efficiency declines slightly compared to that of the DEA model while four banks drop from the efficiency list, all of which are part of group I. For comparison, the minimum among the scores of the two VEA models in which Hachijuni Bank and Higashi-Nippon Bank are used as the sole MPS (i.e., the minimum of the scores in columns (6) and (7) for each bank) is given in column (10). A comparison among the scores in columns (9) and (10) reveals that, for each bank, the former score is lower than or equal to the latter. In particular, there are three banks (Chiba Bank, Joyo Bank, and Aichi Bank) for which the VEA score in column (9) is lower than the minimum among the scores in columns (6) and (7). This is in accordance with Propositions 7.7, 7.8 and 7.10. In addition, by comparing the scores in columns (8) and (9), we see that the efficiency scores of the VEA model, in which Hachijuni Bank and Higashi-Nippon Bank jointly comprise the MPS, are lower than or equal to those of the CR($\mathbb{U}$)-DEA model, in which these two banks are the *models*. This is in accordance with Proposition 7.3.

Lastly, let us assume that DMs choose as Hachijuni Bank and Higashi-Nippon Bank as the *model* banks but also wish to evaluate all banks based on well-defined input and output weights using an EXFA-CR-DEA model. This would require estimating the mixed-integer model in (7.12). Alternatively, following the approach outlined in Thanassoulis *et al.* (2008), one would need to identify all the FDEFs jointly generated by Hachijuni Bank and Higashi-Nippon Bank and the weight vectors that are normal to each of them, estimate the efficiency scores of each banks using each of these weight vectors and selecting the maximum among these scores. Using the approach outlined in Davtablab-Olyaie *et al.* (2014), we identified three FDEFs jointly generated by the two *model* banks and the weight vectors normal to each of them.[115] We estimated the efficiency scores of each evaluated bank using each of these weight vectors and selected the maximum among these scores, which is given in column (11) of Table 7.3. Comparing these with the scores of the VEA model in which the two *model* banks jointly comprise the MPS (column (9)), we see that the efficiency scores of VEA model

[115] These FDEFs are generated by Hachijuni Bank and Higashi-Nippon Bank jointly with (i) Shizuoka Bank and Gunma Bank, (ii) Shizuoka Bank and Kansai Urban Banking Corporation, and (iii) Bank of Kyoto and Gunma Bank.

constitute an upper bound for those of the EXFA-CR-DEA model. This is in accordance with Proposition 7.9.

7.6. Concluding remarks

CR-DEA models are suitable for incorporating DM views about excellent performing DMUs that should serve as role-models for others. However, their estimation has up to today remained rather complicated task as it required to identify all the efficient facets of the DEA frontier and the weight vectors normal to each of them.

In this paper, we elaborated more on the relations between VEA and CR-DEA models including DM preferences in the form of *model* DMUs for two different specifications for the set of *model* DMUs, namely that it comprises of a single DMU and that it contains multiple DMUs. In the latter specification we distinguished between two CR-DEA variants, namely CR($\mathbb{I}$)-DEA and CR($\mathbb{U}$)-DEA, in which the set of feasible weights is respectively specified as the intersection and the union of the sets containing the weight vectors that are optimal in DEA for each *model* DMU. In addition, for both settings we considered the EXFA-CR-DEA model, for which the set of feasible weight vectors contains only those vectors with strictly positive components, each of which is optimal in DEA for *all* the *model* DMUs. Our results suggest that in both specifications, EXFA-CR-DEA provides the minimum efficiency score among all four models, CR($\mathbb{U}$)-DEA provides the maximum score, while VEA and CR($\mathbb{I}$)-DEA provide equal scores to each other. In addition, the only difference in the four models' relations across the two specifications lies in the relation between VEA and CR($\mathbb{U}$)-DEA. The two models provide equal efficiency scores for the case of a single *model* DMU, while when there are multiple *model* DMUs the CR($\mathbb{U}$)-DEA efficiency scores are larger than or equal to those of VEA.

The results of this paper provide a detailed overview of the relations between VEA and CR-DEA models. These extend earlier inferences in the relevant literature and provide a means to simplify the estimation of CR-DEA models. In particular, and regardless of the number of *model* DMUs chosen by the DM, estimating a VEA model or a series of VEA models suffices to obtain the efficiency scores of the CR($\mathbb{I}$)-DEA and the CR($\mathbb{U}$)-DEA models respectively. Our results can accommodate the inclusion of a DEA-inefficient DMU in the set of *model* DMUs, while they also hold for different returns-to-scale assumptions. As regards returns to scale, VRS VEA models have been shown to provide unacceptable (e.g., zero and negative) efficiency scores (Korhonen *et al.*, 2002). A promising avenue for future research would be to investigate, based on the results identified in the paper, the implications of this for the efficiency scores of different CR-DEA models.

CHAPTER 8

Concluding remarks

8.1. Summary

The aim of this thesis is to analyze several theoretical and empirical aspects of VEA. VEA is a method that uses the notion of the MPS to incorporate the preferences of a DM, namely a social planner, regulator, or manager, in the measurement of relative technical efficiency through DEA models. Up to today, its use for decision-making problems in which DM preferences are frequently accounted for, such as the evaluation of effectiveness and cross efficiency, was not considered, while its potential relationships with other approaches incorporating preferences in DEA models were not thoroughly investigated.

Chapters two to four, which comprise the first part of this thesis, are the empirical essays. In the first of them, we used VEA as an alternative for effectiveness assessment by incorporating DMs' views about the DMUs that are "doing the right things" in the choice of the MPS. The VEA efficiency scores were then viewed as effectiveness estimates and further decomposed into one component capturing technical efficiency and another capturing the DMUs' relative distance from the DM's range of desirable structures. We demonstrated the usefulness of the approach by using it to assess the effectiveness of countries in utilizing their income to develop their citizens' social prosperity or human capabilities.

In the second empirical essay, we proposed the use of VEA as a means to incorporate DM preferences in the construction of composite indicators, by developing the VEA-BoD model. This was then used to re-estimate the UN HDI.

In the third empirical essay, we assessed the implications of MPS choice for the VEA efficiency scores. We reviewed the various suggestions proposed for choosing the MPS in the VEA literature and presented some new, which are based respectively on the relative position of efficient DMUs on the DEA frontier, MPSS DMUs, the APU, and common weights. Comparative empirical analysis regarding the effect of alternative MPS choices on the VEA efficiency scores was provided using a dataset of Greek cotton farms. The results provide useful information regarding the MPS choices that are more likely to result in insignificant or excessive differences between the DEA and the VEA efficiency scores, and the choices which are frequently similar to each other in practice.

The second part contains three theoretical essays and comprises of chapters five to seven. In the first theoretical essay, we related VEA to cross efficiency, namely the notion of peer appraisal in DEA. Particularly, we showed that the VEA model is equivalent to the TB cross efficiency model, provided that the "reference" DMU in the TB model if it is an efficient one, or its radial projection on the DEA frontier if it is inefficient, is chosen as the MPS in the VEA model.

In the second theoretical essay, we examined the relationship between VEA and DEA models with weight restrictions and their dual production trade-offs and showed that the VEA model can be viewed as a particular class of DEA models with production trade-offs. The coefficient vectors in these trade-offs, which are dual to Type II assurance region weight restrictions, are equal to the deviations of all evaluated DMUs' input and output quantities from those of the DMUs chosen as the MPS. We also showed that, if these Trade-Offs are considered only for the inputs or the output, then a similar equivalence holds between pure output or input VEA models and DEA models with production trade-offs.

In the last theoretical essay, we elaborated more on the relationship between VEA and CR-DEA models that include preferences on efficient DMUs than DMs consider as examples (*model* DMUs) for the remaining DMUs. We showed that, provided that the *model* DMUs in CR-DEA are those that constitute the MPS in VEA, the efficiency scores from a CR-DEA model in which the cone of feasible weight vectors is specified as the intersection of the sets containing the weight vectors that are optimal in DEA for each *model* DMU are equivalent to the VEA scores. In addition, we showed that the VEA scores provide a lower and an upper bound for the scores obtained from two other CR-DEA models. In the former case, the cone of feasible weights in the CR-DEA model is given as the union of the sets containing the optimal weight vectors for each *model* DMU, while in the latter case it consists only of the strictly positive weight vectors that are jointly optimal in DEA for all the *model* DMUs.

8.2. Implications

The theoretical and empirical results obtained in the previous chapters have several implications: *first,* the detailed and enlarged list of the various MPS choice presents DMs and practitioners with a variety of alternative options to choose the MPS in practice. It also provides insights regarding the rationale related to each choice and its potential usefulness in particular performance evaluation cases. For instance, in the case of a manager interested in reorganizing efficiently a group of retail branches, the use of the APU as the MPS could provide useful insights, as the resulting efficiency scores reflect the performance of DMUs from the perspective of centralized resource allocation. On the other hand, the use of common weights in VEA may be preferred when the DMUs need to be assessed against a common standard or should follow organizational objectives rather than pursuing their own. Lastly, when the DM views a particular input or output as the most important one in assessing performance, these preferences could be reflected through the use of a best-in-input or a best-in-output MPS.

Second, DMs and practitioners are presented with insights regarding the practical implications of using alternative MPSs for the resulting efficiency scores. For instance, using influential peers as the MPS in VEA is not expected to present DMs with useful additional insights compared to those of the DEA model, while alternative ways to select an influential peer as the MPS are also similar to each other in practice. On the other hand, the use of other MPSs (such as the APU, an MPSS DMU, or a combination of DMUs generating an FDEF) might offer interesting insights to management that would complement those of the DEA model. Moreover, “end-point” MPSs are more likely to imply an input/output bundle that is very dissimilar compared to the bundles used by most DMUs, while different kinds of “end-point” MPSs might imply different kinds of extreme bundles from each other.

