TÉCNICO LISBOA

Preference information incorporation for decision-making in a DEA model using the Choquet integral

Performance assessment of secondary healthcare providers

Miguel Alves Pereira

Dissertation to obtain the Master of Science Degree in

Biomedical Engineering

Supervisors: Prof. Rui Domingos Ribeiro da Cunha Marques
Prof. José Rui de Matos Figueira

Examination Committee

Chairperson: Prof. João Pedro Estrela Rodrigues Conde
Supervisor: Prof. Rui Domingos Ribeiro da Cunha Marques
Member of the Committee: Prof. Ana Maria Cunha Ribeiro dos Santos Ponces Camanho

May 2018

Dedicated to me (for my ghastly endurance over the last few years), my family (for always being there for me, in their different ways), and the love of my life (for being my present and, most of all, my future) in chronological order, so there is no hassle.

Declaration

I declare that this document is an original work of my own authorship and that it fulfils all the requirements of the Code of Conduct and Good Practices of the Universidade de Lisboa.

Acknowledgments

The completion of this dissertation was only possible thanks to the encouragement and support of a legion of individuals.

Foremost, I would like to express my gratitude towards my entire family as a whole, for being a source of strength my entire life, and each member in its own fashion, for the inestimable life lessons. Exceptionally, to my mother (for her positively-intended annoyance in pursuance of me focusing my efforts), to my father (for his nonstop buzzkill for the sake of me keeping my feet on the ground), and to my brother (for his everlasting illusive nonchalance beneficial to my personal motivation), my eternal praise and assurance that the journey is just commencing.

Next in order, I would like to bow down to my girlfriend, who picked me up (pun intended) when times were bad and has passed all possible tests to uphold her vital part as my lifelong *innamorata numero uno*.

Undoubtedly, I would also like to recognise the invaluable assistance of my supervisors, Prof. Rui Cunha Marques and Prof. José Rui Figueira, for their guidance and, whether they are aware of it or not, care in lifting the spirits of a once disheartened master's student and leading him to pursue a long-dreamt and now foreseeable future in academic research.

By the same token, a special gesture of appreciation to my office buddies, Diogo Ferreira and José Oliveira, for their hospitality, spirituous insights, and expertise for getting me out of 'keep me on my toes' messes. In particular, Diogo deserves a sole paragraph praising his help and nourishment of my inexperienced and blossoming mind - if it were not for him, I would still be wallowing in methodological mud -, but he is soon to be a Doctor of Philosophy, so my words had to be suitably measured.

At last, but certainly not least, my friends. Coming to Instituto Superior Técnico was an experience that left a deep-rooted impression on me, that is certain. However, not only did this adventure provide me with know-how and shaped me as a forthcoming professional, but also gave me the opportunity to meet quite a handful of generous, humane, and brilliant people whom I can call my friends. João Franco, João Perdiz, Miguel Vieira, Francisco Sargo, Pedro Borges, Miguel Martinho, Mafalda Prazeres, Joana Gonçalves, and so many others who were with me since the very beginning of this enterprise, were the best sidekicks my superhero could have asked for.

Abstract

In a world in permanent (r)evolution that revolves around money, seeking new ways to contain costs, better allocate resources, and, overall, improve performance is a constant across all the fields. Hence, the usage of computational methods based on operational research and statistical science is crucial for achieving an appropriate combination of efficiency and effectiveness, especially in domains where the decision-making process is a complex task. This is where Data Envelopment Analysis (DEA) comes in. However, as a non-parametric and, usually, purely objective technique, DEA makes up for in flexibility and adaptability what it lacks in incorporating preference information, which is particularly important in areas where the judgment of decision-makers is crucial. This work proposes a cutting-edge and original approach to fill in this knowledge gap by linking DEA and multi-criteria decision-making with an additive DEA model that takes into consideration criteria interactivity, by using an inference methodology to determine their weights, and decision-makers' preference information incorporation, by taking advantage of the Choquet multi-criteria preference aggregation model. Thus, PRICDEA (PReference information Incorporation using the Choquet integral in a DEA model) was born, generating credible weights stemmed from the decision-maker's judgments, and yielding acceptable and valid results in a tailor-made case study of performance assessment of Portuguese National Healthcare Service secondary healthcare providers across robustness-testing scenarios.

Keywords: Multi-criteria decision-making, Decision analysis, Data Envelopment Analysis, Choquet integral, Interactive criteria, Preference information incorporation

Resumo

Num mundo em permanente (r)evolução que gira à volta de dinheiro, a procura por novas formas de conter custos, melhor alocar recursos, e, sobretudo, melhorar o desempenho é uma constante transversal a todos os campos. Por conseguinte, a utilização de métodos computacionais baseados em investigação operacional e ciência estatística é crucial para alcançar uma combinação apropriada de eficiência e eficácia, especialmente em domínios onde o processo de tomada de decisão é uma tarefa complexa. É aqui que entra a *Data Envelopment Analysis* (DEA). Contudo, como uma técnica não-paramétrica e, usualmente, puramente objetiva, a DEA compensa em flexibilidade e adaptabilidade o que carece em incorporar a informação de preferências, particularmente importante em áreas onde o julgamento de decisores é crucial. Este trabalho propõe uma abordagem pioneira e original para preencher esta lacuna de conhecimento através da ligação entre a DEA e a tomada de decisão multicritério com um modelo aditivo de DEA que tem em consideração a interatividade de critérios, usando uma metodologia de inferência para determinar os seus pesos, e a incorporação da informação das preferências de decisores, tirando partido do modelo agregador de preferências multicritério de Choquet. Assim, nasceu o PRICDEA (*PReference information Incorporation using the Choquet integral in a DEA model*), gerando pesos credíveis a partir dos julgamentos de um decisor, e produzindo resultados aceitáveis e válidos num caso de estudo feito à medida sobre avaliação de desempenho dos prestadores de cuidados de saúde secundários do Serviço Nacional de Saúde português por meio de cenários testadores de robustez.

Palavras-chave: Tomada de decisão multicritério, Análise de decisão, *Data Envelopment Analysis*, Integral de Choquet, Critérios interativos, Incorporação da informação de preferências

Contents

List of Tables

List of Figures

Nomenclature

Greek symbols

λ Lagrange multiplier.

μ Choquet capacity.

ε Non-Archimedean infinitesimal.

Roman symbols

C_μ Choquet integral.

a Alternative.

d Distance between two alternatives.

g Indicator.

m Möbius coefficient.

s Slack.

$\mathbf{M}$ Möbius coefficients' vector.

$\mathbf{U}$ Utility matrix.

$\mathbf{z}$ Efficiency score vector.

Subscripts

i, j, k, l PRICDEA's Phase 1 computational indexes per indicator.

i, k PRICDEA's Phase 2 computational indexes per alternative j.

j PRICDEA's Phase 2 computational indexes per alternative.

x, y PRICDEA's Phase 1 computational indexes per alternative.

Glossary

ACSS	*Administração Central do Sistema de Saúde* is a Portuguese public institute integrated in the indirect State health administration.
ADD	Additive DEA model is a DEA model that combines both input reduction and output augmentation simultaneously.
CPE	Corporate Public Entity is a public law collective person with a corporate nature, created by the State, but administered by a third party indicated by it.
CRS	Constant Returns-to-Scale is a DEA scaling assumption that reflects the fact that outputs will change in the same proportion as inputs.
DEA	Data Envelopment Analysis is a non-parametric linear-programming-based technique for measuring the relative performance of organisational units in a context of multiple and outputs.
DMU	Decision-Making Unit is an entity subject to evaluation in a decision-making process.
GDP	Gross Domestic Product is the monetary value of all final goods and services produced within a country's borders in a certain period of time.
IPO	*Instituto Português de Oncologia* is a Portuguese institute that regionally provides specialised healthcare in the oncology domain.
LSR	Least Squares Regression is a regression analysis' standard approach that approximates the solution of overdetermined systems.

MAUT Multi-Attribute Utility Theory is a methodology structured to deal with the tradeoffs among multiple objectives.

MCDM Multi-Criteria Decision-Making is a sub-discipline of operational research that explicitly evaluates multiple conflicting criteria for decision-making purposes.

MOLP Multi-Objective Linear Programming is a sub-area of mathematical optimisation whose linear programs have more than one objective function.

NIRS Non-Increasing Returns-to-Scale is a constant added to a VRS DEA model that intends to assess if a DMU is operating in the area of increasing or decreasing returns-to-scale.

PRICDEA PRICDEA is an innovative revamped additive DEA model for multi-criteria decision-making that accounts for interactive criteria and incorporates the preference information of a decision-maker using the Choquet integral.

SFA Stochastic Frontier Analysis is a parametric method of economic modelling that produces efficiency measures using a frontier approach.

SNS *Serviço Nacional de Saúde* is the Portuguese State structure that assures the right of all citizens to health.

TFP Total Factor Productivity is an efficiency measure that accounts for the unexplained portion of outputs by traditionally measured labour and capital inputs used in production.

VRS Variable Returns-to-Scale is another DEA scaling assumption that echoes the fact that production technology may exhibit increasing, constant and decreasing returns-to-scale.

Chapter 1

Introduction

1.1 Motivation

In life, nothing that has value, real value, has no cost. When Dean Kamen, one of the world's most prolific inventors of healthcare technologies (best-known for his invention of the Segway), said this in an interview to an acclaimed science and technology magazine, Popular Mechanics, back in 2009, he was referring, *inter alia*, to healthcare (Meigs & Beilinson, 2009). The American engineer also stated that putting money into the economy, even at a time of crisis, should target areas where problems would be solved, diseases would be cured, and technologies would be created, in order to improve the well-being of the population. In fact, the founder of DEKA Research & Development Corporation is on the right path with his ideas and he is not alone, at least not in his own country, inasmuch that the United States of America's total healthcare spending exceeded 3.2$ trillion in 2015 (National Center for Health Statistics, 2017). Nevertheless, notwithstanding the fact that the "land of opportunity" is one of the largest countries in the world, its healthcare system is somewhat of a trump card in the sense that, when compared to countries with more established legislation regarding this topic, the American case stands on the side of such nations, among which Portugal is an exquisite example.

Portugal, in spite of being a medium-sized country and having always been of minor significance in most circumstances in the European Union frame of reference, has one of the oldest healthcare systems in the world (it dates back to the late 1970s in the era following the revolution that dethroned the *Estado Novo* corporatist authoritarian regime). Nowadays, in line with the previous reasoning, its total healthcare spending reached over 16€ billion in 2015 (Instituto Nacional de Estatística, 2015). Effectively, according to Portugal's National Statistics' Institute, 8.9% of Portugal's gross domestic product (GDP) is devoted to healthcare, placing the westernmost country in mainland Europe among the top healthcare spenders as a percentage of GDP in the European Union, which weakens Portuguese-based business' global competitiveness. As almost 10% of GDP is spent on healthcare year after year, the National Health Service (SNS, from the Portuguese abbreviation of *Serviço Nacional de Saúde*) is being pushed to a limit where the stress of high costs forces policy-makers and healthcare managers to seek new ways to improve efficiency, contain costs, and maintain quality of care. To that end, one can claim that improving the efficiency of healthcare is one of the most important management challenges of the twenty-first century (Ozcan, 2008).

In this context, this dissertation places emphasis on the application of a well-known performance measurement technique, *sc.* data envelopment analysis (DEA), that, alongside the Choquet integral as an aggregation tool that takes into account interaction phenomena among criteria, develops into a state-of-the-art methodology acknowledging preference information incorporation that infers the weights of the preceding interactive criteria from a reference set via a mathematical programming linear optimisation for decision-making in a context of performance assessment of SNS secondary healthcare providers in Portugal.

1.2 Objectives

This dissertation addresses a long-lasting optimisation problem in the Portuguese SNS, specifically in what concerns secondary healthcare providers, and proposes an *avant-garde* approach for including interactive variables and incorporating stakeholders' preferences in order to infer personalised weights that allow the improvement of institutional performance using an overhauled DEA model. On that account, a miscellany of sequentially arranged objectives to be accomplished can be itemised, *viz.*:

- Have a grasp on the circumstances of health in Portugal in a political, economical, and social sense;

- Recognise DEA's advantages and disadvantages as a non-parametric technique, whose contemporaneity and suitability for performance measurement is utterly useful, particularly in the healthcare context;

- Identify the Choquet multiple criteria preference aggregation model as a convenient and fitting tool for incorporating the information of decision-makers' preferences and dealing with interactive criteria;

- Conceive an innovative Choquet integral-based DEA method for performance evaluation that infers the criteria weights and takes into account interactive variables and preference information incorporation for decision-making in healthcare;

- Apply the envisaged model to a case study in the healthcare sector and, after comparing the results obtained with the aforementioned weights and the weights computed using solely the decision-maker's judgments, withdraw applicable recommendations for healthcare managers and policy-makers.

These goals can be rearranged in a more perceptible manner towards the ultimate purpose of this dissertation, as Figure 1.1 demonstrates.

The accomplishment of these aspirations will be addressed in Section 6.1.

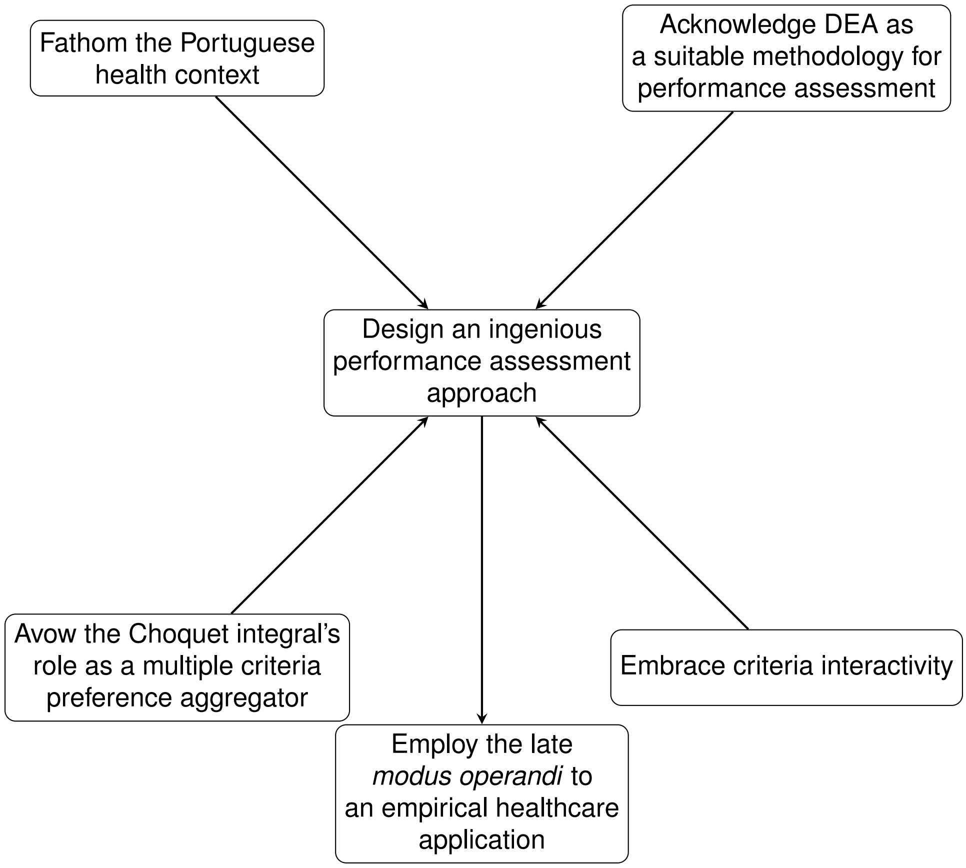

Figure 1.1: Dissertation objectives.