Third, alternative economic interpretations are provided for the DMs’ judgements in VEA and the associated efficiency scores. More specifically, through the equivalence of the VEA model with DEA models with production trade-offs, the DM’s choice of the MPS can be interpreted as incorporating a particular form of additional trade-off relations in a DEA model. These trade-offs aim to restrict the marginal rates of input substitution and output transformation for each assessed DMU to take values among those that are observed in the neighborhood of the selected MPS. On the other hand, one can also interpret the production trade-offs appended in DEA models, when the resulting linear model is equivalent to a VEA one for a particular MPS choice, as reflecting the judgements of a DM regarding the most preferred input/output bundle. Thus, alternative interpretations can be provided for both the VEA efficiency scores and those of DEA models with production trade-offs and their dual weight restrictions.

Through the equivalence of the VEA model to the TB cross efficiency model, the VEA scores can be interpreted as the most favorable (i.e., the TB) cross-efficiency scores from the perspective of a particular "reference" DMU, namely the one chosen as the MPS in VEA, while the TB cross-efficiency scores when a particular DMU is used as a "reference" reflect also the judgements of a DM that views this "reference" DMU has having the most desirable input/output bundle.

Moreover, the VEA efficiency scores can also be interpreted as empirical estimates of the DMUs' effectiveness, namely the extent that DMUs do the "right things" such as follow behavioral or organizational objectives, norms of mandates or abide by certain agreements set up with management. In this case, the DMU or the set of DMUs chosen as the MPS are considered as those "doing the right things", i.e., those aligned the most with the specified objectives or most closely following agreements and mandates.

Fourth, we showed for the first time that cross efficiency scores can be obtained using the envelopment formulation of (VEA) linear models, rather than the multiplier formulation of DEA models, which was the only way to obtain cross efficiency scores up to today. This holds for a particular form of cross efficiency scores, namely those obtained through the TB formulation, and is an implication of the equivalence of the VEA model to the TB cross efficiency model.

Fifth, computational gains are provided in the estimation of particular DEA models that are rather complicated or their estimation is time-consuming. More specifically, the relations identified between VEA and the three CR-DEA models in the seventh chapter allow estimating or approximating the efficiency scores of CR-DEA models by means of VEA. The VEA models are less computationally demanding as they do not require to *a priori* identify the cone of feasible weight vectors. In addition, significant shortcuts can be obtained in the estimation of the TB cross efficiency matrix though the use of VEA models instead of the TB formulation by (i) providing the cross-efficiency scores through the envelopment form of VEA rather than the multiplier one, and (ii) estimating fewer linear models when there exist inefficient DMUs that are projected on the same part of the strongly efficient DEA frontier.

Sixth, the empirical applications of the VEA model presented in the previous chapters place VEA as a useful alternative for several cases of applied performance evaluation. Particularly, as demonstrated in the second chapter, VEA can be used for assessing effectiveness, the measurement of which is crucial in cases of an entity centrally managing or coordinating a set of DMUs (such as branches of the same firm or firms operating within the same sector), where certain directions and mandates are given by the coordinating authority (the firm's general manager or the sector's planner) and abiding by those is essential for the performance of the entity (i.e., the firm or the sector) as a whole. In addition, the VEA-BoD model introduced in the third chapter is a useful alternative for incorporating preference information in the construction of composite indicators. The use of such indicators is increasingly widespread nowadays, mainly attributed to their ability to communicate multifaceted information regarding socio-economic phenomena in a reduced but concise form. Their construction however frequently requires that the preferences of DMs, social planners, or even the public regarding the relative importance of a phenomenon's dimensions is taken into account. This can be facilitated through the choice of the MPS in a VEA-BoD model.

References

Adler, N., Friedman, L., & Sinuany-Stern, Z. (2002). Review of ranking methods in the data envelopment analysis context. *European Journal of Operational Research, 140*, 249-265. https://doi.org/10.1016/S0377-2217(02)00068-1.

Afsharian, M., Ahn, H., & Harms, S.G. (2021). A review of approaches applying a common set of weights: The perspective of centralized management. *European Journal of Operational Research, 294,* 3-15. https://doi.org/10.1016/j.ejor.2021.01.001.

Afsharian, M., Ahn, H., & Thanassoulis, E. (2019). A frontier-based system of incentives for units in organisations with varying degrees of decentralisation. *European Journal of Operational Research, 275,* 224-237. https://doi.org/10.1016/j.ejor.2018.11.036.

Alcaraz, J, Aparicio, Monge, J.F., & Ramón, N. (2022). Weight profiles in cross-efficiency evaluation based on hypervolume maximization. *Socio-Economic Planning Sciences* (in press) https://doi.org/10.1016/j.seps.2022.101270.

Alkire, S. (2005). Why the Capability Approach? *Journal of Human Development and Capabilities, 6*, 115-135. https://doi.org/10.1080/146498805200034275.

Allen, R., Athanassopoulos, A., Dyson, R. G., & Thanassoulis, E. (1997). Weights restrictions and value judgements in data envelopment analysis: evolution, development and future directions. *Annals of Operations Research, 73*, 13-34. https://doi.org/10.1023/A:1018968909638.

Al-Siyabi, M., Amin, G.R., Bose, S. & Al-Masroori, H. (2019). Peer-judgment risk minimization using DEA cross-evaluation with an application in fishery. *Annals of Operations Research*, 274, 39-55 https://doi.org/10.1007/s10479-018-2858-3.

Amado, C.A.F., & Dyson, R.G. (2009). Exploring the use of DEA for formative evaluation in primary diabetes care: an application to compare English practices. *Journal of the Operational Research Society 60,* 1469-1482 https://doi.org/10.1057/jors.2008.160.

Anand, S., & Sen, A. (2000a). Human development and economic sustainability. *World Development*, *28*, 2029-2049. https://doi.org/10.1016/S0305-750X(00)00071-1.

Anand, S., & Sen, A. (2000b). The income component of the human development index. *Journal of Human Development and Capabilities*, *1*, 83-106. https://doi.org/10.1080/14649880050008782.

Andersen, P., & Petersen, N.C. (1993). A procedure for ranking efficient units in data envelopment analysis. *Management Science, 39,* 1261-1264. https://doi.org/10.1287/mnsc.39.10.1261.

Angulo-Meza, L., & Estellita-Lins, M. P. (2002). Review of Methods for increasing discrimination in Data Envelopment Analysis. *Annals of Operations Research, 116*, 225-242. https://doi.org/10.1023/A:1021340616758.

Aparicio, J. & Zofío, J.L. (2021). Economic cross-efficiency. *Omega*, 100, 102374. https://doi.org/10.1016/j.omega.2020.102374.

Asmild, M., Paradi, J.C., Reese, D.N., & Tam, F. (2007). Measuring overall efficiency and effectiveness using DEA. *European Journal of Operational Research, 178,* 305-321. https://doi.org/10.1016/j.ejor.2006.01.014.

Atici, K.B., & Podinovski, V.V. (2015). Using data envelopment analysis for the assessment of technical efficiency of units with different specialisations: An application to agriculture. *Omega*, *54,* 72-83. https://doi.org/10.1016/j.omega.2015.01.015.

Baker, R.C., & Talluri, S. (1997). A closer look at the use of data envelopment analysis for technology selection. *Computers and Industrial Engineering, 32*, 101-108. https://doi.org/10.1016/S0360-8352(96)00199-4.

Banker, R.D. (1984). Estimating most productive scale size using Data Envelopment Analysis. *European Journal of Operational Research, 17,* 35-44. https://doi.org/10.1016/0377-2217(84)90006-7.

Banker, R.D., & Chang, H. (2006). The super-efficiency procedure for outlier identification, not for ranking efficient units. *European Journal of Operational Research, 175*, 1311-1320. https://doi.org/10.1016/j.ejor.2005.06.028.

Banker, R.D., & Natarajan, R. (2011). Statistical Tests Based on DEA Efficiency Scores. In Cooper, W.W., Seiford, L.M., & Zhu, J. (*eds*.). *Handbook on Data Envelopment Analysis*, International Series in Operations Research and Management Science 164, Springer. https://doi.org/10.1007/978-1-4419-6151-8_11.

Banker, R.D., & Thrall, R.M. (1992). Estimation of returns to scale using Data Envelopment Analysis. *European Journal of Operational Research, 62,* 74-84. https://doi.org/10.1016/0377-2217(92)90178-C.

Banker, R.D., Charnes, A., & Cooper, W.W. (1984). Some models for estimating technical and scale inefficiencies in data envelopment analysis. *Management Science*, *30*, 1078-1092. https://doi.org/10.1287/mnsc.30.9.1078.

Barros, C.P., Managi, S., & Matousek, R. (2012). The technical efficiency of the Japanese banks: Non-radial directional performance measurement with undesirable output. *Omega,* 40, 1-8. https://doi.org/10.1016/j.omega.2011.02.005.

Berger, A., & Humphrey, D. (1992). Measurement and efficiency issues in commercial banking. In Z. Griliches (Ed.): *Output measurement in service sectors: National*

Bureau of Economic Research, studies in income and wealth, 56, University of Chicago Press. https://www.nber.org/system/files/chapters/c7237/c7237.pdf

Blancard, S., & Hoarau, J.F. (2013). A new sustainable human development indicator for small island developing states: a reappraisal from data envelopment analysis. *Economic Modelling, 30,* 623-635. https://doi.org/10.1016/j.econmod.2012.10.016

Bogetoft, P., & Otto, L. (2011). *Benchmarking with DEA, SFA, and R*. International Series in Operations Research and Management Science 157, Springer. https://doi.org/10.1007/978-1-4419-7961-2.

Brockett, P.L., Charnes, A., Cooper, W.W., Huang, Z.M., & Sun, D.B. (1997). Data transformations in DEA cone ratio envelopment approaches for monitoring bank performances. *European Journal of Operational Research*, 98, 250-268. https://doi.org/10.1016/S0377-2217(97)83069-X.

Canning, D., French, D., & Moore, M. (2013). Non-parametric estimation of data dimensionality prior to data compression: the case of the human development index. *Journal of Applied Statistics*, *40*, 1853-1863. https://doi.org/10.1080/02664763.2013.798629.

Carrillo, M., & Jorge, J.M. (2018). An alternative neutral approach for cross-efficiency evaluation *Computers & Industrial Engineering*, 120, 137-145. https://doi.org/10.1016/j.cie.2018.04.017.

Cazals, C., Florens, J.-P., & Simar, L. (2002). Nonparametric frontier estimation: a robust approach. *Journal of Econometrics*, *106*, 1-25. https://doi.org/10.1016/S0304-4076(01)00080-X.

Chakravarty, S. (2011). A Reconsideration of the tradeoffs in the new Human Development Index. *Journal of Economic Inequality, 9,* 471-474. https://doi.org/10.1007/s10888-011-9190-3

Chakravarty, S.R. (2003). A generalized human development index. *Review of Development Economics 7,* 99-114. https://doi.org/10.1111/1467-9361.00178.

Charnes, A., & Cooper, W.W. (1962). Programming with linear fractional functionals. *Naval Research Logistics Quarterly, 9,* 181-186. https://doi.org/10.1002/nav.3800090303.

Charnes, A., Cooper, W. W., Golany, B., Seiford, L., & Stutz, J. (1985). Foundations of data envelopment analysis for Pareto–Koopmans efficient empirical production functions. *Journal of Econometrics, 30,* 91-107 https://doi.org/10.1016/0304-4076(85)90133-2.

Charnes, A., Cooper, W.W., & Rhodes, E. (1978). Measuring the efficiency of decision making units. *European Journal of Operational Research, 2,* 429-444. https://doi.org/10.1016/0377-2217(78)90138-8.

Charnes, A., Cooper, W.W., & Thrall, R.M. (1991a). A structure for classifying and characterizing efficiency and inefficiency in Data Envelopment Analysis. *Journal of Productivity Analysis 2*, 197-237. https://doi.org/10.1007/BF00159732.

Charnes, A., Cooper, W.W., Huang, Z.M., & Sun, D.B. (1990). Polyhedral cone- ratio DEA models with an illustrative application to large commercial banks. *Journal of Econometrics 46*, 73-91. https://doi.org/10.1016/0304-4076(90)90048-X.

Charnes, A., Cooper, W.W., Huang, Z.M., & Sun, D.B. (1991b). Relations between half-space and finitely generated cones in polyhedral cone-ratio DEA models. *International Journal of Systems Science*, *22*, 2057-2077. https://doi.org/10.1080/00207729108910773.

Charnes, A., Cooper, W.W., Wei, Q.L., & Huang, Z.M. (1989). Cone ratio data envelopment analysis and multi-objective programming. *International Journal of Systems Science*, *20*, 1099-1118. https://doi.org/10.1080/00207728908910197.

Chen, H. (2018). Average lexicographic efficiency for data envelopment analysis. *Omega*, *74*, 82–91. https://doi.org/10.1016/j.omega.2017.01.008.

Chen, L., Wang, S-H, & Wang, Y-M. (2022) A new data envelopment analysis clustering approach within cross-efficiency framework, *Journal of the Operational Research Society*, *73*, 664-673 https://doi.org/10.1080/01605682.2020.1857667.

Cherchye, L., Moesen, W., Rogge, N., & van Puyenbroeck, T. (2007a). An introduction to 'benefit of the doubt' composite indicators. *Social Indicators Research*, *82*, 111-145. https://doi.org/10.1007/s11205-006-9029-7.

Cherchye, L., Lovell, C.A.K., Moesen, W., & van Puyenbroeck, T. (2007b). One market, one number? A composite indicator assessment of EU internal market dynamics. *European Economic Review*, *51*, 749–779. https://doi.org/10.1016/j.euroecorev.2006.03.011.

Cherchye, L., Moesen, W., Rogge, N., Van Puyenbroeck, T., Saisana, M., Saltelli, A., Liska, R., & Tarantola, S. (2008). Creating composite indicators with DEA and robustness analysis: the case of the Technology Achievement Index. *Journal of the Operational Research Society*, *59*, 239–251. https://doi.org/10.1057/palgrave.jors.2602445.

Chilingerian, J.A., & Sherman, H. (1997). DEA and primary care physician report cards: Deriving preferred practice cones from managed care service concepts and operating strategies. *Annals of Operations Research*, 73, 35-66. https://doi.org/10.1023/A:1018993515090.

Chowdhury, S., & Squire, L. (2006). Setting weights for aggregate indices: An application to the commitment to development index and human development index. *Journal of Development Studies*, *42*, 761-771. https://doi.org/10.1080/00220380600741904.

Çillingirtürk, A.M., & Koçak, H. (2018). Human Development Index (HDI) Rank-Order Variability. *Social Indicators Research*, *137*, 481–504. https://doi.org/10.1007/s11205-017-1605-5.

Contreras, I. (2012). Optimizing the rank position of the DMU as secondary goal in DEA cross-evaluation. *Applied Mathematical Modelling*, *36*, 2642-2648. https://doi.org/10.1016/j.apm.2011.09.046.

Contreras, I., Lozano, S., & Hinojosa, M.A. (2021). A DEA cross-efficiency approach based on bargaining theory. *Journal of the Operational Research Society*, *72*, 1156-1167. https://doi.org/10.1080/01605682.2020.1755898.

Cook, W.D., Ramon, N., Ruiz, J.L., Sirvent, I., & Zhu, J. (2019). DEA-based benchmarking for performance evaluation in pay-for-performance incentive plans. *Omega*, *84*, 45-54. https://doi.org/10.1016/j.omega.2018.04.004.

Cook, W.D., Seiford, L.M., & Zhu, J. (2004). Models for performance benchmarking: measuring the effect of e-business activities on banking performance. *Omega*, *32*, 313-322. https://doi.org/10.1016/j.omega.2004.01.001.

Cooper, W.W., Park, K. S., & Pastor, J.T. (1999). RAM: A Range Adjusted Measure of inefficiency for use with additive models, and relations to other models and measures in DEA. *Journal of Productivity Analysis*, *11*, 5-42. https://doi.org/10.1023/A:1007701304281.

Cooper, W.W., Ruiz, J.L., & Sirvent, I. (2007b). Choosing weights from alternative optimal solutions of dual multiplier models in DEA. *European Journal of Operational Research*, *180*, 443-458. https://doi.org/10.1016/j.ejor.2006.02.037.

Cooper, W.W., Ruiz, J.L., & Sirvent, I. (2011). Choices and uses of DEA weights. In: Cooper, W.W., Seiford, L.M., & Zhu, J. (Eds.), *Handbook on data envelopment analysis* (pp. 93-126) Springer. https://doi.org/10.1007/978-1-4419-6151-8_4.

Cooper, W.W., Seiford, L.M., & Tone, K. (2007a). *Data envelopment analysis: a comprehensive text with models, applications, references and DEA-Solver software* (2nd ed.). Kluwer Academic, Boston. https://doi.org/10.1007/978-0-387-45283-8.

Daraio, C., & Simar, L. (2005). Introducing environmental variables in nonparametric frontier models: A probabilistic approach. *Journal of Productivity Analysis*, *24*, 93-121. https://doi.org/10.1007/s11123-005-3042-8.

Davtalab-Olyaie, M. (2019). A secondary goal in DEA cross-efficiency evaluation: A “one home run is much better than two doubles” criterion. *Journal of the Operational Research Society*, *70*, 807-816. https://doi.org/10.1080/01605682.2018.1457482.

Davtalab-Olyaie, M., Ghandi, F., & Asgharian, M. (2021). On the spectrum of achievable targets in cross-efficiency evaluation and the associated secondary goal models. *Expert Systems with Applications, 177,* 114927. https://doi.org/10.1016/j.eswa.2021.114927.

Davtalab-Olyaie, M., Roshdi, I., Jahanshahloo, G., & Asgharian, M. (2014). Characterizing and finding full dimensional efficient facets in DEA: a variable returns to scale specification. *Journal of Operational Research Society*, 65, 1453-1464. https://doi.org/10.1057/jors.2013.97.

Dellnitz, A., Reucher, E., & Kleine, A. (2021). Efficiency evaluation in data envelopment analysis using strong defining hyperplanes: a cross-efficiency framework, *OR Spectrum*, *43*, 441-465. https://doi.org/10.1007/s00291-021-00623-2.

Desai, M.J. (1991). Human development: concepts and measurement. *European Economic Review*, *35*, 350-357. https://doi.org/10.1016/0014-2921(91)90136-7.

Despotis, D.K. (2005a). A reassessment of the Human Development Index via Data Envelopment Analysis. *Journal of Operational Research Society*, *56*, 969-980. https://doi.org/10.1057/palgrave.jors.2601927.

Despotis, D.K. (2005b). Measuring human development via data envelopment analysis: the case of Asia and the Pacific. *Omega*, *33,* 385-390. https://doi.org/10.1016/j.omega.2004.07.002.

Ding, J., Feng, C., Bi, G., Liang, L., & Khan, R. (2015). Cone ratio models with shared resources and nontransparent allocation parameters in network DEA. *Journal of Productivity Analysis*, 44, 137-155. https://doi.org/10.1007/s11123-014-0420-0.