1.3 Research methodology

In pursuance of the aspirations listed in Section 1.2, the stages schematised in Figure 1.2 were considered as a procedural plan to take on this monumental undertaking.

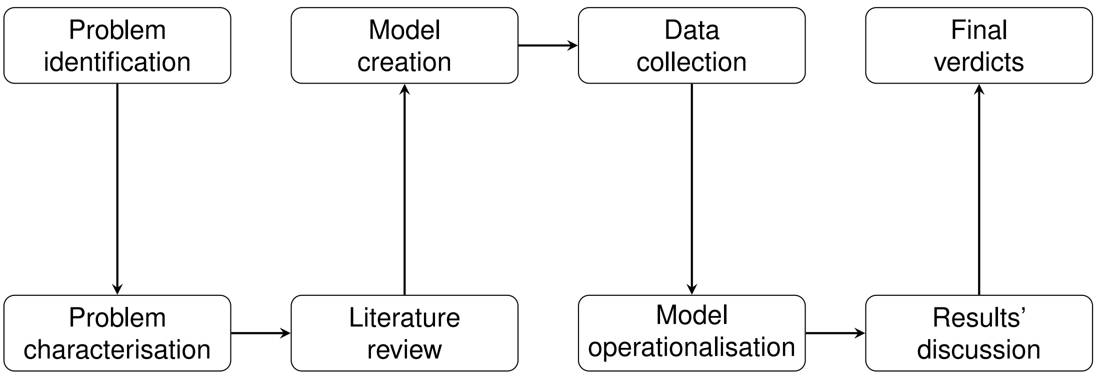

Figure 1.2: Research methodological steps.

More explicitly:

1. The first node consists on the recognition and establishment of the quandary at hand by digging into the background of health in Portugal;

2. The problem is then described and detailed in the next step, being given a fitting framework and

mise en scène;

3. Once the dilemma is authenticated, a literature review phase is required in the interest of erecting a solid theoretical basis in healthcare performance evaluation;

4. With such reliable foundations concerning the previous stage, creating a model that suits the identified problem (inasmuch as there is no other model cut out for the predicament) is in order;

5. Logically, the following block of the diagram is concerned with populating the devised model with *bona fide* data;

6. By the time all data is collected, the next step of operationalisation is in place for sound and coherent results to be obtained;

7. On that occasion, discussing those results according to a series of parametres is the penultimate phase;

8. In the end, withdrawing applicable conclusions is the ultimate resolution of this research.

1.4 Outline

The present dissertation is composed of six chapters aligned with the objectives and methodology detailed above. In this first chapter, an introduction to the already stated dissertation is presented with a focus on the motivation that led to it, its purpose, and the methodological steps that should be followed to achieve it. The second chapter comprises a robust description of the problem at hand by aiming attention at the global health environment before particularising the Portuguese situation. Then, the foundations are laid for a setting analysis regarding the context and the ambience of the case study at hand. A literature review is the third and rational chapter before developing the cherished model. For the sake of demonstrating the logic behind the reasoning of developing this model, a progressively more in-depth literature review is conducted, beginning with performance evaluation in healthcare and moving towards efficiency and effectiveness models, before ending with the desired performance measurement method - DEA. Now that the fundamentals are settled, a methodology chapter is essential in furtherance of describing the methods in which the soon-to-be-named PRICDEA is based on: the 'two-phase method' built on the additive (ADD) DEA model, the Choquet multiple criteria preference aggregation model, and the linear program for inferring the weights of interacting criteria from a reference set. Additionally, the implementation of this choice and the procedures that lead to its application are also devoted some effort to, namely: the creation of the database, the construction of the utilities' interval scale, and the attribution of utility value to the numerical scale levels using a linear interpolation. Moreover, since the PRICDEA model is a neoteric approach, analogous Möbius coefficients relying on the decision-maker's discernment were also computed for comparison purposes. Next, the fifth chapter enraptures the case study *per se*, where an overview of the situation and an *aperçu* of the stakeholders are engaged in, followed by a description of data, sample, and variables. Results and their discussion are the last sections of this chapter, where the PRICDEA model is tested in three different scenarios discussed with the decision-maker. The achievements and limitations of this dissertation, together with some future prospects, are presented in the last chapter.

Chapter 2

Problem description

The healthcare sector is a delicate one. Everything needs to be in the right place and in the right time so that the maintenance of a flying level of care remains uninterrupted. Consequently, there is a need to reduce costs without compromising quality, which translates into an opportunity to implement operational research-based methods in this field (Florez, Aguirre, Amaya, & Velasco, 2008). The problem described in this chapter is an optimisation headache that represents the reality of the Portuguese SNS secondary healthcare providers.

2.1 Background

New challenges present themselves to the healthcare industry on a daily basis and require a proficient management of the entirety of its dedicated GDP resources. Furthermore, new regulations, new technologies, and new organisations are being continuously created as a result of public policies. Hence, healthcare managers need to overcome these obstacles by mastering performance evaluation and decision-making skills.

In truth, the case of healthcare is not an isolated one - management across all types of industries is moving towards a more objective performance evaluation and decision-making (Ozcan, 2008). However, this particular industry is struggling to keep up the pace with other industries in this respect, although, to be more accurate, healthcare has always historically trailed behind major innovations. Such an example can be found in the enactment of the prospective payment system in the United States of America in 1983, where the healthcare industry strived to meet the needs of their clients due to a decrease in Medicare reimbursements (Guterman & Dobson, 1986). Obviously, the immediate reaction of the administrators was to cut costs and avoid high-risk cases, but it did not take long until they realised that the only way to maintain their institutions' financial viability was to improve their performance. Ergo, benchmarking was (re)born. Woefully, the existing benchmarks were built on ancient analytical schemes based on multiple ratios, which resulted in several quandaries. Thus, creating performance evaluation tools based on optimisation techniques and their normative structure not only creates benchmarks, but also provides information for lagging organisations and illustrates how to improve their performance

(Ozcan, 2008). It is precisely this that the healthcare industry is desperate for nowadays, especially nation-wide institutions as, *e.g.*, Portugal's SNS.

2.2 Setting

An everyday trait of the Portuguese common man/woman is to complain. During one's life, he/she moans about a panoply of subjects, many of which related to health. However, health in Portugal is not as bad as the average citizen bewails. In fact, the Portuguese population enjoys good health and increasing life expectancy, still at lower levels than other western European countries (Barros, Machado, & Simões, 2011).

After the 25th of April Carnation Revolution in 1974, Portugal underwent an authentic turmoil (especially at political, social, and economic levels) that impacted on all strata of the country, both in breadth and in depth, and the healthcare sector was no exception. It was not until 1979 that Portugal had reunited the conditions to create a constitutionalised SNS, rendering the country as one of the first among its peers with such a service (Cortes, 2016).

However, on the verge of its fortieth year of existence, the SNS has withstood several blows due to internal factors (*e.g.*, political shifts) and external factors (*e.g.*, global economic crises). Moreover, the consequences of those aspects on the daily lives of the Portuguese citizens turned Portugal into the country with the highest level of poverty in the European Union in 1994, and, although the scenario has changed in the last twenty years - people in Portugal have never been so healthy -, the access to healthcare and the health status are still important issues on the Portuguese political agenda (Santana, 2002). Nevertheless, according to the World Health Organization, Portugal's SNS is ranked as the 9th best healthcare system in Europe and the 12th best in the world, which proves that not only that policy-makers, despite their political orientations, achieved good results in several health indices, but also that the Portuguese society made a tremendous effort in pursuance of a better life (and, consequently, health status) since 1979 (Sakellarides, Castelo-Branco, Barbosa, & Azevedo, 2014). The preceding world ranking top 10 countries including Portugal with regard to health systems is displayed in Table 2.1.

Table 2.1: World's health systems ranking.

Position	**Country**
1	France
2	Italy
3	San Marino
4	Andorra
5	Malta
6	Singapore
7	Spain
8	Oman
9	Austria
10	Japan
(...)	(...)
12	Portugal

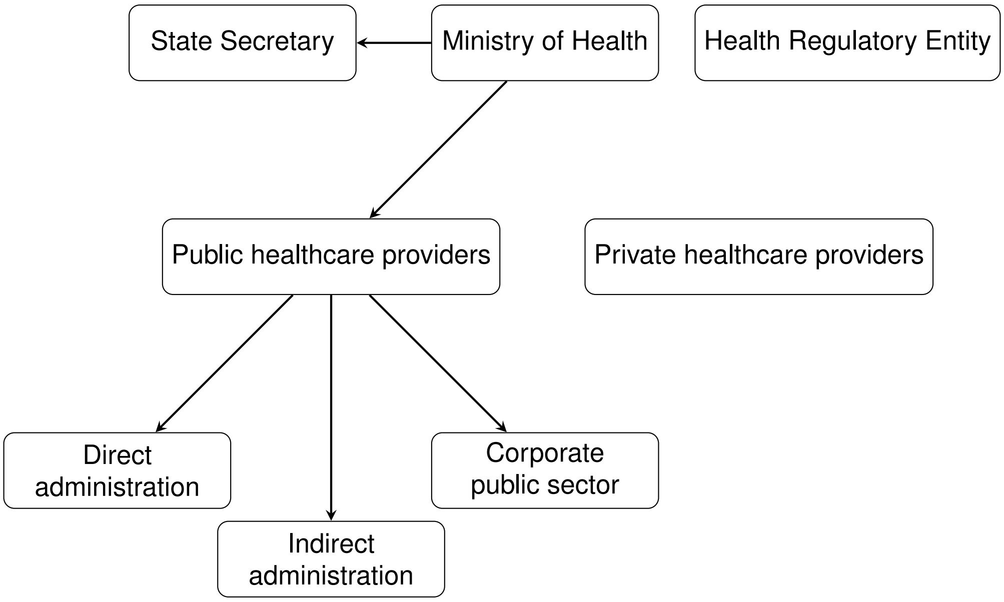

Figure 2.1: Health entities' hierarchy in Portugal.

Interestingly, one would expect that highly developed countries such as the United States of America, the United Kingdom, Germany, or other members of the G20 international forum were ranked in Table 2.1, although, apart from minor exceptions like France and Japan, it is the countries with a relatively minor political and economical impact that make this list. This fact has a multitude of explanations that exceed the scope of this dissertation, but the key evidence to retain is the fact that Portugal is in this top 15 and yet it has numerous performance issues.

Nonetheless, nowadays, health in Portugal is regulated by the Ministry of Health with the assistance of the State Secretary and the Portuguese Health Regulatory Entity. Both public and private healthcare providers are under the sphere of influence of such entities, being the former administrated directly, indirectly, and corporately by assorted institutions. Figure 2.1 depicts this hierarchy. In particular, the SNS comprehends several components of Figure 2.1, both vertical- and horizontally. In line with this reasoning, the schematisation of the institutions comprised by the SNS and their governing entities are clarified in Figure 2.2 (IPO is the Portuguese abbreviation for *Instituto Português de Oncologia.*).

All things considered, evaluating the performance of the corporate public sector institutions of the SNS seems to be an imperative assessment, vital for the sake of the Portuguese health system. Besides, the lack of data or their availability in the other types of SNS institutions' administration entities are hindrances that hamper a more thorough and all-inclusive analysis. Further details on the issue at hand will be disclosed in Chapter 5.

2.3 Summary

In this chapter, the background of the global healthcare sector was characterised and the outline of the case study at hand was focused on. It is clear that the usage of performance measurement techniques is vital for the well-functioning of systems and organisations, which is also unsurprisingly valid for the

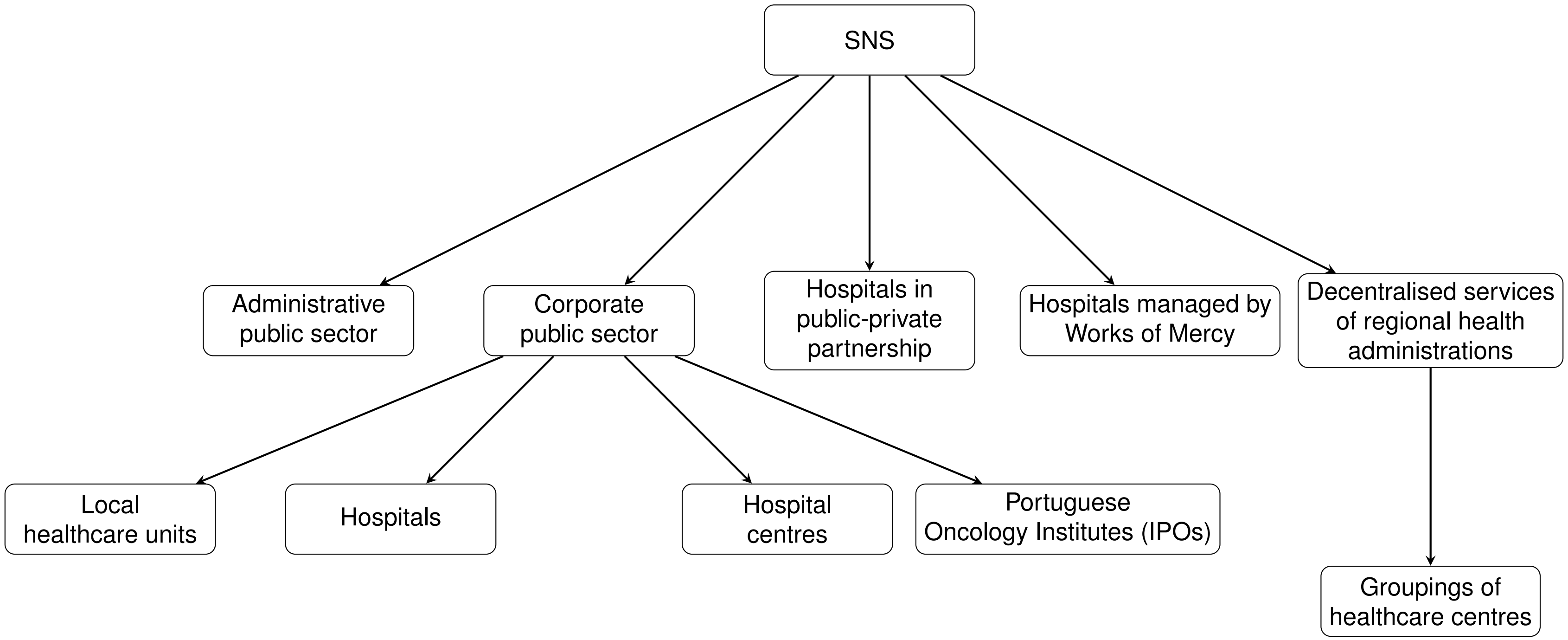

Figure 2.2: SNS institutions grouped by health entity.

case of Portugal's SNS. By delivering a comprehensive historical and political view of it, the practical foundations were laid, with regard to applying in Chapter 5 the ingenious model described in Chapter 4 after an extensive literature review conducted in the following chapter.

Chapter 3

Literature review

This dissertation proposes an innovative performance measurement model applied to a case study in the healthcare sector. Therefore, in Section 3.1, general concepts on performance measurement in healthcare are provided, as well as some typically employed methods for doing so. Section 3.2 provides a guide throughout the pieces that assemble the puzzle of efficiency and effectiveness models. In Section 3.3, performance measurement using DEA is addressed at last, focusing on its components and theories, and finalising with the knowledge gap that justifies the creation of the proposed model.

3.1 Performance evaluation in healthcare

The first section of this literature review is concerned with going through the fundamentals of performance measurement and make advances towards introducing techniques that are popularly applied in the aforestated framework.