Doyle, J. R., & Green, R. (1994). Efficiency and cross efficiency in DEA: Derivation, meanings and uses. *Journal of Operational Research Society*, *45*, 567-578. https://doi.org/10.2307/2584392.

Doyle, J. R., & Green, R. (1995). Cross-evaluation in DEA: Improving discrimination among DMUs. *INFOR*, *33*, 205-222. https://doi.org/10.1080/03155986.1995.11732281.

Doyle, J.R., Arthurs, A.J., Green, R. Mcaulay, L., Pitt, M.R., Bottomley, P.A., & Evans, W. (1996). The judge, the model of the judge, and the model of the judged as judge: Analyses of the UK 1992 research assessment business and exercise data for management studies. *Omega*, *24*, 13-28. https://doi.org/10.1016/0305-0483(95)00044-5.

Drake, L., & Hall, M. (2003). Efficiency in Japanese banking: An empirical analysis. *Journal of Banking and Finance*, 27, 891-917. https://doi.org/10.1016/S0378-4266(02)00240-6.

Drake, L., Hall, M., & Simper, R. (2009). Bank modelling methodologies: A comparative non-parametric analysis of efficiency in the Japanese banking sector. *Journal of International Financial Institutions and Money*, 19, 1-15. https://doi.org/10.1016/j.intfin.2007.05.002.

Dyson, R.G., & Thanassoulis E. (1988). Reducing weight flexibility in Data Envelopment Analysis. *Journal of the Operational Research Society*, 39, 563-576. https://doi.org/10.2307/2582861.

EBA (2021). 2021 EU-wide stress tests: Results. Report available at: https://www.eba.europa.eu/risk-analysis-and-data/eu-wide-stress-testing/2021. Accessed April 2021.

Edvardsen, D.F., Førsund, F.R., & Kittelsen, S.A.C. (2008). Far out or alone in the crowd: a taxonomy of peers in DEA. *Journal of Productivity Analysis*, *29*, 201-210. https://doi.org/10.1007/s11123-007-0066-2.

El Gibari, S., Gómez, T., & Ruiz, F. (2019). Building composite indicators using multicriteria methods: a review. *Journal of Business Economics*, *89*, 1-24. https://doi.org/10.1007/s11573-018-0902-z.

Emrouznejad, A., Anouze, A.L., & Thanassoulis, E. (2010). A semi-oriented radial measure for measuring the efficiency of decision making units with negative data, using DEA. *European Journal of Operational Research*, *200*, 297-304. https://doi.org/10.1016/j.ejor.2009.01.001.

Epstein, M.K., & Henderson, J.C. (1989). Data Envelopment Analysis for managerial control and diagnosis. *Decision Sciences* 20, 90-119. https://doi.org/10.1111/j.1540-5915.1989.tb01399.x.

Färe, R., & Karagiannis G. (2014). Benefit-of-the-doubt aggregation and the diet problem. *Omega*, *47*, 33-35. https://doi.org/10.1016/j.omega.2014.03.001.

Färe, R., & Karagiannis, G. (2017). The denominator rule for share-weighting aggregation. *European Journal of Operational Research*, *260*, 1175-1180. https://doi.org/10.1016/j.ejor.2017.02.008.

Färe, R., & Zelenyuk, V. (2003). On aggregate Farrell efficiencies. *European Journal of Operational Research*, *146*, 615-620. https://doi.org/10.1016/S0377-2217(02)00259-X.

Färe, R., Grosskopf, S., & Lovell, C.A.K. (1985). The measurement of efficiency of production (Kluwer-Nijhoff Publishing, Boston). https://doi.org/10.1007/978-94-015-7721-2.

Farrell, M.J. (1957). The measurement of productive efficiency. *Journal of the Royal Statistical Society Series A*, 120, 253-290. https://doi.org/10.2307/2343100.

Fethi, M.D., & Pasiouras, F. (2010). Assessing bank efficiency and performance with operational research and artificial intelligence techniques: A survey. *European Journal of Operational Research,* 204, 189-198. https://doi.org/10.1016/j.ejor.2009.08.003.

Filippetti, A., & Peyrache, A. (2011). The patterns of technological capabilities of countries: A dual approach using composite indicators and Data Envelopment Analysis. *World Development, 39*, 1108-1121. https://doi.org/10.1016/j.worlddev.2010.12.009.

Førsund, F.R. (2017). Measuring effectiveness of production in the public sector. *Omega*, *73,* 93-103. https://doi.org/10.1016/j.omega.2016.12.007.

Førsund, F.R., & Hjalmarsson, L. (1979). Generalised Farrell measures of efficiency: an application to milk processing in Swedish dairy plants, *Economic Journal, 89,* 294-315. https://doi.org/10.2307/2231603.

Fukuda-Parr, S. (2003). The human development paradigm: Operationalizing Sen's ideas on capabilities. *Feminist Economics, 9,* 301-317. https://doi.org/10.1080/1354570022000077980.

Fukuyama, H. (1993). Technical and scale efficiency of Japanese commercial banks: A non-parametric approach. *Applied Economics*, 25, 1101-1112. https://doi.org/10.1080/00036849300000090.

Fukuyama, H., & Sekitani, K. (2012). Decomposing the efficient frontier of the DEA production possibility set into a smallest number of convex polyhedrons by mixed integer programming. *European Journal of Operational Research*, *221*, 165-174. https://doi.org/10.1016/j.ejor.2012.02.035.

Fukuyama, H., & Weber, W.L. (2002). Estimating output allocative efficiency and productivity change: Application to Japanese banks. *European Journal of Operational Research,* 137, 177-190. https://doi.org/10.1016/S0377-2217(01)00054-6.

Gaaloul, H., & Khalfallah, S. (2014). Application of the "Benefit-of-the-Doubt" approach for the construction of a digital access indicator: A revaluation of the "Digital Access Index". *Social Indicators Research, 118,* 45-56. https://doi.org/10.1007/s11205-013-0422-8.

Gerami, J., Mozaffari, M.R., Wanke, P.F., & Correa, H.L. (2022). Improving information reliability of non-radial value efficiency analysis: An additive slacks based measure approach. *European Journal of Operational Research*, 298, 967-978. https://doi.org/10.1016/j.ejor.2021.07.036.

Golany, B., & Thore, S. (1997). Restricted best practice selection in DEA: An overview with a case study evaluating the socio-economic performance of nations. *Annals of Operations Research*, *73*, 117-140. https://doi.org/10.1023/A:1018916925568.

Golany, B., Phillips, F.Y., & Rousseau, J.J. (1993). Models for improved effectiveness based on DEA efficiency results, *IIE Transactions*, *25*, 2-10. https://doi.org/10.1080/07408179308964322.

Gonzalez, E., Carcaba, A., & Ventura, J. (2010). Value efficiency analysis of health systems: does public financing play a role? *Journal of Public Health*, *18*, 337-350. https://doi.org/10.1007/s10389-009-0311-4.

Gonzalez, E., Carcaba, A., & Ventura, J. (2018). Weight constrained DEA measurement of the quality of life in Spanish municipalities in 2011. *Social Indicators Research*, *136*, 1157-1182. https://doi.org/10.1007/s11205-016-1426-y.

Greco, S., Ishizaka, A., Tasiou, M., & Torrisi, G. (2019). On the methodological framework of composite indices: a review of the issues of weighting, aggregation, and robustness. *Social Indicators Research*, *141*, 61-94. https://doi.org/10.1007/s11205-017-1832-9.

Halme, M., & Korhonen, P. (2000). Restricting weights in value efficiency analysis. *European Journal of Operational Research*, *126*, 175-188. https://doi.org/10.1016/S0377-2217(99)00290-8.

Halme, M., & Korhonen, P. (2015). Using value efficiency analysis to benchmark nonhomogeneous units. *International Journal of Information Technology & Decision Making*, *14*, 727-747. https://doi.org/10.1142/S0219622014500916.

Halme, M., Joro, T., Korhonen, P., Salo, S., & Wallenius, T. (1999). A value efficiency approach to incorporating preference information in data envelopment analysis. *Management Science*, *45*, 103-115. https://doi.org/10.1287/mnsc.45.1.103.

Halme, M., Korhonen, P., & Eskelinen, J. (2014). Non-convex value efficiency analysis and its application to bank branch sales evaluation *Omega*, *48*, 10-18 https://doi.org/10.1016/j.omega.2014.04.002.

Hamzeh, R., & Xu, X. (2019). Technology selection methods and applications in manufacturing: A review from 1990 to 2017. *Computers & Industrial Engineering*, *138*, 106123. https://doi.org/10.1016/j.cie.2019.106123.

Hanson, T. (2018). Estimating output mix effectiveness: An applied scenario approach for the Armed Forces. *Omega*, *83*, 39-49. https://doi.org/10.1016/j.omega.2018.02.002.

Hatefi, S.M., & Torabi, S,A, (2010). A common weight MCDA–DEA approach to construct composite indicators. *Ecological Economics*, *70*, 114-120. https://doi.org/10.1016/j.ecolecon.2010.08.014.

Herrero, C., Martínez, R., & Villar, A. (2010). Multidimensional social evaluation: An application to the measurement of human development. *Review of Income and Wealth*, *56*, 483-497. https://doi.org/10.1111/j.1475-4991.2009.00375.x.

Herrero, C., Martínez, R., & Villar, A. (2012) A newer Human Development Index. *Journal of Human Development and Capabilities, 13,* 247-268. https://doi.org/10.1080/19452829.2011.645027.

Herrero, C., Martínez, R., & Villar, A. (2019). Population structure and the Human Development Index. *Social Indicators Research, 141,* 731-763. https://doi.org/10.1007/s11205-018-1852-0.

Hicks, D.A. (1997). The inequality-adjusted human development index: a constructive proposal. *World Development, 25,* 1283-1298. https://doi.org/10.1016/S0305-750X(97)00034-X.

Hirai, T. (2017). *The Human Development Index and its evolution. The creation of the human development approach*. Palgrave Macmillan. https://doi.org/10.1007/978-3-319-51568-7.

Hopkins, M. (1991). Human development revisited: a new UNDP report. *World Development*, *19,* 1469-1473. https://doi.org/10.1016/0305-750X(91)90089-Z.

Hoshi, T., & Kashyap, A.K. (2010). Will the U.S. bank recapitalization succeed? Eight lessons from Japan. *Journal of Financial Economics*, 18, 3-26. https://doi.org/10.1016/j.jfineco.2010.02.005.

Hwang, C.L., & Yoon, K. (1981). *Multiple attribute decision making*, Springer-Verlag, Berlin. https://doi.org/10.1007/978-3-642-48318-9.

Japanese Banking Association (2019). List of members, Japanese Bankers Association, Japan (2019) https://www.zenginkyo.or.jp/en/outline/list-of-members/.

Joro, T., & Korhonen, P. (2015). *Extension of data envelopment analysis with preference information: Value efficiency*. International Series in Operations Research and Management Science 218, Springer. https://doi.org/10.1007/978-1-4899-7528-7.

Joro, T., & Viitala, E.-J. (2004). Weight-restricted DEA in action: from expert opinions to mathematical models. *Journal of the Operational Research Society, 55*, 814-821. https://doi.org/10.1057/palgrave.jors.2601752.

Joro, T., Korhonen, P., & Zionts, S. (2003). An interactive approach to improve estimates of value efficiency in data envelopment analysis. *European Journal of Operational Research, 149,* 688-699. https://doi.org/10.1016/S0377-2217(02)00458-7.

Kahneman, D., & Deaton, A. (2010). High income improves evaluation of life but not emotional well-being. Psychological and cognitive Sciences. *Proceedings of National Academy of Sciences, 107,* 16489–16493. https://doi.org/10.1073/pnas.1011492107.

Kao, C., & Hung, H.T. (2005) Data envelopment analysis with common weights: the compromise solution approach. *Journal of the Operational Research Society, 56*(10), 1196–1203. https://doi.org/10.1057/palgrave.jors.2601924.

Karagiannis, G. (2014). Modeling issues in applied efficiency analysis: agriculture. *Economic and Business Letters, 3,* 12-18. https://doi.org/10.17811/ebl.3.1.2014.12-18.

Karagiannis, G. (2015). On structural and average technical efficiency. *Journal of Productivity Analysis*, *43*, 259-267. https://doi.org/10.1007/s11123-015-0439-x.

Karagiannis, G. (2017) On aggregate composite indicators. *Journal of Operational Research Society*, *68*, 741-746. https://doi.org/10.1057/jors.2015.81.

Karagiannis, G. (2021). DEA Models Without Inputs or Outputs: A Tour de Force. In: C. F. Parmeter & R. C. Sickles (Eds.), *Advances in Efficiency and Productivity Analysis. NAPW 2018* (pp. 211-232), Springer Proceedings in Business and Economics. Springer, Cham. https://doi.org/10.1007/978-3-030-47106-4_10.

Karagiannis, R., & Karagiannis, G. (2020). Constructing composite indicators with Shannon entropy: The case of Human Development Index. *Socio-Economic Planning Sciences*, *70*, 100701. https://doi.org/10.1016/j.seps.2019.03.007

Kelley, A.C. (1991). The human development index: 'handle with care'. *Population and Development Review, 17,* 315-324. https://doi.org/10.2307/1973733.

Khalili, M., Camanho, A.S., Portela, M.C.A.S., & Alirezaee, M.R. (2010a). The measurement of relative efficiency using data envelopment analysis with assurance regions that link inputs and outputs. *European Journal of Operational Research*, *203*, 761-770. https://doi.org/10.1016/j.ejor.2009.09.002.

Khalili, M., Camanho, A.S., Portela, M.C.A.S., & Alirezaee, M.R. (2010b). An improvement on the Tracy and Chen model 'A generalized model for weight restrictions in DEA', *Journal of the Operational Research Society*, *61*, 1789-1793. https://doi.org/10.1057/jors.2009.140.

Khouja, M. (1995). The use of data envelopment analysis for technology selection *Computers & Industrial Engineering* *28*(1), 123-132. https://doi.org/10.1016/0360-8352(94)00032-I.

Kittelsen, S.A.C., & Førsund, F.R. (1992). Efficiency analysis of Norwegian district courts. *Journal of Productivity Analysis, 3,* 277-306. https://doi.org/10.1007/BF00158357.

Klugman, J., Rodríguez, F., & Choi, H.-J. (2011). The HDI 2010: new controversies, old critiques. *Journal of Economic Inequality, 9,* 249–288. https://doi.org/10.1007/s10888-011-9178-z.

Korhonen, P. (1988). A visual reference direction approach to solving discrete multiple criteria problems. *European Journal of Operational Research, 34,* 152-159. https://doi.org/10.1016/0377-2217(88)90350-5.

Korhonen, P. (2002). Searching the efficient frontier in Data Envelopment Analysis. In: Bouyssou, D., Jacquet-Lagrèze, E., Perny, P., Słowiński, R., Vanderpooten, D., & Vincke, P. (eds), *Aiding Decisions with Multiple Criteria*, Kluwer Academic, Boston, 2002. https://doi.org/10.1007/978-1-4615-0843-4_24.

Korhonen, P., & Syrjanen, M.J. (2005). On the Interpretation of Value Efficiency. *Journal of Productivity Analysis, 24,* 197-201. https://doi.org/10.1007/s11123-005-4704-2.

Korhonen, P., & Wallenius, J. (1988). A Pareto race. *Naval Research* Logistics, 35, 615–623. https://doi.org/10.1002/1520-6750(198812)35:6<615::AID-NAV3220350608>3.0.CO;2-K

Korhonen, P., Siljamaki, A., & Soismaa, M. (1998). Practical aspects of value efficiency analysis. IIASA Interim Report IR-98-042/ June.

Korhonen, P., Soismaa, M., & Siljamaki, A. (2002). On the use of value efficiency analysis and some further developments. *Journal of Productivity Analysis*, *17*, 49-65. https://doi.org/10.1023/A:1013532219759.

Korhonen, P., Tainio, R., & Wallenius, J. (2001). Value efficiency analysis of academic research. *European Journal of Operational Research, 130,* 121-132. https://doi.org/10.1016/S0377-2217(00)00050-3.

Kourtzidis, S., Matousek, R., & Tzeremes, N.G. (2021). Modelling a multi-period production process: Evidence from the Japanese regional banks. *European Journal of Operational Research,* 294, 327-339. https://doi.org/10.1016/j.ejor.2021.01.036.

Krishnakumar, J. (2018). Trade-Offs in a multidimensional Human Development Index. *Social Indicators Research, 138,* 991-1022. https://doi.org/10.1007/s11205-017-1679-0

Krivonozhko, V.E., Førsund, F.R., & Lychev, A.V. (2015). Terminal units in DEA: definition and determination. *Journal of Productivity Analysis, 43,* 151-164. https://doi.org/10.1007/s11123-013-0375-6.

Lam, K.F. (2010). In the determination of weight sets to compute cross efficiency ratios in DEA. *Journal of the Operational Research Society*, *61*, 134-143. https://doi.org/10.1057/jors.2008.138.

Lam, K.F., & Bai, F. (2011) Minimizing deviations of input and output weights from their means in data envelopment analysis. *Computers & Industrial Engineering, 60,* 527-533. https://doi.org/10.1016/j.cie.2010.12.007.

Li, S.K., & Ng, Y.C. (1995). Measuring the productive efficiency of a group of firms. *International Advances in Economic Research, 1,* 377–390. https://doi.org/10.1007/BF02295790.

Li, S.K., & Zhao, L. (2015). The competitiveness and development strategies of provinces in China: a data envelopment analysis approach. *Journal of Productivity Analysis, 44,* 293-307. https://doi.org/10.1007/s11123-015-0445-z.

Liang L, Wu J, Cook WD, & Zhu J (2008b). The DEA game cross-efficiency model and its Nash equilibrium. *Operations Research, 56,* 1278–1288. https://doi.org/10.1287/opre.1070.0487.

Liang, L., Wu, J., Cook, W. D., & Zhu, J. (2008a). Alternative secondary goals in DEA cross-efficiency evaluation. *International Journal of Production Economics*, *113*, 1025-1030. https://doi.org/10.1016/j.ijpe.2007.12.006

Lim, S. (2012). Minimax and maximin formulations of cross-efficiency in DEA. *Computers and Industrial Engineering, 62,* 726-731. https://doi.org/10.1016/j.cie.2011.11.010.

Lin, R., Chen, Z., & Xiong, W. (2016). An iterative method for determining weights in cross efficiency evaluation. *Computers & Industrial Engineering*, *101*, 91-102. https://doi.org/10.1016/j.cie.2016.08.024.

Lind, N. (2010). A calibrated index of human development. *Social Indicators Research, 98,* 301-319. https://doi.org/10.1007/s11205-009-9543-5.

Lind, N. (2019). A development of the Human Development Index. *Social Indicators Research*, *146,* 409-423. https://doi.org/10.1007/s11205-019-02133-9.