3.1.1 Performance measurement

In 1925, Ronald Fisher, one of the most important figures in twentieth-century statistics, created the foundations of modern statistical science with his work on parametric statistics (Fisher, 1925). Parametric statistics assume that data comes from a population that follows a probability distribution based on a set of fixed parametres; in other words, data must meet certain assumptions, otherwise parametric statistics cannot be calculated. On the contrary, non-parametric statistics is not based solely on parametrised families of probability distributions; in fact, the parameter set can increase or decrease in the face of new information, yielding more flexible, but less robust, statistical tests. Both parametric and non-parametric methods have been applied to performance measurement and analysis of healthcare over the past few decades.

Given all the challenges healthcare is confronted with and discussed in Chapter 2, it is imperative healthcare managers readjust their practice to integrate such methods in order to capitalise the resources at their disposal and achieve a higher performance, *i.e.*, a more appropriate combination of efficiency and effectiveness. Nonetheless, these terms are often used with an ill-defined and dubious meaning in this context, therefore it is imperious to distinguish them unequivocally.

Definition 3.1 (Efficiency, Ozcan, 2008)**.** Efficiency refers to using the minimum number of inputs for a given number of outputs.

This means that an efficient healthcare facility produces a certain level of care that meets a stipulated quality standard, using the minimum combination of resources. However, in performance literature, the concepts of efficiency and productivity are conversely used in a recurrent way. In spite of productivity typically connoting a broader meaning than efficiency, both concepts are constituents of a more comprehensive notion - performance - as conceptualised in Figure 3.1.

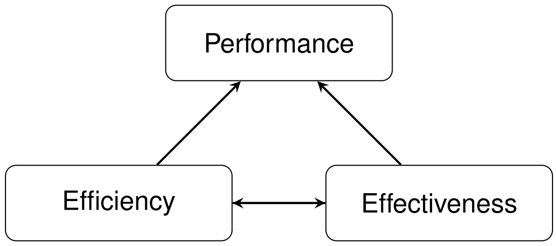

Figure 3.1: Performance components (adapted from Ozcan, 2008).

Definition 3.2 (Effectiveness, Ozcan, 2008)**.** Effectiveness regards the extent of medical care outcomes and may influence or be influenced by efficiency, on top of having a repercussion in healthcare performance.

As a matter of fact, effectiveness concerns the utilisation of the necessary amounts of inputs for the sake of producing the most superlative outcomes. Indeed, a deeper distinction could be introduced between effectiveness and efficacy in what concerns the pragmatism and idealism of each concept, respectively (Kim, 2013), but this matter will be slightly simplified by referring to effectiveness whenever necessary.

Nevertheless, one detail is clear: an healthcare facility can be simultaneously efficient and ineffective, or inefficient and effective. The goal is to be both efficient and effective, evidently; this way, performance will certainly be enhanced as suggested in Figure 3.1. Thorny obstacles and fierce competition will keep on haunting the dreams of healthcare managers, but improving performance within their institutions is the key to their survival. There might not be a magical formula for doing so - each healthcare organisation must be individually examined -, but it is up to healthcare managers to determine the appropriate input and output mix (Ozcan, 2008).

Literature shows that the relationship between efficiency and quality of care is heterogeneous, since some authors found that increasing efficiency results in a quality care (Helling, Nelson, Ramirez, & Humphries, 2006) and other researchers established that the contrary is not always achieved (Singaroyan, Seed, & Egdell, 2006). Yet, Mobley and Magnussen (2002) concluded that poor quality outcomes are associated with an inferior efficiency and Ferrando *et al.* (2005) indicated that healthcare organisations can increase efficiency without impacting the quality of care. For this reason, performance needs to be evaluated and compared across healthcare organisations for reasons that comprehend, according to Ozcan (2008):

- Detecting changes from one period to another;
- Determining how organisations are functioning in relation to others in a given competitive market;
- Investigating deviations from strategic planning scenarios.

This means that performance in the healthcare context is a rather relative phenomenon across healthcare organisations, *i.e.*, a *rara avis* that can be compared at one or multiple points in time within the same organisation or between organisations. The interrogation remains on determining the values that yield efficiency and effectiveness scores, which will be addressed in Subsection 3.1.2.

3.1.2 Performance evaluation methods

In agreement with Ozcan (2008), a comparative performance analysis can be undertaken among various parametric and non-parametric methods, *viz.*:

- Ratio analysis;
- LSR;
- TFP;
- SFA;
- DEA.

3.1.2.1 Ratio analysis

One of the simplest and oldest performance calculation methods found in literature is the ratio analysis. Mainly used for efficiency (productivity), it associates one input and one output, and produces information defined as the number of output units per unit of input:

$$\text{Efficiency (Productivity)} = \frac{\text{Output}}{\text{Input}} \tag{3.1}$$

Given the number of performance dimensions that have to be round up among compatible units or within one unit over different time periods, several ratios have to be computed, which is particularly valid in the context of healthcare. To that end, comparative benchmark and performance statistics are provided by organisations such as Medicare Hospital Compare and Healthcare Effectiveness Data and Information Set (Paddock, 2014).

Anyhow, the gargantuan amount of ratios frequently produces mixed results that disorient more than they aid healthcare managers in comparative performance analysis. This epitomises the weakness of using a ratio analysis in a framework as complex as healthcare, in light of the difficulties healthcare managers have in identifying a consistent benchmark that incorporates all inputs and outputs of an healthcare organisation (Ozcan, 2008).

3.1.2.2 Least-squares regression (LSR)

By harmonising multiple inputs and outputs, as well as taking noise into consideration, the LSR is one of the most popular parametric methods in performance evaluation. Customarily, the LSR is formulated as

$$y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + ... + \beta_n x_n + e, \tag{3.2}$$

where y has a normal distribution for any fixed value of x (being the y values independent of one another), the mean value of y is a straight-line function of x that includes the error term e, the variance of y is the same for any x, and, for any fixed value of x, y is a random variable $(y|x) = \beta_0 + \beta_1 x$.

This method has benefits when a technical change of a time-series data is required and when one is investigating scale economies. However, not everything is rosy, on the grounds that the employed central tendency measures are not necessarily efficient relationships, the LSR does not identify individual inefficient units nor the best performances, and it demands a pre-specified production function due to its parametric formulation (Ozcan, 2008). Ergo, more suitable methodologies that are capable of conveying a more robust performance evaluation in the healthcare context will be addressed in the following paragraphs.

3.1.2.3 Total factor productivity (TFP)

In Subsection 3.1.2.1, the limitations of ratio analysis in incorporating multiple inputs and/or outputs into a single performance ratio were exposed. Thus, TFP overcomes this weakness due to its index numbering measurement usage. As a matter of fact, index numbers may be used to measure price and quantity changes over time, which makes them able to measure differences across healthcare organisations. A general formulation of TFP is usually defined as

$$\text{TFP}_{ab} = \frac{\sum_{i=1}^{N} p_{ib} q_{ib}}{\sum_{i=1}^{N} p_{ia} q_{ia}}, \tag{3.3}$$

where the TFP_{ab} index measures the change in the value of selected quantities q of N outputs from a period a to b, considering the prices p of those outputs.

The most bourgeois indices found in literature include the Laspeyres index, the Pasche index, the Fisher index, the Törnqvist index, and the Malmquist index (Ozcan, 2008). The first four are non-parametric methods that can be used with cross-sectional data to measure the performance of two healthcare organisations in one time period or the performance of one healthcare organisation in two time periods. Be that as it may, the comparison of more than two healthcare organisations at the same time or over time is impractical. Hence, the Malmquist index rises among the TFP measures as the most frequently used method in healthcare, since it gets around some of the deficiencies of the other indices. This index can be computed by way of applying frontier techniques (*e.g.*, SFA or DEA) using either an input-oriented approach or an output-oriented approach. An interesting feature of the Malmquist index is that it decomposes overall efficiency into two mutually exclusive components: one that measures change in technical efficiency and the other that measures change in technology (Tone, 2004). Nevertheless, the Malmquist index will not be focused on in this dissertation, but it would certainly be a thrilling prospect.

3.1.2.4 Stochastic frontier analysis (SFA)

Another parametric technique is SFA. When compared with previously described performance evaluation methods, SFA assumes that deviations from the efficient frontier are due to a noise factor. A generic SFA model is ordinarily formulated as

$$\mathrm{TC} = \mathrm{TC}(Y, W) + V + U, \tag{3.4}$$

where TC stands for total cost, Y represents outputs, W connotes input prices, V depicts the random error (assumed to be normally distributed with zero mean and zero variance), and U acts as the inefficiency residual.

All things considered, SFA is ideal for hypotheses' testing and to measure not only technical and allocative efficiencies, but also scale economies, for instance (Ozcan, 2008). Nonetheless, SFA's drawbacks are centred on specifying functional form and distributional form for U. Additionally, the use of both price information and quantity information may induce measurement errors in the results, culminating in inefficiency of technical and/or allocative origin (Kooreman, 1994).

3.1.2.5 Data envelopment analysis (DEA)

The only non-parametric technique in Subsection 3.1.2 is DEA. In opposition to SFA, DEA assumes that not all units are efficient and allows the usage of multiple inputs and outputs in a linear programming model, which will then compute a single efficiency score per observation. Such efficiency measures may comprise various types, which will be described in Subsection 3.3.1.1. The computations behind such efficiency types will not be addressed thoroughly in this dissertation.

Still, DEA does not incorporate noise as SFA due to its deterministic nature, otherwise units would become inefficient due to deviations from the efficient frontier, in spite of researchers being currently developing stochastic variants of DEA models that account for a random error component, such as Tsionas (2003). Forasmuch as DEA is thought-out to be the main performance evaluation methodology, its application in the healthcare context embellishes the work developed throughout this dissertation.

3.2 Efficiency and effectiveness models

Section 3.2 intends to particularise performance measurement models into efficiency and effectiveness ones, arranging information by measure and respective essential components.

3.2.1 Efficiency measures

In Subsection 3.1.2.1 elementary efficiency was presented as being measured as an output/input ratio. Thus, to improve efficiency, one has to either: increase the outputs, decrease the inputs, assess if the rate of increase for outputs is greater than the rate of increase for inputs whenever both outputs and inputs increase, or assess if the rate of decrease for outputs is lower than the rate of decrease for inputs whenever both outputs and inputs decrease. Higher efficiency may also be attained by introducing technological changes or re-engineering service processes (Ozcan, 2005, 2008), but these transformations exceed the scope of this dissertation.

Constructed on linear programming foundations, DEA's optimisation methodology stands up to the aforementioned methods since it pinpoints the optimal ways of performance rather than mere averages, which is particularly compelling in a world where no one wants to be average. Therefore, the identification of superlative performances normatively points the way towards benchmarking to the delight of healthcare managers, forasmuch as they become able not only to identify top performers, but also to conceive alternative ways to navigate their organisations through the stormy healthcare seas.

A myriad of DEA research initiatives has been played around with since the original work of Charnes, Cooper, & Rhodes (1978). Be that as it may, DEA has not yet been adopted as a canonical tool for benchmarking and decision-making in healthcare and its use has been limited to academia, peculiarly because experts fall short of bridging the theory-practice gap, although DEA's formulation is quite intricate.

3.2.2 Efficiency model components

When engaging in an empirical efficiency analysis in healthcare, one will certainly encounter copious theoretical and practical issues that call for illumination.

It has been established throughout this dissertation that the efficiency of an healthcare organisation corresponds to the ratio of outputs it generates to the value of inputs it depletes and it is precisely this outlook that is encapsulated in Figure 3.2.
The schematised organisation consumes M physical resources valued as X by society and produces S outputs aggregated as Y by society. The transformation process is not relevant at this stage. The ratio Y/X is the essential efficiency measure that could be referred to as cost-effectiveness.

Healthcare efficiency models typically contemplate multiple inputs and/or outputs. Consequently, the relative weights V_m and U_s attached to each input m and output s, respectively, are pivotal to the calculation of X and Y (weighting vectors $\mathbf{V}$ and $\mathbf{U}$ indicate the relative importance of an additional unit of input or output). These weights allows one to compute the valuation of inputs $X = \sum_{m=1}^{M} V_m X_{m0}$ and the valuation of outputs $Y = \sum_{s=1}^{S} U_s Y_{s0}$ for an hypothetical healthcare organisation named 0. In competitive markets, both $\mathbf{V}$ and $\mathbf{U}$ might be already available, since they are able of being considered as prices, which propitiates the immediate calculation of the efficiency measure. Otherwise, in complex contexts like healthcare, this is rarely the case, particularly on the output side, and that is where analytic techniques such as DEA can be deployed to provide akin information.

Despite the fact that this framework is misleadingly simple, several complex problems quickly arise

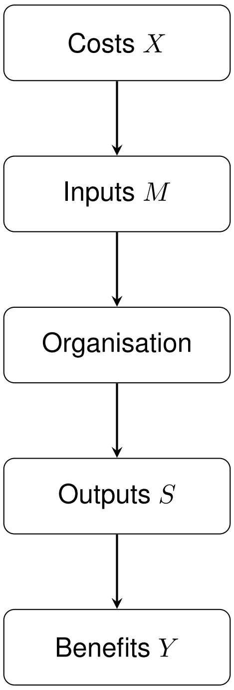

Figure 3.2: Naïve model of organisational performance (adapted from Jacobs, Smith, & Street, 2006).

when one intends to use it in the development of organisational efficiency operational models in healthcare. Indeed, the complexity involved in such an endeavour merely reflects the complexity of the production process. Rather than a production line, most healthcare is perfectly fitted to the needs of the individual recipient (Harris, 1977). This means that heterogeneity and lack of clarity regarding the production process are abundant. Additionally, multiple stakeholders are involved over a legion of time periods and in different settings (Jacobs, Smith, & Street, 2006). Therefore, prior to addressing DEA itself, a few model-building principles must first have a go at.

3.2.2.1 Unit under investigation

A legitimate efficiency analysis should have a clear idea of the entity it is examining, but should also acknowledge that its performance might be influenced by other entities or by factors beyond its control. On that account, three criteria should guide the choice of units (Jacobs et al., 2006):

1. They should capture the entire production process of interest;
2. They should be DMUs, *i.e.*, their function should be to convert inputs to outputs and to be discrete about the technological conversion process that takes place;
3. They should be comparable, especially in the sense that they are seeking to produce the same set of outputs.

This brings about the generic definition of DMU.

Definition 3.3 (Decision-making unit)**.** A DMU is an organisation subject to evaluation in the light of DEA literature. Hospitals, clinics, primary care centres, and other entities are examples of facilities whose performance is adequate for DEA.

Moreover, the interoperationality between all of these criteria is not necessarily elementary. Even by themselves, one criterion is a handful (*e.g.*, almost every organisation claims to be "unique", so comparing the same set of healthcare outputs is difficult to achieve). An example of an interoperability friction is the conflict between the first and the second criteria, which is most seemingly to happen where the production process is characterised by varying degrees of vertical integration. Nevertheless, under some circumstances, organisations may vertically integrate and assume control of the entire production process; under others, they may prefer to buy in inputs from organisations further down the process or sell on to those further up the chain.

Bottom line, minimising transaction costs as a factor in determining what range of production process might be under control of a single organisation was found to be desirable (Coase, 1937; Williamson, 1973).

However, guaranteeing that the analytical DMU fully circumscribes jointness in production is particularly important in contexts where there is variation in how relative contributions to joint production are defined, which raises the question of where the boundaries of the production process should be drawn. On one hand, the DMU could be thought of as the entire health system. As loose as this definition might be, it is also infeasible and often unhelpful - defining the system and identifying its fundamental decision-makers and inputs is not entirely clear. On the other hand, taking interest in single actions or individuals is much more preferable in comparison to a whole-system approach, because their activities are more limited, the inputs can be more easily identified, and assigning personal responsibility for performance is more likely to be higher. For all that, there are handicaps to this approach, since outputs are generally produced by teams and entities might be inherently linked with each other.

In essence, a larger collection of individuals is more appropriate for analytical purposes, when outputs result from joint production decisions (Jacobs et al., 2006).

3.2.2.2 Healthcare inputs

The less puzzling side among what goes in and out of an healthcare organisation, as presented on Figure 3.2, is, effectively, what enters that system, *i.e.*, inputs. Inputs' physical nature allows a more rigorous measurement than outputs, as will be seen in Subsection 3.2.2.3, especially since they can be summarised as costs, notwithstanding genuine conceptual and practical dilemmas that input specification institutes on efficiency analysis.

One indispensable resolution that must be regarded is the level of disaggregation of inputs to be specified (viewed as a long-term perspective). The use of a single measure of costs essentially transforms the efficiency model into a cost function and it will be directly reflected in price efficiency. Another crucial decision is to acknowledge that a specific number of aspects of the input mix is outside the sphere of influence of the organisation (viewed as a short-term perspective). Consequently, disaggregating the inputs to some extent is likewise required, for the sake of capturing the contrasting inherited input mixes by healthcare organisations. For these reasons, two input categories come into sight, in consonance with Jacobs *et al.* (2006):

- **Labour inputs**

  Labour inputs can be measured by skill level disaggregation. Aggregating labour types into a single measure of labour input and weighing the various labour inputs by relative wages is an appropriate approach, except for when there is a distinct interest in the deployment of different types. Nevertheless, if such evidence does not exist, disaggregating labour inputs is not praiseworthy. There are models that might be willing to consider skill mix disaggregation, producing a valuable recommendation for policy-making.

  Additionally, labour inputs are commonly measured in either physical units or costs of labour, and their independent use builds upon the context, resting assured that each has its advantages and disadvantages. At an organisational level, labour inputs can be promptly measured, but headaches occur when the efficiency analysis is focused on sub-units of healthcare organisations, since attributing labour inputs becomes problematic when the unit of observation becomes smaller.

  Another reflection depends on the extent to which healthcare organisations buy in certain services, instead of directly hiring new workers. This complicates the comparison between organisations, forcing analysts to make use of a single measure of inputs - total costs.

- **Capital inputs**

  Integrating capital measures as inputs of an efficiency analysis is more demanding, because of the arduousness of measuring capital stock and the issues in attributing its usage to a particular time period, both to a certain degree. Generally, capital measures are quite vulgar and ambiguous (*e.g.*, the number of hospital beds as a proxy for physical capital). Besides, healthcare organisations usually invest in non-physical capital inputs, like health promotion.

  All things considered, an efficiency model ordinarily uses the capital consumed in the current time period as an input to the production process. However, capital is used across time by nature. The diagram laid out in Figure 3.3 encapsulates the accounting struggles capital inputs give rise to.

  In each time period t, inputs are consumed and outputs are produced by and from the healthcare organisation, respectively. If an organisation has been blessed with the benefits of past investments in period t, then the output/input ratio for that time period will not be a *bona fide* efficiency measurement.

  In spite of all, when treating capital inputs, one must take into consideration if the desired efficiency measurement is probing a short-term or a long-term, in favour of avoiding imprecise evaluations.

3.2.2.3 Healthcare outputs

In competitive industries, the physical output of an organisation is customarily a traded product. Naturally, despite how homogeneous a market is, such products variability is tremendous on various dimensions of quality (*e.g.*, reliability, conformance, and serviceability). The quality of a product is intrinsic to its social value, but that value can be read between lines by scrutinising people's purchasing behaviour (Jacobs et al., 2006).

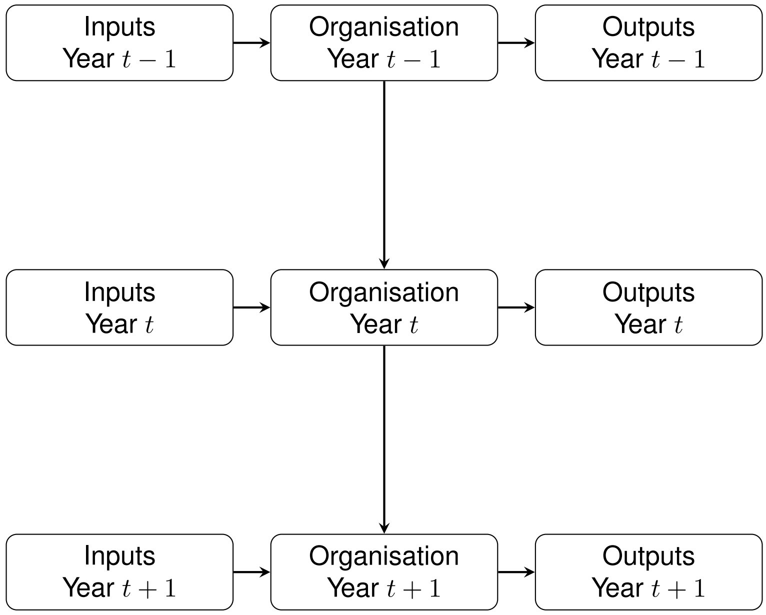

Figure 3.3: Dynamic nature of organisational performance (adapted from Jacobs *et al.*, 2006).

Anyhow, in profuse sectors of the economy, not only do prices not exist, but also defining outputs is a challenging task, particularly in the case of governmental spending on goods and services (Atkinson, 2005). The context of healthcare is no different, since valuating the concept *per se* and not existing a conventional sense of market make the undertaking slightly more tortuous. Seeing that healthcare makes a positive contribution to the health status, rather than being demanded for its own sake, healthcare outputs should be defined in terms of produced healthcare outcomes, even though organisations do not commonly collect periodic information about their own outcomes.

Two comprehensive classes can be considered when talking about healthcare outcomes (Jacobs et al., 2006), transforming them into quality-adjusted physical outputs:

1. Additional health conferred on the patient;

2. Broader patient satisfaction over and above that related to the health effect.

On one hand, for most patients and caregivers, health gain is the key intervention success indicator, in that it focuses attention on the patient on behalf of the services provided by the organisation. Fundamentally, measuring an health outcome should point out the added value to health, due to contact with the health system. Goldstein & Spiegelhalter (1996) successfully applied this 'value-added' measurement in the school education sector to assess pupils' ability at entry to each school and then comparing their exam grades in relation to that baseline, yielding the preceding measure. However, when attempts were made towards applying this concept in healthcare, obstacles have rapidly risen, by cause of a much greater service users' heterogeneity and intrinsic measurement difficulties. In essence, finding a baseline is extraordinarily unattainable. On that account, involving comparisons of health status before and after an intervention is acceptable for these purposes, although the aforementioned measures tend to introduce bias into the comparative analysis (Jacobs et al., 2006). Changes in health status with and without an intervention could also be considered, but, without treatment, health profiles are infrequently observable. Collecting before/after measures of treatment effects have been developed and instruments like EQ5D and SF36 (EuroQol Group, 1990; Ware & Sherbourne, 1992) are deep-rooted nowadays, whose limitations will not be addressed in this dissertation.

On the other hand, and in all respects aside from health gains, the concerns with “patient experience” (covering issues such as autonomy, privacy, and response times) are becoming progressively more vocal. However, efficiency studies do not usually integrate such information (Pedraja-Chaparro, Salinas-Jiménez, & Smith, 1999), with the exception of the World Health Organisation’s development of the concept of responsiveness on the examination of national health systems’ efficiency (2000). Nevertheless, in spite of how useful this United Nation’s specialised agency’s report was, it was crippled by the usage of weak measurement methods. In recent past, Üstün, Chatterji, Mechbal, Murray, & World Health Surveys Collaborating Groups (2003) addressed this issue in a more adequate way. In fact, several tools have been applied to a regular extent, in spite of still being difficult to integrate an operation efficiency analysis, and, with proper data condensation, a small number of responsiveness measures can be obtained (Coulter & Magee, 2003).

Without regard to everything that was mentioned above, health outcomes are not the only type of measure analysts deal with. Healthcare activities are another kind of measure, despite how crude they are in capturing effectiveness variations of healthcare delivery. Indeed, some impracticalities concerning healthcare outcomes lead to the emergence of healthcare activities as proxies (Jacobs et al., 2006). However, measuring activities has its drawbacks: identifying the degree of variation in outcomes is straightforwardly traceable to healthcare organisations’ actions. In truth, relying on them, in light of sound research evidence, gives rise to health improvement, and measuring such activities provides a powerful indicator of expected health outcomes, even if one inherently assumes that all organisations have the same effectiveness when implementing a procedure. At last, the employment of activity measures is recurrently the sole option at hand, but one must bear in mind of its limitations. Particularly, a pair of misjudgements rely on, *ceteris paribus*, healthcare organisations that engage in more activities being rated as efficient and the effectiveness of healthcare delivery not being captured by a handful of activities (Jacobs et al., 2006).

In the end, one question emerges: how does one assess the value of different kinds of healthcare outputs? Given the intricacy of healthcare organisations and the wide judgmental variation among stakeholders, the answer lies in attributing the decision-maker role to someone who has to decide what is valued on behalf of society. More than healthcare managers, it is up to politicians to provide valid political statements on the goals of health systems or their associated organisations (Jacobs et al., 2006).

3.2.2.4 Environmental determinants of performance

Organisational capacity is influenced by a plethora of factors, but, in plentiful contexts, environmental constraints rise as an isolated class of exogenous forces that affect an organisation’s performance beyond its control. These effects reflect the external environment within which it must operate, being often greatly dependent on population characteristics. For instance, population mortality rates, surgical outcomes, and emergency ambulatory services performance heavily build upon demography, community, or geography (Jacobs et al., 2006).

There is a good deal of debates concerning the controllability of such determinants. In fact, even a short-run and tactical analysis or a longer-run and strategic analysis require different data and are constrained by distinctive factors. Additionally, the performance of several healthcare organisations relies upon inputs from external agencies, and this dependency should be recognised in efficiency modelling. Jacobs *et al.* (2006) propose three wide-ranging manners of taking environmental constraints into account in efficiency analysis:

- Restrict the comparison to organisations with a similarly constrained environment;
- Model the constraints as being analogous to production process' factors;
- Undertake risk adjustment.

The first resolution clusters organisations into comparable classes, using cluster analysis (Everitt, Landau, Leese, & Stahl, 2011). However, the issue about what criteria should be used to create these categories is imperative. On one hand, it is impossible to conclude if performance is correlated with exogenous influences or efficiency variations. On the other hand, clustering reduces sample size, as an extrapolation of performance is ruled out.

The second approach attempts to incorporate environmental performance determinants alike labour or capital inputs. This concept solves the extrapolation problem of the previous point, by generalising the clustering procedure.

At last, the final method adjusts organisational outputs for differences in circumstances before they are set up in an efficiency model. This astute approach allows the analyst to fine-tune each output for only those factors that apply specifically for that output, in lieu of the contrary, and have been cultivated to a high level of precision (Iezzoni, 2003).

Wistfully, environmental constraints were not studied in this dissertation, but would have certainly been an even more formidable addition.

3.2.3 Effectiveness measures

In healthcare, effectiveness is measured by outcomes or quality and is of vital importance to numerous stakeholders. Notwithstanding, measuring effectiveness is more puzzling than measuring efficiency, on the grounds that a multiplex of angles on outcomes and quality introduce further measurement adversities, despite the acquaintance of processes' inputs and outputs. Moreover, the majority of outcomes measures and quality measures were not systematically reported, which entangled the task of healthcare managers. Nevertheless, measuring effectiveness is not the purpose of this dissertation, the focus remains on more compelling subjects.

3.3 Performance measurement using data envelopment analysis (DEA)

The last section of Chapter 3 devotes its efforts to describing DEA methodology, its usage in the healthcare context, and the specific case of preference information incorporation.

3.3.1 DEA methodology

In opposition to the parametric approach governed by economic theory, DEA is a data-driven technique. It is data that determines the location and the shape of the efficiency frontier, constructed by observing the highest output/input ratios and joining those observations up in the input-output space. Inefficient units are enveloped by this efficient frontier and such inefficiency is calculated relatively to this surface (Grosskopf & Valdmanis, 1987; Charnes, Cooper, Lewin, & Seiford, 1994; Cooper, Seiford, & Tone, 2000).

As it was seen in Definition 3.3, DEA literature conventionally uses the DMU terminology to invoke each of the units of analysis under scrutiny (Charnes et al., 1978), which can range from entire healthcare systems (Puig-Junoy, 1998a) to health regions (Ozcan & Cotter, 1994; Gerdtham, Rehnberg, & Tambour, 1999), from whole hospitals (Grosskopf & Valdmanis, 1987) to specific hospital services (Puig-Junoy, 1998b; Hollingsworth & Parkin, 2001), or even reach individual physicians (Chilingerian, 1994) or other healthcare professionals.

In this subsection, the various notions of efficiency will be reviewed within the context of DEA, as well as all the paraphernalia involving this technique's formulation and considerations.

3.3.1.1 Efficiency evaluations using DEA

Efficiency is one the major components of performance, as depicted in Figure 3.1, and is drawn up as a ratio of outputs over inputs. In DEA, it is computed relatively to healthcare organisations in a particular assessment, hence the score for benchmark healthcare organisations uniquely represents the set of units considered in the analysis. This naturally influences the results of future evaluations, not only because each organisation may perform differently from one time period to another, but also due to the arrival or departure of DMUs (Ozcan, 2008).

Innumerable factors may also act upon overall efficiency in such manner that it is reasonable to define four components, as shown in Figure 3.4, upon which efficiency classification relies on that are *ad hoc* of healthcare evaluation (Othman, Mohd-Zamil, Rasid, Vakilbashi, & Mokhber, 2016).

Definition 3.4 (Technical efficiency)**.** Also known as global efficiency, technical efficiency refers to the physical relation between resources and health outcomes. A technically efficient position is achieved when the maximum possible improvement in outputs is obtained with fewer inputs (Palmer & Torgerson, 1999).

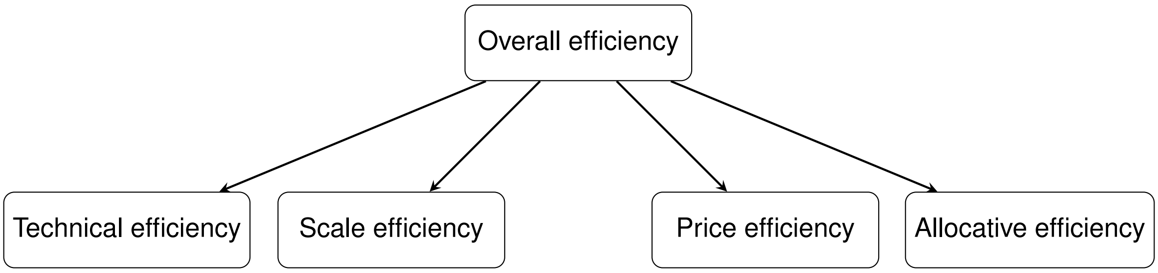

Figure 3.4: Efficiency components (adapted from Sherman & Zhu, 2006).