Liu, H.-H., Song, Y.Y., & Yang, G.L. (2019). Cross-efficiency evaluation in data envelopment analysis based on prospect theory. *European Journal of Operational Research, 273*, 364-375. https://doi.org/10.1016/j.ejor.2018.07.046.

Liu, W., Wang, Y-M, & Lv, S. (2017). An aggressive game cross-efficiency evaluation in data envelopment analysis. *Annals of Operations Research*, *259*, 241-258 https://doi.org/10.1007/s10479-017-2524-1.

Lovell, C.A.K. & Pastor, J.T. (1999). Radial DEA models without inputs or without outputs. *European Journal of Operational Research, 118,* 46-51. https://doi.org/10.1016/S0377-2217(98)00338-5.

Lovell, C.A.K., Pastor, J.T., & Turner, J.A. (1995). Measuring Macroeconomic Performance in the OECD: A Comparison of European and Non-European Countries, *European Journal of Operational Research*, *87*, 507-518. https://doi.org/10.1016/0377-2217(95)00226-X.

Lozano, S., & Gutiérrez, E. (2008). Data envelopment analysis of the human development index. *International Journal of Society Systems Science*, *1*, 132-150. https://doi.org/10.1504/IJSSS.2008.021916.

Luque, M., Perez-Moreno, S., & Rodriguez, B. (2016). Measuring human development: A multi-criteria approach. *Social Indicators Research*, *125*, 713-733. https://doi.org/10.1007/s11205-015-0874-0.

Mahlberg, B., & Obersteiner, M. (2001). Remeasuring the HDI by data envelopment analysis. IIASA interim report IR-01-069, Luxemburg: 2001.

Mariano, E.B., & Rebelatto, D.A.N. (2014). Transformation of wealth produced into quality of life: analysis of the social efficiency of nation-states with the DEA's triple index approach. *Journal of the Operational Research Society*, *65*, 1664-1681. https://doi.org/10.1057/jors.2013.132.

Marshall, E., & Shortle, J. (2005). Using DEA and VEA to evaluate quality of life in the Mid-Atlantic States. *Agricultural and Resource Economics Review*, *34*, 185-203. https://doi.org/10.1017/S1068280500008352.

Mazumdar, K. (2003). A new approach to human development index. *Review of Social Economy*, *61*, 535-549. https://doi.org/10.1080/0034676032000160895.

Mishra, S., & Nathan, H.S.K. (2018). A MANUSH or HUMANS characterization of the Human Development Index. *Journal of Human Development and Capabilities*, *19*, 398-415. https://doi.org/10.1080/19452829.2017.1422703.

Neumayer, E. (2001). The human development index and sustainability-a constructive proposal. *Ecological Economics*, *39*, 101-114. https://doi.org/10.1016/S0921-8009(01)00201-4.

Nguefack-Tsangue, G., Klasen, S., & Zucchini, W. (2011). On weighting the components of the human development index: a statistical justification. *Journal of Human Development and Capabilities*, *12*, 183-202. https://doi.org/10.1080/19452829.2011.571077.

Noorbakhsh, F. (1998a). A modified Human Development Index. *World Development*, *26*, 517-528. https://doi.org/10.1016/S0305-750X(97)10063-8.

Noorbakhsh, F. (1998b). The human development index: some technical issues and alternative indices. *Journal of International Development*, *10*, 589-605. https://doi.org/10.1002/(SICI)1099-1328(199807/08)10:5<589::AID-JID484>3.0.CO;2-S

OECD, *Handbook on constructing composite indicators: Methodology and user guide*, Paris, 2008.

Ogwang, T. (1994). The choice of principal variables for computing the human development index. *World Development, 22,* 2011-2014. https://doi.org/10.1016/0305-750X(94)90189-9.

Ogwang, T., & Abdou, A. (2003). The choice of principal variables for computing some measures of human well-being. *Social Indicators Research, 64,* 139-152. https://doi.org/10.1023/A:1024788123572.

Olesen, O.B. (2018). Cross Efficiency Analysis and Extended Facets. *The Data Envelopment Analysis Journal, 4,* 27-65. http://dx.doi.org/10.1561/103.00000023.

Olesen, O.B., & Petersen, N.C. (1996). Indicators of ill-conditioned data sets and model misspecification in Data Envelopment Analysis: An Extended Facet approach. *Management Science, 42*, 205-219. https://doi.org/10.1287/mnsc.42.2.205.

Olesen, O.B., & Petersen, N.C. (2003). Identification and use of efficient faces and facets in DEA. *Journal of Productivity Analysis, 20,* 323-360. https://doi.org/10.1023/A:1027303901017.

Olesen, O.B., & Petersen, N.C. (2015). Facet Analysis in Data Envelopment Analysis. In: Zhu J (eds.) Data Envelopment Analysis: A Handbook of Models and Methods International Series in Operations Research & Management Science, 221, Springer. https://doi.org/10.1007/978-1-4899-7553-9_6.

Oral, M. (2010). E-DEA: enhanced data envelopment analysis. *European Journal of Operational Research, 207,* 916-926. https://doi.org/10.1016/j.ejor.2010.05.035.

Oral, M. (2012). Action research contextualizes DEA in a multi-organizational decision-making process. *Expert Systems with Applications, 39,* 6503-6513. http://dx.doi.org/10.1016/j.eswa.2011.12.054.

Oral, M., & Yolalan, R. (1990). An empirical study on measuring operating efficiency and profitability of bank branches. *European Journal of Operational Research, 46,* 282-294. https://doi.org/10.1016/0377-2217(90)90002-S.

Oral, M., Amin, G.R., Oukil, A. (2015). Cross-efficiency in DEA: A maximum resonated appreciative model. *Measurement, 63,* 159-167. https://doi.org/10.1016/j.measurement.2014.12.006.

Oral, M., Kettani, O., & Lang, P. (1991). A methodology for collective evaluation and selection of industrial R&D projects. *Management Science, 37,* 871-885. https://doi.org/10.1287/mnsc.37.7.871.

Oral, M., Kettani, O., & Yolalan, R. (1992). An empirical study on analyzing the productivity of bank branches. *Iie Transactions, 24,* 166-176. https://doi.org/10.1080/07408179208964257.

Oral, M., Oukil, A., Malouin, J. L., & Kettani, O. (2014). The appreciative democratic voice of DEA: a case of faculty academic performance evaluation. *Socio-Economic Planning Sciences, 48,* 20-28. https://doi.org/10.1016/j.seps.2013.08.003.

Örkcü, H.H., & Bal, H. (2010). Goal programming approaches for data envelopment analysis cross efficiency evaluation. *Applied Mathematics and Computation, 218,* 346-356. https://doi.org/10.1016/j.amc.2011.05.070.

Oukil, A. (2020). Exploiting value system multiplicity and preference voting for robust ranking. *Omega, 94,* 102048. https://doi.org/10.1016/j.omega.2019.03.006.

Oukil, A., & Govindaluri, S.M. (2017). A systematic approach for ranking football players within an integrated DEA-OWA framework. *Managerial and Decision Economics, 38,* 1125-1136. https://doi.org/10.1002/mde.2851.

Oyinlola, M.A., Adedeji, A.A., & Bolarinwa, M.O. (2020). Exploring the nexus among natural resource rents, human capital and industrial development in the SSA region. *Economic Change and Restructuring, 53,* 87-111. https://doi.org/10.1007/s10644-018-09243-3.

Palazzi, P., & Lauri, A. (1998). The human development index: suggested corrections. *BNL Quarterly Review, 51*, 193-221. https://doi.org/10.13133/2037-3643/10596.

Pan, L., Xu, Z. & Ren, P. (2021): DEA cross-efficiency framework for efficiency evaluation with probabilistic linguistic term sets. *Journal of the Operational Research Society, 72,* 1191-1206. https://doi.org/10.1080/01605682.2020.1848360.

Panigrahi, R., & Sivramkrishna, S. (2002). An adjusted Human Development Index: robust country rankings with respect to the choice of fixed maximum and minimum indicator values. *Journal of Human Development and Capabilities, 3,* 301-311. https://doi.org/10.1080/14649880220147365.

Paradi, J.C., Asmild, M., & Simak, P.C. (2004). Using DEA and Worst Practice DEA in Credit Risk Evaluation. *Journal of Productivity Analysis,* 21, 153-165. https://doi.org/10.1023/B:PROD.0000016870.47060.0b.

Pinar, M., Stengos, T., & Topaloglou, N. (2013). Measuring human development: a stochastic dominance approach. *Journal of Economic Growth, 18,* 69-108. https://doi.org/10.1007/s10887-012-9083-8.

Pinar, M., Stengos, T., & Topaloglou, N. (2017). Testing for the implicit weights of the dimensions of the Human Development Index using stochastic dominance. *Economics Letters, 161*, 38-42. https://doi.org/10.1016/j.econlet.2017.09.023.

Podinovski, V.V. (2004) Production trade-offs and weight restrictions in data envelopment analysis. *Journal of the Operational Research Society, 55*(12), 1311-1322. https://doi.org/10.1057/palgrave.jors.2601794.

Podinovski, V.V. (2005) The explicit role of weight bounds in models of data envelopment analysis. *Journal of the Operational Research Society 56*(12), 1408-1418. https://doi.org/10.1057/palgrave.jors.2601969.

Podinovski, V.V. (2007a). Improving data envelopment analysis by the use of production trade-offs. *Journal of the Operational Research Society, 58*(10), 1261-1270. https://doi.org/10.1057/palgrave.jors.2602302.

Podinovski, V.V. (2007b). Computation of efficient targets in DEA models with production trade-offs and weight restrictions. *European Journal of Operational Research 181*(2), 586-591. https://doi.org/10.1016/j.ejor.2006.06.041.