Definition 3.5 (Scale efficiency)**.** Scale efficiency has been developed in different ways in the DEA framework (Daraio & Simar, 2007). Howbeit, a general notion concerns the performance of a unit operating in optimum scale: a unit is scale efficient when it works at the most productive scale size, *i.e.*, when its input and output sets are optimal so that any modifications will render them less efficient (Eken & Kale, 2014; Shanmugam, 2014).

Definition 3.6 (Price efficiency)**.** Reducing input cost while maintaining quality is a fitting and concise explanation for price efficiency (Othman et al., 2016).

Definition 3.7 (Allocative efficiency)**.** A breviloquent interpretation of allocative efficiency appertains to the measure of the optimal input mix in order to produce outputs (Sherman & Zhu, 2006; Othman et al., 2016).

Some other efficiency components can be classified (*e.g.*, X-efficiency and profit efficiency), but these four fundamental constituents comprise a natural and common categorisation of overall productivity.

3.3.1.2 DEA formulation

Generally, the efficiency concepts discussed in the previous paragraph assume one knows the production function of the organisation. In the real world, this is not always true. In line with Jacobs *et al.* (2006), DEA assesses efficiency in two stages:

1. Identify a frontier based on either those organisations using the lowest input mix to produce their outputs or those organisations achieving the highest output mix given their inputs, *i.e.*, the input orientation or the output orientation;

2. Form a piecewise linear envelope of surfaces in multidimensional space by assigning an efficiency score to each organisation and comparing their respective output/input ratio to that of efficient organisations.

Using the same nomenclature as in Figure 3.2, if there are M inputs and S outputs, then the production frontier becomes a surface in $(M + S)$ dimensional space. Consequently, the efficiency of a DMU is the distance it lies from this surface, which is defined as the ratio of a weighted sum of outputs of a DMU divided by a weighted sum of its inputs. Hence, technical efficiency is computed by solving, for each

DMU, the mathematical program

$$\begin{aligned}
\max \quad & \frac{\sum_{s=1}^{S} u_s y_{s0}}{\sum_{m=1}^{M} v_m x_{m0}} \\
\text{s.t.} \quad & \frac{\sum_{s=1}^{S} u_s y_{si}}{\sum_{m=1}^{M} v_m x_{mi}} \leqslant 1, \\
& u_s, v_s \geqslant 0 \\
& i = 1, ..., I
\end{aligned} \tag{3.5}$$

where y_{s0} is the quantity of output s for DMU$_0$, u_s is the weight attached to output s ($u_s > 0, \quad s = 1, ..., S$), x_{s0} is the quantity of input m for DMU$_0$, and v_m is the weight attached to output m ($v_m > 0, \quad m = 1, ..., M$).

This mathematical program tracks down DMU$_0$, the set of output weights u_s, and the set of input weights v_m that maximise the efficiency of DMU$_0$, subject to the constraint that no DMU can have efficiency greater than 1. As for the weights, they can accept any non-negative value and, ordinarily, each DMU has its own distinctive set of weights, fashioning u_s and v_m as a key DEA characteristic - they are chosen to single out the higher level of efficiency for a certain DMU. Equation (3.5) can be rewritten as

$$\begin{aligned}
\max \quad & \frac{u' y_0}{v' x_0} \\
\text{s.t.} \quad & \frac{u' y_i}{v' x_i} \leqslant 1 \\
& i = 1, ..., I \\
& u', v' \geqslant 0
\end{aligned} \tag{3.6}$$

where u' and v' are output and input weight vectors, respectively. In the interest of obtaining the optimal weights, Equation (3.6) is converted into a system of linear equations, erected in a way that a linear objective function can be maximised while subjected to a set of linear constraints. Nevertheless, this ratio formulation has an infinite number of solutions (Coelli, Rao, & Battese, 1998), so an additional constraint stating that either the numerator or the denominator of the efficiency ratio has to be equal to 1 had to be enforced. Thus, the problem becomes one of two: maximise the weighted output subject to weighted input being equal to 1 or minimise weighted input subject to weighted output being equal to 1 (Parkin & Hollingsworth, 1997). To echo this metamorphosis, Equation (3.6) can be reformulated in the multiplier form as

$$\begin{aligned} \max \quad & \mu' y_0 \\ \text{s.t.} \quad & \nu' x_0 = 1 \\ & \mu' y_i - \nu' x_i \leqslant 0 \\ & i = 1, ..., I \\ & \mu, \nu \geqslant 0 \end{aligned} \tag{3.7}$$

which could be also expressed as the correspondent minimisation problem (Coelli et al., 1998)

$$\begin{aligned} \min \quad & \theta_0 \\ \text{s.t.} \quad & -y_0 + Y\lambda \geqslant 0 \\ & \theta x_0 - X\lambda \geqslant 0 \\ & \lambda \geqslant 0 \end{aligned} \tag{3.8}$$

that has the advantage of involving fewer constraints, where x_i and y_i are input and output column vectors for each of the I-DMUs, X and Y are the input and output matrices representing the data for all the I-DMUs, θ is a scalar, and λ is a $I \times 1$ vector of variables. The value of θ corresponds to the efficiency score for DMU$_0$ and satisfies $\theta \leqslant 1$.

This linear program must be solved for each individual DMU in the data set in pursuance of the value of θ for each one (Coelli et al., 1998). Ergo, this DEA formulation yields weights that are specific to each DMU and its objective is to scout for the minimum θ that reduces the input vector x_i to θx_i while assuring the output level y_i. However, λ_i reflects the weights attached to DMU$_i$ in forming the efficient benchmark for DMU$_0$ - indeed, DMU$_0$ is compared to the point on the frontier formed by creating a composite peer DMU encompassing a linear combination of all other DMUs (Jacobs et al., 2006). Naturally, non-zero weights will be designated only for efficient DMUs in the peer group, producing the so-called efficient peers or DMU$_0$ comparators. Further considerations on model orientation and returns-to-scale will be addressed shortly in Subsection 3.3.1.3 and Subsection 3.3.1.4.

3.3.1.3 Returns-to-scale

Until now, all of the *vide supra* formulations were envisaged under a constants returns-to-scale (CRS), as postulated in the seminal DEA paper (Charnes et al., 1978). Curiously, the CRS model was known as the CCR model, because of its developers surname's initials, as would the subsequent BCC model. Indeed, a few years later, Banker, Charnes, & Cooper (1984) broadened this concept in the interest of integrating a more malleable scale - variable returns-to-scale (VRS) - model, which could be applied when not all DMUs were considered to be operating at an optimal scale. In healthcare, operating at an ineficient scale occurs on a daily basis due to countless factors, *e.g.*, imperfect competition, financial constraints, and regulatory constraints (Jacobs et al., 2006). Therefore, choosing between CRS and

VRS is an important decision.

In line with Equation (3.8), the addition of a single convexity constraint turns this CRS linear programming problem into a VRS one:

$$\sum_{i=1}^{I} \lambda_i = 1 \tag{3.9}$$

As a deduction, the calculation of scale inefficiency could be computed by running both CRS and VRS DEA models on the same data and attributing any change in measured efficiency to it (Coelli, 1996; Parkin & Hollingsworth, 1997). With a view to collect evidence of whether a DMU is operating in the area of increasing or decreasing returns-to-scale, a non-increasing returns-to-scale (NIRS) constant can be added to Equation (3.8), modifying Equation (3.9):

$$\sum_{i=1}^{I} \lambda_i \leqslant 1 \tag{3.10}$$

Scale inefficiencies can then be calculated by comparing the DMU's technical efficiency score over the NIRS constraint to their technical efficiency score under the NIRS constraint. As reported by Coelli, Rao, & Battese (1998), if they are not equal, then increasing returns-to-scale exist; if they are equal, then decreasing returns-to-scale apply.

At last, the choice between CRS or VRS depends on the context and purpose of the analysis or whether short-run or long-run efficiency is under scrutiny (Jacobs et al., 2006). For instance, a productivity-based perspective may be more appropriately built using a CRS model and a managerial-based perspective may prefer a VRS model.

3.3.1.4 Model orientation

It was mentioned in Subsection 3.1.2.1 that efficiency in ratio analysis is calculated using an output over input ratio, where input reduction is usually stressed in order to improve efficiency. In the DEA context, this approach is named input orientation and can be properly ascertained in Definition 3.8.

Definition 3.8 (Input-oriented DEA, Ozcan, 2008)**.** DEA's input orientation assumes healthcare managers hold the reins more over the inputs rather than outputs.

Howbeit, the contrary might also be verified, since outputs also play their part in this analysis, as stated in Definition 3.9.

Definition 3.9 (Output-oriented DEA, Ozcan, 2008)**.** In case healthcare managers run the show with regards to outputs, *i.e.*, when they can attract patients to their facilities by marketing or other means, and increase their organisation's efficiency given their input set, DEA has an output orientation.

To bridge the gap between this paragraph and the last one, this subject can be elaborated to some extent. Under CRS, input-oriented or output-oriented DEA results are the same. However, the same cannot be said when one talks about VRS. It is incorrect to think that the choice of orientation affects which observations are identified as fully efficient (Coelli, 1996), but the distance of an ineffective DMU from the frontier depends on the horizontal distance in the input orientation and on the vertical distance on the output orientation, thus varying the technical efficiency score for that DMU (Jacobs et al., 2006).

Additionally, a miscellaneous of DEA models has been developed over the years assuming either input orientation or output orientation, highlighting proportional reduction of excessive inputs (input slacks) or proportional augmentation of lacking outputs (output slacks). Nonetheless, for the sake of healthcare managers, some models were created combining both input reduction and output augmentation simultaneously by decreasing input slacks and increasing output slacks. In DEA literature, such models are designated as ADD models or non-oriented models (Ozcan, 2008).

Finally, when choosing between input orientation or output orientation, analysts view on which parametres healthcare managers are able to control is crucial (Ozcan, 2008). For example, a hospital may fix the amount of inputs in a given period for a specific service, so technical efficiency is measured by analysing the output expansion without altering the input set. Hence, an output-oriented model is advised. Another case is related with a hospital's contractual arrangement in terms of the number of patients treated. By considering the amount of inputs that could be reduced in the name of maintaining the output target, an input-oriented model is recommended.

3.3.1.5 Basic frontier models

Taking into account everything asserted in the previous two paragraphs, a handful of DEA basic frontier models can now be briefly presented. The use of these types of models depends, naturally, on the conditions of the problem at hand, but they all are able to be identified based on the scale and orientation of the model.

If one assumes that scale economies do not change with the increase in service facility size and that all DMUs are operating at an optimum scale, then CRS-type DEA models are a convenient choice. Otherwise, VRS models ought to be applied. Input orientation or output orientation depend on the output or input restrictions bestowed by healthcare managers, as described in Subsection 3.3.1.4. As promised, Figure 3.5 schematises the four essential envelopment models described above.

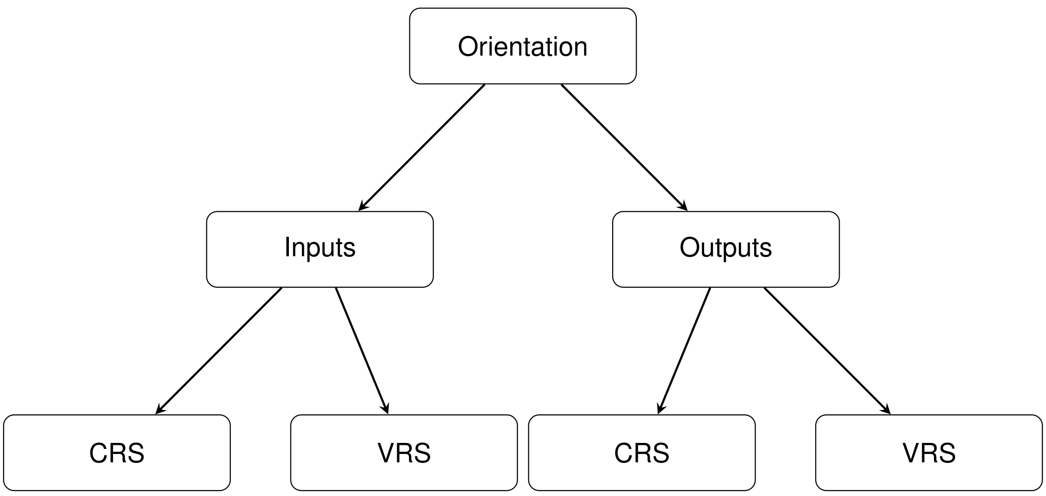

Figure 3.5: Basic DEA model classifications (adapted from Jacobs *et al.*, 2006).

Note that the scheme of Figure 3.5 is a rather primitive and elementary approach on the choosing of a DEA model. In practice, copious other factors are taken into consideration in an order that does not necessarily follow the one depicted in the design. For instance, the model selected for this research bore in mind the need for slacks and the usage of utilities, so a VRS ADD model was the chosen concept.

3.3.2 DEA in healthcare

Challenges such as those mentioned in Section 2.1 were overflowing healthcare organisations in the 1980s just as healthcare managers were pursuing efficiency improvements, particularly in the United States of America. A few years later, governmental initiatives extended the fixed pricing mechanism of diagnostic-related groupings (Fetter, Shin, Freeman, Averill, & Thompson, 1980) of the previous decade to physicians' services through resource-based relative schedule. In line with Ozcan (2008), albeit suchlike pricing mechanisms influenced the control of the amounts paid to healthcare organisations and healthcare professionals, a deeper insight into physician practice behaviour and disease treatment protocols development also took place, and it was the theoretical progress related to these advancements that led to the DEA approach flowering outset by Charnes *et al.* (1978). The efficiency measurement of DMUs using this non-parametric linear programming methodology was the primary focus of these researchers, and the efficiency frontier development, given an optimised weighted output/input ratio of each DMU (accounting for the fact that this ratio can be equal, but never exceed, unity for any other DMU in the data set), was a tremendous breakthrough at the time.

In the context of healthcare, the measurement of routine nursing service efficiency (Nunamaker, 1983) was DEA's first incursion in this domain and, from then onward, this frontier technique has become the dominant approach to efficiency measurement in healthcare, among other areas (Hollingsworth & Smith, 2003). *De facto*, assessing overall hospital efficiency was carried out forthwith with the efforts of Sherman (1984) and, from then on, multiple other researchers followed the same path. Indeed, one of the goals of the research conducted in this dissertation is to assess the overall performance of several secondary healthcare providers.

3.3.3 Preference information incorporation in DEA

Originally, since Charnes *et al.* (1978), DEA calculations were value-free, *i.e.*, they were purely objective and free from criteria imposed by subjective standards. However, attempts were made in the direction of incorporating preference information. According to Halme, Joro, Korhonen, Salo, & Wallenius (1999), such models are broken down into two classes:

- Models that use preference information to set targets for inefficient DMUs;
- Models that use preference information to produce more meaningful efficiency scores.

The development of target-setting models was cultivated primarily by Golany (1988) and Thanassoulis & Dyson (1992), the former with a view to allow the decision-maker to select the preferred set of output levels given the input levels of a DMU and the latter to estimate alternative input/output target levels in favour of rendering efficient relatively inefficient DMUs.

As for the efficiency score models, restricting weight flexibility was the conventional approach (Thompson, Singleton, Thrall, & Smith, 1986; Dyson & Thanassoulis, 1988; Charnes, Cooper, Wei, & Huang, 1989, 1990; Thompson, Langemeier, Lee, Lee, & Thrall, 1990; Wong & Beasley, 1990). This resulted in the reduction of the number of efficient DMUs and can be categorised into three groups (Joro & Korhonen, 2015):

- Direct and relative weight restriction methods (absolute values or weight ratios are the typical configurations used in multiplier DEA model);
- Cone ratio DEA (seen as a generalisation of relative weight restrictions);
- Contribution restriction methods (absolute or relative restrictions are based on products of the dual weights and the input and/or output quantities).

Nonetheless, Zhu (1996) initiated preference information incorporation in DEA to calculate efficiency scores, in spite of Golany & Roll (1994) having created a DEA model that incorporated preference information in the form of hypothetical DMUs a couple of years earlier.

The origins of preference information incorporation in Multi-Criteria Decision-Making (MCDM) literature are planted in the aforementioned interactive multi-objective linear programming (MOLP) procedure by Golany (1988).

Additionally, numerous arguments can be found against using importance weights to elicit and represent the decision-maker's preference information (Steuer, 1986; Wierzbicki, 1986; Korhonen & Wallenius, 1989; Korhonen, Silvennoinen, Wallenius, & Öörni, 2012), since a greater importance is not always denoted by a larger weight. That is how Halme *et al.* (1999) came up with the idea of the *Most Preferred Solution* (MPS) - an input/output vector on the efficient frontier that explicitly incorporates the decision-maker's preference information in the efficiency analysis. Theoretically, the MPS is the point at which the decision-maker's implicitly known value function has its maximum value and the interactive MOLP procedure terminates. It is with the knowledge of the MPS that the value function is approximated using tangent cones, being the efficiency of each DMU computed with respect to their respective tangent cone. This approach yields the so-called (in)efficiency scores, due to the fact that the efficiency of each DMU is determined by approximation.

On a slightly different topic, in the last few years, a handful of researchers have been employing a different method to DEA in the interest of considering the interactions between inputs and outputs - the Choquet integral. Ji, Liu, Qiu, & Lin (2015) and Xia & Chen's (2017) work is in line with other authors who have been paying attention to combine the Choquet integral with classical MCDM methods, such as ELECTRE (Figueira, Greco, & Roy, 2009) and PROMETHEE (Corrente, Figueira, & Greco, 2014). However, none of these investigators contemplated the usage of the Choquet integral method both as a way to deal with the aforementioned interactions and, most of all, as a tool to incorporate preference information of a body of decision-makers. It is in this knowledge gap that this dissertation dwells.

In Chapter 4, the way the decision-maker's preference information is incorporated in the model developed in this dissertation will be enlightened, but, suffice it to say, the Choquet integral is the tool integrated in DEA that considers variable interaction and is involved in producing more meaningful efficiency scores.

3.4 Summary

This chapter has introduced a handful of concepts of performance measurement in healthcare, including two dimensions of performance: efficiency and effectiveness. To evaluate performance, a series of methods were introduced alongside their strengths and weaknesses. Ultimately, the DEA technique was likewise made acquainted with, not to mention its model orientation and returns-to-scale notions, as well as the singular circumstances of preference information incorporation. The next chapter will be devoted to explaining in detail how everything is assembled into the model conceived in this dissertation.

Chapter 4

Methodology

Chapter 4 lays out the reasoning behind the construction of the model formulated in the scope of this dissertation. After an extensive section (Section 4.1) describing the methodologies in which the PRICDEA model was built on, a second section (Section 4.2) reports the implementation of its process and inherent procedures.

4.1 Methodology choice

In healthcare, DEA is considered to be the leading performance evaluation approach, as stated in Subsection 3.3.2. Indeed, as a method that evaluates the relative efficiency of DMUs using multiple inputs to produce multiple outputs, there are several types of models, *vide supra*, that can be used to determine the DMUs that form the efficient frontier. Particularly, given the complexity of the underlying application area, the choice in this work relied on a variant of the ADD model (which assumes a VRS) based on the 'two-phase method' developed by Gouveia, Dias, & Antunes (2008) in healthcare DEA. However, these researchers did not consider a crucial aspect of decision-making: interaction among criteria. Thus, by making use of a renowned multiple criteria preference aggregator - the Choquet integral (Choquet, 1954) -, interaction between a few numbers of pairs of criteria can now be taken into account directly, based on the work by Bottero, Ferretti, Figueira, Greco, & Roy (2018), and indirectly, adapted from the linear program developed by Marichal & Roubens (2000). Moreover, a method for the computation of utilities is also presented.

In this fashion, a 'Choquet integral-modified two-phase method' was developed, bearing in mind preference information incorporation, in order to provide a state-of-the-art non-parametric efficiency evaluation of secondary healthcare providers of the Portuguese SNS.

4.1.1 The 'two-phase method'

Quite a few authors have exploited the links between DEA and MCDM. For example, Joro, Korhonen, & Wallenius (1998) and Halme *et al.* (1999) related DEA with MOLP, Doyle & Green (1993), Stewart (1996), and Bouyssou (1999) related DEA and discrete multiple criteria problems, and Athanasopoulos

& Podinovski (1997) have found relationships between DEA and MCDM with imprecise information. Specifically, Gouveia *et al.* (2008) have linked DEA and MCDM with DMUs playing the role of decision alternatives. These researchers developed an innovative approach on the use of the additive DEA model in the interest of overcoming some of its flaws employing concepts from Multi-Attribute Utility Theory (MAUT) with imprecise information. In essence, their idea was to convert input and output factors into utilities aggregated by additive utility functions (Keeney, Raiffa, & Rajala, 1979) and then let each DMU choose the weights associated with those functions to minimise utility difference to the best DMU. The result was an ADD DEA model with oriented projections with a straightforward rationale for its efficiency measures, whose reasoning will be explained thenceforth.

At heart, the main goal of the VRS ADD model is to maximise the L_1 distance of the DMU under analysis (X_k, Y_k) to the projected point $(\hat{X}_k, \hat{Y}_k)$ on the efficient frontier. If and only if the optimal value of the primal ADD model is null and the slack values are also null, the projected point and the point under analysis coincide. In this case, the DMU$_k$ is efficient. If DMU$_k$ does not lie on the efficient frontier, it is deemed inefficient. This process is repeated n-times, once for each DMU$_k$ to be evaluated, partitioning the DMU set into two subsets: the efficient ones and the inefficient ones. In the case of the dual ADD model, a geometric interpretation is yielded. The linear programming formulations of these primal ADD$_{\text{P}}$ and dual ADD$_{\text{D}}$ ADD models (envelopment and multiplier forms, respectively) to evaluate the efficiency of DMU$_k$ are

$$\begin{aligned} \min \quad & z_k = -(\mathbf{1}s + \mathbf{1}e) \\ \text{s.t.} \quad & \mathbf{Y}\lambda - s = Y_k \\ & -\mathbf{X}\lambda - e = X_k \\ & \mathbf{1}\lambda = 1 \\ & \lambda \geqslant 0,\, e \geqslant 0,\, s \geqslant 0 \end{aligned} \tag{4.1}$$

and

$$\begin{aligned} \max \quad & w_k = \mu^T Y_k - \nu^T X_k + \omega_k \\ \text{s.t.} \quad & \mu^t \mathbf{Y} - \nu^T \mathbf{X} + \omega_k \mathbf{1} \leqslant 0 \\ & \mu^T \geqslant \epsilon \mathbf{1},\, \nu^T \geqslant \epsilon \mathbf{1}\,. \end{aligned} \tag{4.2}$$

Nevertheless, three issues can be identified regarding the ADD DEA model (Gouveia et al., 2008):

- All DEA models are projection mechanisms, and the projections of the inefficient DMUs on the efficient frontier depend on the scales used to measure each input or output;
- The efficiency measurement is quite pessimistic, due to the L_1 distance maximisation;
- The efficiency measure does not have an intuitive interpretation.

On that account, Ali, Lerme, & Seiford (1995) presented a variant of the ADD model with oriented projections, which Gouveia *et al.* (2008) named 'weighted ADD model', whose primal ADD$_{\text{WP}}$ and dual

ADD$_{\text{WD}}$ linear programming formulations to evaluate the efficiency of DMU$_k$ are

$$\begin{aligned} \min \quad & z_k = -(\mathbf{u}^k s + \mathbf{v}^{\mathbf{k}} e) \\ \text{s.t.} \quad & \mathbf{Y}\lambda - s = Y_k \\ & -\mathbf{X}\lambda - e = -X_k \\ & \mathbf{1}\lambda = 1 \\ & \lambda \geqslant 0,\, e \geqslant 0,\, s \geqslant 0 \end{aligned} \tag{4.3}$$

(4.4)

and

$$\begin{aligned} \min \quad & \mu^T Y_k - \nu^T X_k + \omega_k \\ \text{s.t.} \quad & \mu^t \mathbf{Y} - \nu^T \mathbf{X} + \omega_k \mathbf{1} \leqslant 0 \\ & \mu^T \geqslant \mathbf{u}^k \\ & \nu^T \geqslant \mathbf{v}^k\,. \end{aligned} \tag{4.5}$$

This alternative introduces the new parameters $\mathbf{u}^k$ and $\mathbf{v}^k$. These vectors are the coefficients of the objective function in the ADD$_{\text{WP}}$ formulation and are responsible for providing and determining the directions of the projection. Curiously, setting equal weights in the ADD$_{\text{WP}}$ (*e.g.*, $\mathbf{u}^k = 1$ and $\mathbf{u}^k = 1$) yields the ADD$_{\text{P}}$ model. Bottom line, the 'weighted ADD model', when compared to the ADD model, has the upper hand of allowing the flexibility of changing the direction of the projection, possibly less pessimistic than the one proposed by the latter.

Yet, the question arises: how are $\mathbf{u}^k$ and $\mathbf{v}^k$ chosen? The answer lies in the 'two-phase method'. Gouveia *et al.* (2008) begin to address this method by considering the set of DMUs as being a set of alternatives to be evaluated according to an additive MAUT model. The utility of each DMU is

$$U(\text{DMU}_i) = \sum_{j=1}^{q} w_j u_j(\text{DMU}_i), \tag{4.6}$$

where the scale coefficients $w_1, ..., w_q$ are the weights of the utility functions. Then, let d denote the distance between the utility difference between DMU$_k$ and the remaining alternatives for the sake of computing the the vector w of utility function weights that minimises this distance. Accordingly, the following program emerges:

$$\begin{aligned} \min \quad & d \\ \text{s.t.} \quad & U(\text{DMU}_i) - U(\text{DMU}_k) \leqslant d \\ & i = 1, ..., n \end{aligned} \tag{4.7}$$

which can be written as

$$\begin{aligned} \min \quad & d \\ \text{s.t.} \quad & \sum_{j=1}^{q} w_j u_j(\mathsf{DMU}_i) - \sum_{j=1}^{q} w_j u_j(\mathsf{DMU}_k) \leqslant d \\ & i = 1, ..., n \\ & \sum_{j=1}^{q} w_j = 1 \\ & w_j \geqslant 0, \forall j = 1, ..., q\,.n \end{aligned} \tag{4.8}$$

If the optimal value d^* of the objective function is null, then DMU_k is efficient; otherwise, it is inefficient and d^* is the minimum utility difference to the best DMU (*i.e.*, the DMU with the highest global utility).

From now on, the two obvious phases of the 'two-phase method' (Gouveia et al., 2008) are in a position to be addressed, *viz.*:

Phase 1 Convert inputs and outputs in utility scales, and compute the efficiency measure d of each DMU and the vector of weights associated with this utility difference using Equation (4.7);

Phase 2 Solve the 'weighted ADD model' using the weighting vector resulting from Phase 1, and determine the corresponding projected point of the DMU under evaluation.

Furthermore, since utilities are regarded as outputs, the primal formulation actually becomes

$$\begin{aligned} \min \quad & z_k = -\sum_{j=1}^{q} w_j s_j \\ \text{s.t.} \quad & \mathbf{Y}\lambda - s = Y_k \\ & \mathbf{1}\lambda = 1 \\ & \lambda \geqslant 0,\ s \geqslant 0 \end{aligned} \tag{4.9}$$

Thus, (4.7) is solved in Phase 1 and the simplified 'weighted ADD model' is solved in Phase 2 using the optimal weighting vector w^* obtained in Phase 1.

4.1.2 The Choquet integral

The concept of measurement was initiated in the 1950s by Choquet (1954). In his seminal book chapter, the French mathematician considered that non-linear relationships, especially interactions among attributes, could use non-linear integrals as data aggregation tools. Indeed, the weighted arithmetic mean was the most common aggregation tool used in MCDM until recently. Since not only it, but also none of the classical aggregators incorporate interactions among criteria, Sugeno (1974, 1977) came up with the concept of fuzzy integrals, which are defined from the concept of fuzzy measure. As a matter of fact, the Choquet integral is an example of a fuzzy integral, that was developed as a generalisation of the Lebesgue integral, whose fuzzy measure is known as the Choquet capacity.

Distinctively from the usual weighted-sum, the Choquet integral uses those capacities μ to compute an overall weight $\mu(T)$ of each subset T of the criteria set G. If the overall weight is different from the sum of weights $\mu(g_i)$ of the criteria belonging to subset T, one must infer that it is the result of some sort of interaction among criteria. Effectively, considering a generous amount of interacting criteria in an analysis is an ambitious choice that leads to a challenging interpretation and a significant reasoning effort from the decision-maker, so only the interaction between a number of pairs of criteria is acknowledged. In fact, according to Bottero *et al.* (2018), interaction between a pair of criteria g_i and g_j may be categorised as:

- Having no interaction: $\mu(g_i, g_j) = \mu(g_i) + \mu(g_j)$;
- Having a mutually-strengthening effect (synergy): $\mu(g_i, g_j) > \mu(g_i) + \mu(g_j)$;
- Having a mutually-weakening effect (redundancy): $\mu(g_i, g_j) < \mu(g_i) + \mu(g_j)$.

Theoretical and practical issues regarding interacting criteria phenomena have a solid foundation thanks to the works of Murofushi & Soneda (1993), Grabisch (1996a), and Grabisch & Roubens (1999).

Mathematically, there are plenty of intrinsic concepts to address about the Choquet integral. Let A denote a set containing m actions $a_1, ..., a_k, ..., a_m$ and G the criteria set with n criteria, $g_1, ..., g_i, ..., g_n$. The performance of action a_k on criterion g_i is written as $g_i(a_k)$ and the utility of this performance is drafted as $u_i(g_i(a_k))$. In literature, authors often employ a simpler notation for the sake of clarity, but, as everything in life, the concept of clarity depends on the mind needing clarification and the author of this dissertation found this notation quite more enlightening. Moving on, the theory whirling capacities is based on Grabisch (1996b), who considers that a capacity is a set function $\mu : 2^G \to [0, 1]$ on the power set 2^G, *i.e.*, the set of all subsets of G, that satisfies the following properties:

i. Boundaries: $\mu(\emptyset) = 0$ and $\mu(G) = 1$;

ii. Monotonicity: $\forall S \subseteq T \subseteq G : \mu(S) \leqslant \mu(T)$.

As a deduction from the first property, the values $\mu(S)$ assigned by the capacity μ to all other $2^{|G|} - 2$ subsets S of G have to be defined.

All things considered, the Choquet integral of an action $a_k \in A$ and a capacity μ on 2^G is defined as

$$C_\mu(a_k) = \sum_{i=1}^{n} \left(u_i(g_i(a_k)) - u_{i-1}(g_{i-1}(a_k))\right) \mu(G_i), \tag{4.10}$$

where $u_1(g_1(a_k)), ..., u_n(g_n(a_k))$ are the utilities of criteria from G, reordered in such a way that $u_1(g_1(a_k)) \leqslant ... \leqslant u_i(g_i(a_k)) \leqslant ... \leqslant u_n(g_n(a_k))$, and $G_i = g_i, ..., g_n$ for $i = 1, ..., n$, with $u_0(g_0(a_k)) = 0$.