Podinovski, V.V., & Bouzdine-Chameeva, T. (2013). Weight restrictions and free production in data envelopment analysis. *Operations Research 61*(2), 426-437. https://doi.org/10.1287/opre.1120.1122.

Podinovski, V.V., & Førsund, F. R. (2010). Differential characteristics of efficient frontiers in data envelopment analysis. *Operations Research, 58*(6), 1743-1754. https://doi.org/10.1287/opre.1090.0794.

Portela, M.C.A.S (2014). Value and quantity data in economic and technical efficiency measurement. *Economics Letters,* 124, 108-112. https://doi.org/10.1016/j.econlet.2014.04.023.

Portela, M.C.A.S, & Thanassoulis, E. (2006). Zero weights and non-zero slacks: Different solutions to the same problem. *Annals of Operations Research, 145,* 129-147. https://doi.org/10.1007/s10479-006-0029-4.

Prados de la Escosura, L. (2010). Improving human development: A long-run view. *Journal of Economic Surveys*, *24*, 841-894. https://doi.org/10.1111/j.1467-6419.2010.00639.x

Prieto, A.M., & Zofio, J.L. (2001). Evaluating effectiveness in public provision of infrastructure and equipment: the case of Spanish municipalities. *Journal of Productivity Analysis, 15,* 41-58. https://doi.org/10.1023/A:1026595807015.

Raheem, I.D., Isah, K.O., & Adedeji, A.A. (2018). Inclusive growth, human capital development and natural resource rent in SSA. *Economic Change and Restructuring*, *51*, 29-48. https://doi.org/10.1007/s10644-016-9193-y.

Ramon, N., Ruiz, J.L., & Sirvent, I. (2010). On the choice of weights profiles in cross-efficiency evaluations *European Journal of Operational Research*, *207*, 1564-1572. https://doi.org/10.1016/j.ejor.2010.07.022.

Ramon, N., Ruiz, J.L., & Sirvent, I. (2011). Reducing differences between profiles of weights: a ‘‘peer-restricted cross-efficiency evaluation. *Omega*, *39*, 634-641. https://doi.org/10.1016/j.omega.2011.01.004.

Ranis, G., Stewart, F. & Ramirez, A. (2000). Economic growth and human development. *World Development, 28*, 197-219. https://doi.org/10.1016/S0305-750X(99)00131-X.

Ranis, G., Stewart, F., & Samman, E. (2006). Human development: beyond the human development index. *Journal of Human Development and Capabilities, 7*, 323-358. https://doi.org/10.1080/14649880600815917.

Räty, T. (2002). Efficient facet based efficiency index: A variable returns to scale specification. *Journal of Productivity Analysis* 17, 65-82. https://doi.org/10.1023/A:1013584203829.

Ravallion, M. (2012). Troubling tradeoffs in the Human Development Index. *Journal of Development Economics, 99*, 201-209. https://doi.org/10.1016/j.jdeveco.2012.01.003.

Rödder, W., & Reucher, E. (2012). Advanced X-efficiencies for CCR- and BCC-models-towards Peer-based DEA controlling. *European Journal of Operational Research, 219*, 467-476. https://doi.org/10.1016/j.ejor.2011.12.024.

Rogge N, Self R (2019) Measuring regional social inclusion performances in the EU: Looking for unity in diversity. *Journal of European Social Policy* 29:325-344. https://doi.org/10.1177/0958928718792135.

Rogge, N. (2012). Undesirable specialization in the construction of composite policy indicators: The Environmental Performance Index. *Ecological Indicators, 23*, 143-154. https://doi.org/10.1016/j.ecolind.2012.03.020.

Roll, Y., Cook, W.D., & Golany, B. (1991). Controlling factor weights in Data Envelopment Analysis. *IIE Transactions, 23*, 2-9. https://doi.org/10.1080/07408179108963835.

Ruiz, J.L., & Sirvent, I. (2019). Performance evaluation through DEA benchmarking adjusted to goals. *Omega, 87*, 150-157. https://doi.org/10.1016/j.omega.2018.08.014.

Saaty, T.L. (1988) Multicriteria Decision Making: The Analytic Hierarchy Process, RWS Publications, Pittsburgh, PA.

Sagar, A.D., & Najam, A. (1998). The human development index: a critical review. *Ecological Economics, 25*, 249-264. https://doi.org/10.1016/S0921-8009(97)00168-7.

Sahoo, B.K., & Acharya, D. (2010). An alternative approach to monetary aggregation in DEA. *European Journal of Operational Research, 204*(3), 672-682. https://doi.org/10.1016/j.ejor.2009.11.035.

Saisana, M., Saltelli, A., & Tarantola, S. (2005). Uncertainty and sensitivity analysis techniques as tools for quality assessment of composite indicators. *Journal of the*

Royal Statistical Society Series A, *168*, 307-323. https://doi.org/10.1111/j.1467-985X.2005.00350.x.

Santos, S.P., Amado, C.A.F., & Rosado, J.R. (2011). Formative evaluation of electricity distribution utilities using data envelopment analysis. *Journal of the Operational Research Society 62*(7), 1298-1319. https://doi.org/10.1057/jors.2010.66.

Sarrico, C.S., & Dyson, R.G. (2004). Restricting virtual weights in data envelopment analysis. *European Journal of Operational Research*, *159*, 17-34. https://doi.org/10.1016/S0377-2217(03)00402-8.

Sayed, H., Hamed, R., Ramadan, M.A.-G., & Hosny, S. (2015). Using Meta-goal Programming for a New Human Development Indicator with Distinguishable Country Ranks. *Social Indicators Research*, *123*, 1-27. https://doi.org/10.1007/s11205-014-0723-6.

Seiford, L.M., & Zhu, J. (1999). Infeasibility of super-efficiency Data Envelopment Analysis Models. *INFOR: Information Systems and Operational Research*, *37*, 174-187. https://doi.org/10.1080/03155986.1999.11732379.

Sen, A. (1983). Development: Which way now? *The Economic Journal*, *93*, 745-762. https://doi.org/10.2307/2232744.

Sen, A. (1993). Capability and well-being. In: Nussbaum, M., & Sen, A. (Eds.), The Quality of Life. Clarendon, Oxford. https://doi.org/10.1093/0198287976.001.0001.

Sen, A. (1999). Development as Freedom. Oxford University Press, Oxford.

Sen, A. (2003). Development as Capability Expansion. In: Fukuda-Parr, S., Shiva Kumar, A.K. (eds.) Readings in Human Development, Concepts, Measures and Policies for a Development Paradigm. Oxford University Press, New York.

Sexton, T. R., Silkman, R. H., & Hogan, A. J. (1986). Data envelopment analysis: Critique and extensions. In R. H. Silkman (Ed.), *Measuring Efficiency: An Assessment of Data Envelopment Analysis* (pp. 73-105). Jossey-Bass. https://doi.org/10.1002/ev.1441.

Shang, J., & Sueyoshi, T. (1995). A unified framework for the selection of a flexible manufacturing system. *European Journal of Operational Research*, *85*, 297-315. https://doi.org/10.1016/0377-2217(94)00041-A.

Shi, H.-L., Chen S.-Q., Chen L., & Wang, Y.-M. (2021). A neutral cross-efficiency evaluation method based on interval reference points in consideration of bounded rational behavior. *European Journal of Operational Research*, *290*, 1098–1110 https://doi.org/10.1016/j.ejor.2020.08.055

Shi, H.-L., Wang, Y.-M, & Chen L., &. (2019). Neutral cross-efficiency evaluation regarding an ideal frontier and anti-ideal frontier as evaluation criteria. *Computers*

& Industrial Engineering, *132*, 385-394. https://doi.org/10.1016/j.cie.2019.04.035

Soares de Mello, J.C.C.B., Angulo Meza, L. A., Quintanilla da Silveira, J., & Gonzales Gomes, E. (2013). About negative efficiencies in cross efficiency BCC input oriented models. *European Journal of Operational Research, 229,* 732-737. https://doi.org/10.1016/j.ejor.2013.02.020.

Srinivasan, T.N. (1994). Human development: A new paradigm or reinvention of the wheel? *American Economic Review, 84,* 238-243. https://www.jstor.org/stable/2117836

Stapleton, L.M., & Garrod, G.D. (2007). Keeping things simple: why the Human Development Index should not diverge from its equal weights assumption. *Social Indicators Research*, *84*, 179-188. https://doi.org/10.1007/s11205-006-9081-3.

Stewart, T.J. (2010). Goal directed benchmarking for organizational efficiency. *Omega*, *38*, 534-539. https://doi.org/10.1016/j.omega.2010.01.004.

Suri, T., Boozer, M.A., Ranis, G. & Stewart, F. (2011). Paths to success: The relationship between human development and economic growth. *World Development*, 39, 506-522. https://doi.org/10.1016/j.worlddev.2010.08.020.

Talluri, S., & Yoon, K.P. (2000). A cone-ratio DEA approach for AMT justification. *International Journal of Production Economics*, 66, 119-129. https://doi.org/10.1016/S0925-5273(99)00123-1.

Thanassoulis, E., Portela, M.C.S, & Despíc, O. (2008). Data envelopment analysis: the mathematical programming approach to efficiency analysis. In: H.O. Fried, C.A.K. Lovell, & S.S. Schmidt (Eds.) *The measurement of productive efficiency and productivity growth* (pp. 251-420). Oxford University Press, New York. https://doi.org/10.1093/acprof:oso/9780195183528.003.0003.

Thanassoulis, E., Portela, M.C.S., & Allen, R. (2004). Incorporating value judgements in DEA. In: W.W. Cooper, L.M. Seiford & J. Zhu (Eds.). *Handbook on data envelopment analysis* (pp. 99-138). Kluwer Academic, Boston. https://doi.org/10.1007/1-4020-7798-X_4.

Thiele, L., Miettinen, K., Korhonen, P.J., & Molina, J. (2009). A preference-based evolutionary algorithm for multi-objective optimization. *Evolutionary Computation*, *17*, 411-436. https://doi.org/10.1162/evco.2009.17.3.411.

Thompson, R.G., Dharmapala, P.S., & Thrall, R.M. (1995). Linked-cone DEA profit ratios and technical efficiency with application to Illinois coal mines. *International Journal of Production Economics*, 39, 99-115. https://doi.org/10.1016/0925-5273(94)00064-H.

Thompson, R.G., Dharmapala, P.S., Humphrey, D.B., Taylor, W.L. & Thrall, R.M. (1996). Computing DEA/AR efficiency and profit ratio measures with an illustrative bank application. *Annals of Operations Research*, 68, 301-327. https://doi.org/10.1007/BF02207220.

Thompson, R.G., Dharmapala, P.S., Rothenberg, L.J., & Thrall, R.M. (1994). DEA ARs and CRs applied to worldwide major oil companies. *Journal of Productivity Analysis*, 5, 181-203. https://doi.org/10.1007/BF01073854.

Thompson, R.G., Langemeier, L.M., Lee, C.-T., Lee, E., & Thrall, R.M. (1990). The role of multiplier bounds in efficiency analysis with application to Kansas farming. *Journal of Econometrics, 46*, 93-108. https://doi.org/10.1016/0304-4076(90)90049-Y.

Thompson, R.G., Singleton, F., Thrall, R., & Smith, B. (1986). Comparative site evaluations for locating a high-energy physics lab in Texas. *Interfaces, 16*, 35-49. https://doi.org/10.1287/inte.16.6.35.

Tofallis, C. (2013). An automatic-democratic approach to weight setting for the new human development index. *Journal of Population Economics*, 26, 1325-1345. https://doi.org/10.1007/s00148-012-0432-x.

Tofallis, C. (2014). On constructing a composite indicator with multiplicative aggregation and the avoidance of zero weights in DEA. *Journal of the Operational Research Society, 65*, 791-792. https://doi.org/10.1057/jors.2013.137.

Tone., K. (1997). Several algorithms to determine multipliers for use in cone-ratio envelopment approaches to efficiency evaluations in DEA In: Amman, H., Rustem, B., Whinston, A. (eds). *Computational Approaches to Economic Problems* (pp. 91-109). Advances in Computational Economics, 6. Springer, Boston, MA. https://doi.org/10.1007/978-1-4757-2644-2_7.

Torgensen, A.M., Førsund, F.R., & Kittelsen, S.A.C. (1996). Slack-adjusted efficiency measures and ranking of efficient units. *Journal of Productivity Analysis*, 7, 379-398. https://doi.org/10.1007/BF00162048.

Tracy, D.L., & Chen, B. (2005). A generalized model for weight restrictions in data envelopment analysis, *Journal of the Operational Research Society, 56*, 390-396. https://doi.org/10.1057/palgrave.jors.2601823.

UNDP (1990). Human Development Report 1990. New York: Oxford University Press.

UNDP (1994). *Human Development Report 1994*, New York: Oxford University Press.

UNDP (2015). Human Development Report 2015, New York: Palgrave Macmillan.

UNDP (2016). Human Development Report 2016, New York: Palgrave Macmillan.

UNDP (2018). Human Development Report 2018, New York: Palgrave Macmillan.

UNDP (2019). Human Development Report 2019, New York: Palgrave Macmillan.

Van Puyenbroeck, T. (2018) On the Output Orientation of the Benefit-of-the-Doubt-Model. *Social Indicators Research, 139,* 1-17. https://doi.org/10.1007/s11205-017-1734-x.

Van Puyenbroeck, T., & Rogge, N. (2017). Geometric mean quantity index numbers with Benefit-of-the-Doubt weights. *European Journal of Operational Research*, *256*, 1004-1014. https://doi.org/10.1016/j.ejor.2016.07.038.

Wang, Y.M., & Chin, K.S. (2010a). A neutral DEA model for cross-efficiency evaluation and its extension. *Expert Systems with Applications*, *37*, 3666-3675. https://doi.org/10.1016/j.eswa.2009.10.024.

Wang, Y.M., & Chin, K.S. (2010b). Some alternative models for DEA cross-efficiency evaluation. *International Journal of Production Economics*, *128*, 332-338. https://doi.org/10.1016/j.ijpe.2010.07.032.

Wang, Y.M., & Luo, Y. (2006). DEA efficiency assessment using ideal and anti-ideal decision making units. *Applied Mathematics and Computation*, 1*73*, 902-915. https://doi.org/10.1016/j.amc.2005.04.023.

Wang, Y.M., Chin, K. S., & Jiang, P. (2011a). Weight determination in the cross-efficiency evaluation. *Computers and Industrial Engineering*, *61*, 497-502. https://doi.org/10.1016/j.cie.2011.04.004.

Wang, Y.M., Chin, K. S., & Luo, Y. (2011b). Cross-efficiency evaluation based on ideal and anti-ideal decision making units. *Expert Systems with Applications*, *38*, 10312-10319. https://doi.org/10.1016/j.eswa.2011.02.116.

Wang, Y.M., Chin, K.S., & Wang, S. (2012). DEA models for minimizing weight disparity in cross-efficiency evaluation. *Journal of the Operational Research Society*, *63*, 1079-1088. https://doi.org/10.1057/jors.2011.116.

Wilson, P.W. (1995). Detecting influential observations in Data Envelopment Analysis. *Journal of Productivity Analysis*, *6*, 27-45. https://doi.org/10.1007/BF01073493.

Wong, Y.H.B., & Beasley, J.E. (1990). Restricting weight flexibility in data envelopment analysis. *Journal of Operational Research Society*, *41*, 829-835. https://doi.org/10.1057/jors.1990.120.

Wu, J., Chu, J., Sun, & Zhu, Q. (2016c) DEA cross-efficiency evaluation based on Pareto improvement. *European Journal of Operational Research*, *248*, 571-579. https://doi.org/10.1016/j.ejor.2015.07.042

Wu, J., Chu, J., Sun, J., Zhu, Q., & Liang, L. (2016a). Extended secondary goal models for weights selection in DEA cross-efficiency evaluation. *Computers and Industrial Engineering*, *93*, 143-151. https://doi.org/10.1016/j.cie.2015.12.019.

Wu, J., Chu, J., Zhu, Q., Yin, P. & Liang, L. (2016b). DEA cross-efficiency evaluation based on satisfaction degree: an application to technology selection. *International Journal of Production Research*, *54*, 5990–6007. http://dx.doi.org/10.1080/00207543.2016.1148278.

Wu, J., Liang, L., & Chen, Y. (2009b). DEA game cross-efficiency approach to Olympic rankings. *Omega 37*, 909-918. https://doi.org/10.1016/j.omega.2008.07.001.

Wu, J., Liang, L., Yang, F. & Yan, H. (2009c). Bargaining game model in the evaluation of decision making units. *Expert Systems with Applications*, *36*, 4357-4362. https://doi.org/10.1016/j.eswa.2008.05.001.

Wu, J., Liang, L., Zha, Y., & Yang, F. (2009a). Determination of cross-efficiency under the principle of rank priority in cross-evaluation. *Expert Systems with Applications, 36*, 4826-4829. https://doi.org/10.1016/j.eswa.2008.05.042.

Yamada, Y., Matui, T., & Sugiyama, M. (1994). An inefficiency measurement method for management systems. *Journal of the Operations Research Society of Japan*, 37:158-167 (in Japanese). https://doi.org/10.15807/jorsj.37.158.

Yang, F., Ang, S., Xia, Q., & Yang, S. (2012). Ranking DMUs by using interval DEA cross efficiency matrix with acceptability analysis. *European Journal of Operational Research, 223*, 483-488. https://doi.org/10.1016/j.ejor.2012.07.001.

Yang, J.-B., Wong, B.Y.H., Xu, D.-L., & Stewart, T.J. (2009). Integrating DEA-oriented performance assessment and target setting using interactive MOLP methods. *European Journal of Operational Research*, 195, 205-222. https://doi.org/10.1016/j.ejor.2008.01.013.

Yu, M.M., & Lin, E.T.J. (2008). Efficiency and effectiveness in railway performance using a multi-activity network DEA model. *Omega*, 36, 1005-1017. https://doi.org/10.1016/j.omega.2007.06.003.

Zanella, A, Camanho, A.S, & Dias, T.G. (2013). Benchmarking countries' environmental performance. *Journal of the Operational Research Society 64*, 426-438. https://doi.org/10.1057/jors.2012.62.

Zhou, P., Ang, B.W., & Poh, K.L. (2007). A mathematical programming approach to constructing composite indicators. *Ecological Economics*, *62*, 291-297. https://doi.org/10.1016/j.ecolecon.2006.12.020.

Zhou, P., Ang, B.W., & Zhou, D.Q. (2010). Weighting and aggregation in composite indicator construction: a multiplicative optimization approach. *Social Indicators Research*, 96, 169-181. https://doi.org/10.1007/s11205-009-9472-3.

Zhu, J. (1996). DEA/AR analysis of the 1988-1989 performance of the Nanjing textiles corporation. *Annals of Operations Research*, 66, 311-335. https://doi.org/10.1007/BF02188949.

Zhu, J. (2001) Multidimensional quality-of-life measure with an application to Fortune's best cities. *Socio-Economic Planning Sciences*, 35, 263-284. https://doi.org/10.1016/S0038-0121(01)00009-X.