For a matter of clarity, a simple practical application of the Choquet integral in an apocryphal hospital services' scenario is delineated below, not only for providing an empirical elucidation on the concepts inherent to this multiple criteria preference aggregator, but also for giving a taste of its actual application further ahead.

Example 4.1. Let

$$C_{\mu}(x) = \sum_{i=1}^{n} \left(x_{\tau(i)} - x_{\tau(i-1)}\right) \mu(\tau(i), ..., \tau(n)),$$

be the Choquet integral with $x : (x_1, ..., x_n) \rightarrow \mathbb{R}$ and a certain fuzzy measure μ, where τ is a permutation of N such that $x_{\tau(1)} \leqslant x_{\tau(2)} \leqslant ... \leqslant x_{\tau(n-1)} \leqslant x_{\tau(n)}$ and $x_{\tau(0)=0}$ will be disclosed. Suffice it to say that this example is entirely fictitious. Accordingly, Table 4.1 displays the performance values of three hospital services (A, B, and C) being analysed with regard to three criteria (I, II, and III), so that the performance assessment of each hospital is conducted using the aforesaid multiple criteria preference aggregation model. Criteria I and II vary in the interval $[0, 20]$ and criterion III in $[0, 100]$. Be that as it may, commensurate scales are needed to the correct application of the Choquet integral, in such manner that Table 4.2 presents the data of Table 4.1 with scales whose values are contained in $[0, 20]$. If the Choquet capacity $\mu : 2^N \rightarrow [0, 1]$ is defined by $\mu(N) = \mu(\{I, II\}) = 1$, $\mu(\emptyset) = \mu(\{I\}) = 0$, and $\mu(\{II\}) = \mu(\{III\}) = \mu(\{II, III\}) = \mu(\{I, III\}) = \frac{1}{2}$, then the results are rather straightforward:

Table 4.1: Example 4.1 performance table with incommensurate scales.

<table>
<tr><th></th><th>I: Occupancy</th><th>II: Readmissions</th><th>III: Waiting time before surgery</th></tr>
<tr><td>A: Cardiology</td><td>17</td><td>||||| ||||</td><td>60</td></tr>
<tr><td>B: Neurology</td><td>11</td><td>||||| ||</td><td>70</td></tr>
<tr><td>C: Rheumatology</td><td>10</td><td>||||| |||</td><td>45</td></tr>
</table>

Table 4.2: Example 4.1 performance table with commensurate scales.

	I: **Occupancy**	II: **Readmissions**	III: **Waiting time before surgery**
A: **Cardiology**	17	9	12
B: **Neurology**	11	7	14
C: **Rheumatology**	10	8	9

$$\begin{aligned} C_{\mu}(A) &= \sum_{i=1}^{3} \left(x_{\tau(i)} - x_{\tau(i-1)}\right) \mu(\tau(i), ..., \tau(3)) = \\ &= (9-0) \times \mu(N) + (12-9) \times \mu(\{I, III\}) + (17-12) \times \mu(\{I\}) = \\ &= 10.5 \end{aligned}$$

$$\begin{aligned} C_{\mu}(B) &= \sum_{i=1}^{3} \left(x_{\tau(i)} - x_{\tau(i-1)}\right) \mu(\tau(i), ..., \tau(3)) = \\ &= (7-0) \times \mu(N) + (11-7) \times \mu(\{I, III\}) + (14-11) \times \mu(\{III\}) = \\ &= 10.5 \end{aligned}$$

$$
\begin{aligned}
C_\mu(C) &= \sum_{i=1}^{3} \left(x_{\tau(i)} - x_{\tau(i-1)}\right) \mu(\tau(i), ..., \tau(3)) = \\
&= (8-0) \times \mu(N) + (9-8) \times \mu(\{I, III\}) + (10-9) \times \mu(\{I\}) = \\
&= 8.5
\end{aligned}
$$

As a deduction, and in line with the MCDM relation of preference $\succeq_x$ over N, it is possible to state that $A \sim_x B$, $A \succ_x C$, and $B \succ_x C$.

Moreover, it is usual in Choquet integral literature to formulate Equation (4.10) in its Möbius representation. Hence, given a capacity μ on 2^G, its Möbius representation is a function $m : 2^G \to \mathbb{R}^n$ such that, for all $S \subseteq G$,

$$\mu(S) = \sum_{T \subseteq S} m(T), \tag{4.11}$$

thus one has that

$$m(S) = \sum_{T \subseteq S} (-1)^{|S-T|} \mu(T), \tag{4.12}$$

and the properties stated above can be redefined as

i'. Boundaries: $m(\emptyset) = 0$ and $\sum\limits_{T \subseteq G} m(T) = 1$;

ii'. Monotonicity: $\forall i \in G$ and $\forall R \subseteq G : m(g_i) + \sum\limits_{T \subseteq R} m(T \cup g_i) \geqslant 0$;

where R is another subset of criteria that, alongside with S, implies that $\mu(R \cup S) = \mu(R) + \mu(S)$ whenever $R \cap S = \emptyset$, meaning that a fuzzy measure is additive. On that account, the Choquet integral's Möbius representation m of the capacity μ is expressed as

$$C_\mu(a_k) = \sum_{T \subseteq G} m(T) \min_{g_i \in T} \{u_i(g_i(a_k))\}. \tag{4.13}$$

Furthermore, it seems evident that a decision-maker is not able to provide information on 2^n coefficients of n criteria to the extent of defining the fuzzy measure μ on every subset. To overpower this problem, Grabisch (1997) came up with the idea of a k-order fuzzy measure. A fuzzy measure's polynomial expression has a noticeable additive linear representation (*e.g.*, $f(a_k) = \sum_{i=1}^{n} u_i(g_i(a_k))\omega_i$, where ω is a weighting vector). By extension, it is not unwise to think of a fuzzy measure having a polynomial representation of degree 2, 3, or any fixed integer k. Ergo, a k-order fuzzy measure represents a k-order approximation of its polynomial expression in the neighbourhood of the origin. In line with Grabisch (1997) and Marichal & Roubens (2000), this approach is confined to the 2-order case, because it allows to model criteria interaction while not increasing mathematical and computational complexity. That being said, let O denote the set of pairs of criteria g_i and g_j. For all $S \subseteq G$, the coefficients required to define the fuzzy measure are

$$\mu(g_i) = m(g_i), i \in G \text{ and}$$

$$\mu(g_i, g_j) = m(g_i) + m(g_j) + m(g_i, g_j),\ g_i, g_j \subseteq G,$$

being the other coefficients computed by

$$\mu(S) = \sum_{g_i \in S} m(g_i) + \sum_{g_i, g_j \subseteq S \; g_i, g_j \in O} m(g_i, g_j), S \subseteq N, |S| \geqslant 2. \tag{4.14}$$

and

$$\mu(G) = \sum_{g_i \in G} m(g_i) + \sum_{g_i, g_j \in O} m(g_i, g_j) = 1, \tag{4.15}$$

so the Choquet integral can be reformulated as

$$C_\mu(a_k) = \sum_{g_i \in G} m(g_i) u_i(g_i(a_k)) + \sum_{g_i, g_j \in O} m(g_i, g_j) \min\{u_i(g_i(a_k)), u_j(g_j(a_k))\}. \tag{4.16}$$

At long last, in this context, the conditions to define a 2-order fuzzy measure become

$$\begin{cases} m(\emptyset) = 0 & \\ \sum_{g_i \in G} m(g_i) + \sum_{g_i, g_j \subseteq G} m(g_i, g_j) = 1 & \\ m(g_i) \geqslant 0, & \forall g_i \in G \\ m(g_i) + \sum_{j \in O} m(g_i, g_j) \geqslant 0, & \forall g_i \in G, \forall O \subseteq G \setminus g_i. \end{cases} \tag{4.17}$$

With this in mind, Marichal & Roubens (2000) formalised a linear program to obtain the 2-order fuzzy measure contemplating the role of a decision-maker able to rank the criteria and the interactions between them. In terms of the Möbius representation, their model is given as follows:

$$
\begin{aligned}
\max \quad & z = \varepsilon && \text{(4.18)}\\
\text{s.t.} \quad & \text{Partial semiorder with threshold } \delta\text{:}\\
& \bullet\ C_\mu(a_a) - C_\mu(a_b) \geqslant \delta + \varepsilon, \text{ if } a_a \succ_A a_b\\
& \bullet\ -\delta \leqslant C_\mu(a_a) - C_\mu(a_b) \leqslant \delta, \text{ if } a_a \sim_A a_b\\
& \text{Ranking of criteria:}\\
& \bullet\ m(g_i) - m(g_j) \geqslant \varepsilon, \text{ if } g_i \succ_G g_j\\
& \bullet\ m(g_i) = m(g_j), \text{ if } g_i \sim_G g_j\\
& \text{Ranking of pairs of criteria:}\\
& \bullet\ m(g_{ij}) - m(g_{kl}) \geqslant \varepsilon, \text{ if } g_{ij} \succ_P g_{kl}\\
& \bullet\ m(g_{ij}) = m(g_{kl}), \text{ if } g_{ij} \sim_P g_{kl}\\
& \text{Sign of some interactions:}\\
& \bullet\ m(g_{ij}) \geqslant \varepsilon \text{ (resp. } \leqslant -\varepsilon), \text{ if } m(g_{ij}) > 0 \text{ (resp. } < 0)\\
& \bullet\ m(g_{ij}) = 0, \text{ if } m(g_{ij}) = 0\\
& \text{Boundary and monotonicity conditions:}\\
& \bullet\ \sum_{g_i \in G} m(g_i) + \sum_{g_i, g_j \subseteq G} m(g_i, g_j) = 1\\
& \bullet\ m(g_i) \geqslant 0, \forall g_i \in G\\
& \bullet\ m(g_i) + \sum_{j \in O} m(g_i, g_j) \geqslant 0, \forall g_i \in G, \forall O \subseteq G \setminus g_i\\
& \text{Definition of } C_\mu\text{:}\\
& C_\mu(a_a) = \sum_{g_i \in G} m(g_i) u_i(g_i(a_a)) + \sum_{g_i, g_j \in O} m(g_i, g_j) \min\{u_i(g_i(a_a)), u_j(g_j(a_a))\}.
\end{aligned}
$$

where a_a and a_b are universal DMUs, g_i and g_j are common criteria, and g_{ij} and g_{kl} are generic interactions between two criteria, respectively. As for the ranks elected by the decision-maker, $\succ_A$ is the partial preorder on the set of alternatives A, $\succ_G$ is the partial preorder on the set of criteria G, and $\succ_P$ is the partial preorder on the set of pairs of interacting criteria P.

4.1.3 The PRICDEA (PReference information Incorporation using the Choquet integral in a DEA model) model

Now that both the 'two-phase method' and the Choquet integral have been thoroughly detailed, the heavy groundwork is laid so that the PRICDEA model is promptly defined at this instant. The concept behind the PRICDEA model is quite transparent and has been meticulously explained in the early chapters of this dissertation. Essentially, integrating the Choquet integral preference aggregation tool in the non-parametric performance evaluation DEA 'weighted ADD model' yields a utility-apt preference-information-incorporated interactive-variable-equipped method. Without delay, the PRICDEA model is implemented as follows:

Phase 1 Convert the indicators' performances in utility scales using the preference information of the decision-maker and infer the optimal Möbius coefficients based not only on those utilities, but also on the criteria interactions and ranks provided by the same expert via the adapted conjoint version of Equation (4.8) and Equation (4.18)

$$\begin{aligned}
\min \quad & d \\
\text{s.t.} \quad & C_\mu(a_x) - C_\mu(a_y) \leqslant d \\
& \text{Ranking of criteria:} \\
& \bullet\ m(g_i) - m(g_j) \geqslant \varepsilon, \text{ if } g_i \succ_G g_j \\
& \bullet\ m(g_i) = m(g_j), \text{ if } g_i \sim_G g_j \\
& \text{Ranking of pairs of criteria:} \\
& \bullet\ m(g_{ij}) - m(g_{kl}) \geqslant \varepsilon, \text{ if } g_{ij} \succ_P g_{kl} \\
& \bullet\ m(g_{ij}) = m(g_{kl}), \text{ if } g_{ij} \sim_P g_{kl} \\
& \text{Boundary and monotonicity conditions:} \\
& \bullet\ \sum_{g_i \in G} m(g_i) + \sum_{g_i, g_j \subseteq G} m(g_i, g_j) = 1 \\
& \bullet\ m(g_i) \geqslant 0, \forall g_i \in G \\
& \bullet\ m(g_i) + \sum_{j \in O} m(g_i, g_j) \geqslant 0, \forall g_i \in G, \forall O \subseteq G \setminus g_i
\end{aligned} \tag{4.19}$$

for $x = 1, ..., m$ and $y = 1, ..., m$, and a preset $\varepsilon = 0.001$, and using the Choquet integral definition declared in Equation (4.16).

Phase 2 Solve the modified rendition of the 'weighted ADD model' of Equation (4.9)

$$\begin{aligned}
\min \quad & z_k = -\sum_{j=1}^{n} M_j s_j \\
\text{s.t.} \quad & \sum_{i=1}^{m} U_{ij} \lambda_i - s_j = U_{kj} \leqslant 0 \\
& \sum_{i=1}^{m} \lambda_i = 1 \\
& \lambda_i \geqslant 0,\ s_j \geqslant 0 \\
& i = 1, ..., m,\ j = 1, ..., n\,.
\end{aligned} \tag{4.20}$$

using the Möbius coefficients M_j and the U_{ij} utilities' matrix from Phase 1, and determine the efficient frontier, the benchmarks for the inefficient DMUs, and the corresponding slacks.

The PRICDEA model can be more pleasantly presented graphically, by using a block diagram to discriminate its phases and inherent steps. Figure 4.1 contains the essential information for an unequivocal and incontrovertible explanation.

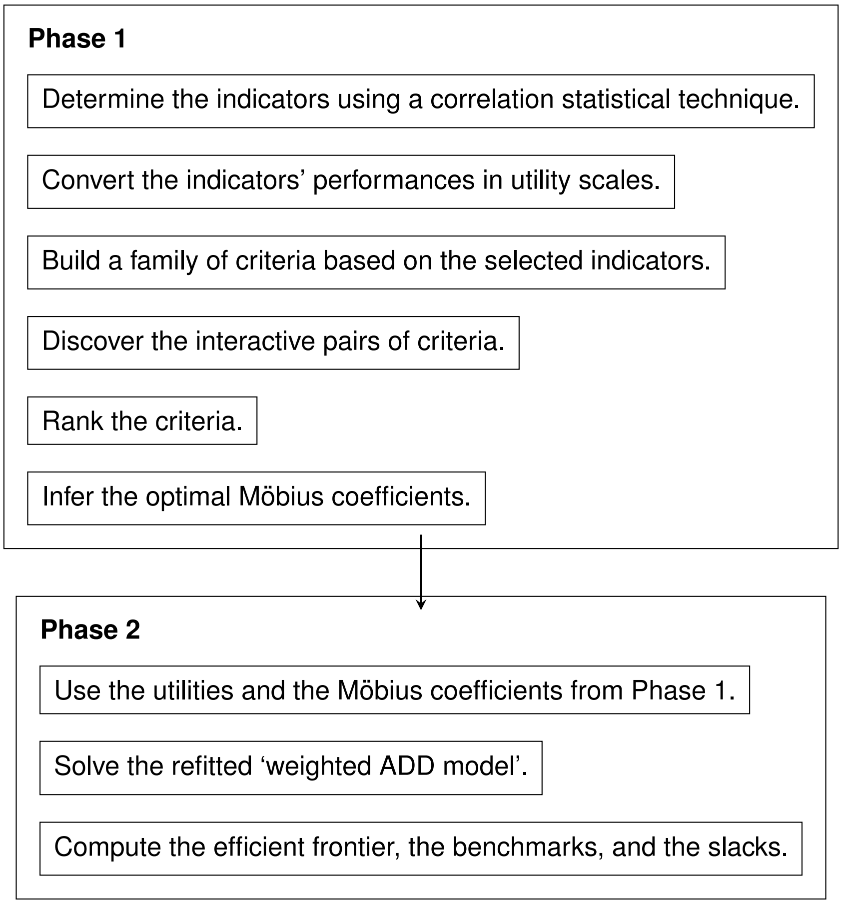

Figure 4.1: PRICDEA's graphical schematisation.

PRICDEA and built-in procedures will be exhaustively illuminated in the remainder of this chapter and the following one.

4.2 Methodology implementation

This section discusses the methodology of PRICDEA by and of itself, starting with an uncomplicated how-to and finishing with the methodologies behind the methodology.

4.2.1 PRICDEA how-to

Methodologically, delineating theoretical notions is essential to the understanding of the intrinsic mechanisms of a certain approach, but it is only half of the way. Perceiving how the method operates constitutes the other half. In Subsection 4.1.3, the logical and formalistic steps of PRICDEA were addressed, and, in this subsection, its operationalisation will be in focus.

Firstly, PRICDEA's Phase 1 is concerned with data conversion and preference information incorporation. As will be dug into in the following chapter, setting up a database is crucial for valuable results. With that taken care of, it is time to build the interval scale for each criterion, taking into account the decision-maker's preferences, and assign the utility values to the performances in-between the defined levels, so that all of the indicators can be compared to one another on the same scope. Finally, assessing the decision-maker's propensity towards those indicators and respective exposed interactions allows the computation of the optimal Möbius coefficients per DMU with the linear program in Equation (4.19), which will be used in the next phase of the PRICDEA model. These procedures are duly explained in Subsection 4.2.2.

Secondly, based on the optimal Möbius-transformed Choquet capacities and the utility values computed in the previous stage, the modified version of the 'weighted ADD model' is solved using the linear program in Equation (4.20) in Phase 2. This program returns not only the efficient frontier, but also the benchmark values for the inefficient DMUs and the slack values per indicator.

Nonetheless, since PRICDEA is a computationally demanding and experimental method, a choice was made to corroborate the results of its Phase 1 using the procedure adopted by Bottero *et al.* (2018) and described in Subsection 4.2.2.2 to compute what can be named as 'analogous Möbius coefficients', thus authenticating the validity of the model wielding the decision-maker's preference information likewise for increased consistency. Moreover, it is the questioning procedure described in this paragraph that is used in the ranking step of PRICDEA's Phase 1 to build the constraints of Equation (4.19) and infer the optimal Möbius coefficients, but, since it is already precised in that site, there is no need to revisit it.

Further details on the implementation of PRICDEA *per se* will be thoroughly unfolded in Section 5.5, being this subsection a clear and honest explanatory guide.

4.2.2 Procedures *de rigueur*

4.2.2.1 Creating the database

"*Scientia potentia est*" is a Latin aphorism that means "knowledge is power". Nowadays, in a digital world where information technologies are being developed at an astonishing level, this astute 1597 observation uttered by Sir Francis Bacon symbolises even more than back then. In fact, knowledge is the power of information, which means that "information is power" - the predominant source of power in modern society (Toffler, 1991).

Particularly, in the interest of evaluating the performance of Portuguese SNS secondary healthcare providers in the methodological context of this dissertation, it is indispensable to be in possession of the most accurate, legitimate, and well-founded data in order to generate *bona fide* information. Hence, creating a more fitting database resting on the *vide infra* benchmarking database is essential to a sound analysis. The results of this creation procedure will be addressed in Section 5.3.

Nonetheless, for now, the database is elementarily arranged in the way displayed in Table 4.3, with A the set of m institutions $a_1, ..., a_k, ..., a_m$ and G the set of n criteria $g_1, ..., g_i, ..., g_n$, following the nomenclature used by Bottero *et al.* (2018). It is not relevant to single out the identities of both the institutions and the criteria at this moment.

Table 4.3: Research database elementary representation.

		Access		Performance assistance			Safety			Volume and usage			Productivity			Economic-financial		
		g_1	g_2	g_3	(...)	g_{10}	g_{11}	(...)	g_{16}	g_{17}	(...)	g_{22}	g_{23}	(...)	g_{26}	g_{27}	(...)	g_{35}
Group A	-	-	-	-	(...)	-	-	(...)	-	-	(...)	-	-	(...)	-	-	(...)	-
Group B	a_1	96.16%	100.00%	76.09%	(...)	-	0.001182	(...)	-	0	(...)	0	79.971016	(...)	0.934539	2640	(...)	3.64%
	(...)	(...)	(...)	(...)	(...)	(...)	(...)	(...)	(...)	(...)	(...)	(...)	(...)	(...)	(...)	(...)	(...)	(...)
	a_9	66.24%	94.16%	83.20%	(...)	-	0.000000	(...)	-	0	(...)	1	72.831441	(...)	0.762694	-	(...)	-
Group C	a_{10}	96.02%	97.77%	85.16%	(...)	35.68%	0.001130	(...)	0.12%	0	(...)	0	86.396287	(...)	0.758759	-	(...)	-
	(...)	(...)	(...)	(...)	(...)	(...)	(...)	(...)	(...)	(...)	(...)	(...)	(...)	(...)	(...)	(...)	(...)	(...)
	a_{25}	50.75%	89.41%	79.93%	(...)	37.91%	0.000354	(...)	0.54%	0	(...)	0	68.346135	(...)	0.562452	2824	(...)	6.99%
Group D	a_{26}	83.66%	80.94%	87.79%	(...)	37.72%	0.000967	(...)	1.06%	25	(...)	540	96.453441	(...)	1.008700	2738	(...)	5.04%
	(...)	(...)	(...)	(...)	(...)	(...)	(...)	(...)	(...)	(...)	(...)	(...)	(...)	(...)	(...)	(...)	(...)	(...)
	a_{33}	54.42%	83.49%	75.11%	(...)	36.21%	0.003055	(...)	0.57%	1	(...)	550	81.311964	(...)	0.716776	2868	(...)	8.59%
Group E	a_{34}	95.92%	74.86%	85.11%	(...)	37.62%	0.002141	(...)	0.80%	89	(...)	740	75.164420	(...)	1.273103	3125	(...)	0.56%
	(...)	(...)	(...)	(...)	(...)	(...)	(...)	(...)	(...)	(...)	(...)	(...)	(...)	(...)	(...)	(...)	(...)	(...)
	a_{39}	67.64%	84.63%	77.36%	(...)	41.47%	0.000803	(...)	0.12%	32	(...)	473	87.765426	(...)	1.264281	2861	(...)	0.56%
Group F	a_{40}	99.73%	72.57%	83.66%	(...)	-	0.001666	(...)	-	0	(...)	0	148.935462	(...)	1.709612	2249	(...)	1.01%
	(...)	(...)	(...)	(...)	(...)	(...)	(...)	(...)	(...)	(...)	(...)	(...)	(...)	(...)	(...)	(...)	(...)	(...)
	a_{42}	92.96%	66.92%	69.47%	(...)	-	0.003124	(...)	-	0	(...)	0	143.625186	(...)	0.971468	2500	(...)	1.88%

4.2.2.2 Determining the analogous Möbius coefficients

Ascertaining the weights of criteria within the context of outranking methods was the result of a work conducted by Figueira & Roy (2002), where the two researchers proposed a modified version of the Simos' deck of cards method (Simos, 1990a, 1990b). One of the conclusions of this study asserted that the method *supra* could be extended to build other ratio and interval scales. In such wise, considering the context of this dissertation, this method was adapted to bear in mind the chosen indicators. Notwithstanding, the dialogue between the analyst and the decision-maker took into account the procedure ensued in Figure 4.2, based on Bottero *et al.*'s (2018) one.

Resting on the decision-maker's preferences assayed in the procedure above, the next logical step is to construct the Choquet capacities' ratio scale. The method for determining these capacities from a reference set of projects was adapted from Bottero *et al.* (2018), who established their process in Marichal & Roubens' (2000) extension of the swing weighting procedure (Von Winterfeldt & Edwards, 1986). In consonance with this, and assuming that the minimal and the maximal utilities of each criterion are 0 and 1, the reference set of projects was built as follows:

- n projects denoted by p_j, for all $j \in G$, and characterised by the vector $(0, ..., 0, u_j(p_j) = 1, 0, ..., 0)$, where p_j has the highest performance on criterion j and the lowest elsewhere;
- $|O|$ projects denoted by $p_k = p_{ij}$, for $k = n+1, ..., n+|O|$, and characterised by a vector with the same form $(0, ..., 0, u_i(p_i), 0, ..., 0, u_j(p_j) = 1, 0, ..., 0)$, for all $i, j \in O$, where p_k has the highest utilities on criteria i and j and the lowest elsewhere.

With that in mind, the method for assessing Choquet capacities μ and their Möbius representation m can be bestowed as a step-by-step procedure:

1. Consider the finite set of reference projects $P = p_1, p_2, ..., p_k, ..., p_t$, where $t = n + |O|$;
2. Consider the ranking provided by the decision-maker in the questioning procedure and denoted by $R_1, ..., R_h, ..., R_v$ (from the least to the most preferred projects); let r_h denote a project representative of projects in the equivalence class R_h, for $h = 1, ..., v$, and let e_h denote the number of blank cards between the equivalence classes R_h and R_{h+1}, for $h = 1, ..., v-1$;
3. Attribute a value to project r_1, *e.g.*, $w(r_1) = l$, assuming none of the projects has null utility;

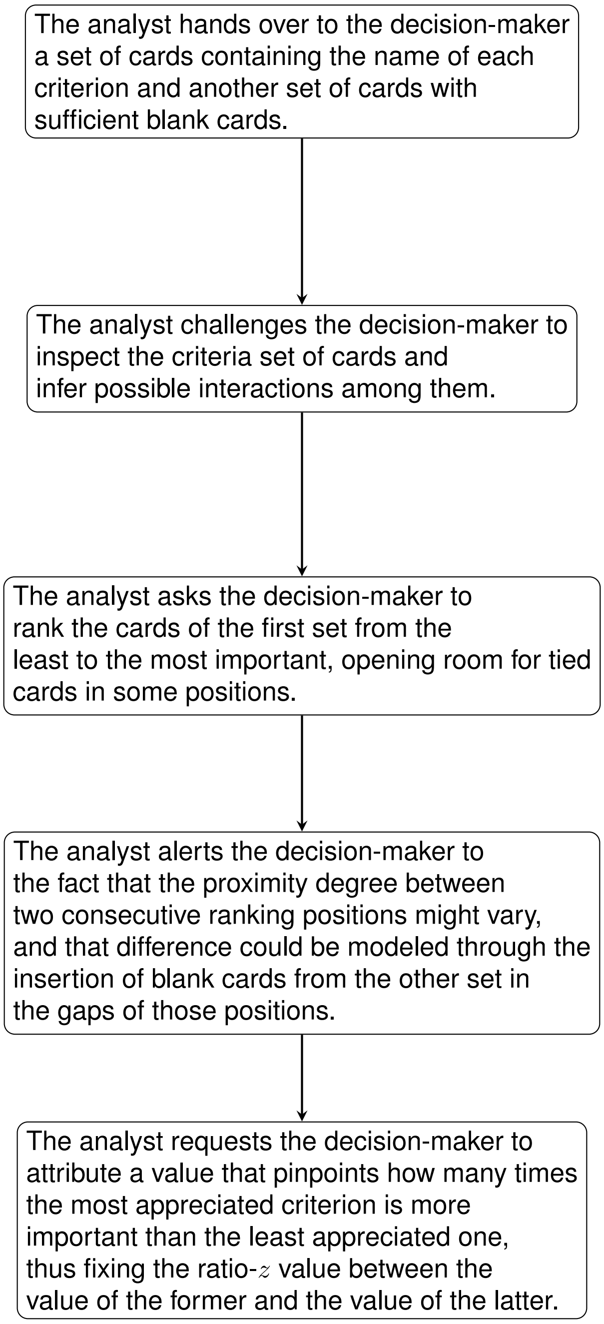

Figure 4.2: Questioning procedure for the decision-maker's project ranking.

4. Compute the value of each unit as

$$\alpha = \frac{l(z-1)}{s},$$

where

$$s = \sum_{h=1}^{v-1}(e_h + 1),$$

which means that α is computed by dividing the difference between the values of the most preferred projects and the least preferred projects;

5. Compute the values $w(r_h)$, for $h = 2, ..., v$, using $w(r_h) = l + \alpha \left(\sum_{j=1}^{h-1}(e_h + 1) \right)$;

6. Compute the value of each project $w(p_k) = w(r_h)$, for all $p_k \in R_h$ with $h = 1, ..., v$;

7. Compute the modified values $\overline{w}(p_k)$ as:

$$\overline{w}(p_k) = \begin{cases} w(p_k) \text{ if } k = i \in G \\ w(p_k) - w(p_i) - w(p_j) \text{ if } p_k = p_{ij}, \text{ for } i, j \in O \text{ and } k \geqslant n+1 \end{cases}$$

8. Compute the Möbius coefficients m_k using

$$m_k = \frac{\overline{w}(p_k)}{\sum_{j=1}^{t} \overline{w}(p_j)}$$

and the Choquet capacities μ_k as

$$\mu_k = \frac{w(p_k)}{\sum_{j=1}^{t} \overline{w}(p_j)},$$

where the coefficients m_k must meet the 2-order fuzzy measure terms established by Equation (4.17).

Note that the computations of step 3 to step 6 were performed by DecSpace - a web application that makes use of MCDM methods to support a decision process by giving a possible solution to a given problem - using the SRF software developed by Roy & Figueira (1998).

4.2.2.3 Building the interval scale

The Choquet integral uses utility values that are the levels of a common interval scale, generally within $[0, 1]$, which means that decoding the original criteria scales into a single common interval scale calls for the employment of a procedure that takes into consideration the decision-maker's preference intensity between consecutive intervals of the scale. In this dissertation, the procedure used by Bottero *et al.* (2018) to define an interval scale based on the underlying concepts of the deck-of-cards method by Figueira & Roy (2002) was adopted. In pursuance of building an interval scale, at least two reference levels must be defined to anchor the computations. Hence, the various steps of the procedure are subsequently enumerated (Bottero et al., 2018):

1. Consider a discrete scale of criterion g,

$$E_g = l_1, ..., l_k, ..., l_t,$$

where $l_1 \prec l_2 \prec ... \prec l_k \prec ... \prec l_t$ (once more, in this context, "$\prec$" means "strictly less preferred than");

2. Define two reference levels l_p and l_q, and assign them two frequently used utility values $u(l_p) = 0$ and $u(l_q) = 1$ (note that l_p and l_q very often coincide with l_1 and l_t);

3. Consider the ranking of the levels with a certain number of blank cards e_k in the intervals between every two consecutive levels l_k and l_{k+1}, with $k = 1, ..., t-1$, such that

$$l_1 e_1 ... l_p e_p l_{p+1} e_{p+1} ... l_k e_k l_{k+1} ... l_{q-1} e_{q-1} l_q ... l_{t-1} e_{t-1} l_t;$$

4. Consider only the levels between l_p and l_q and compute the unit valuation

$$\alpha = \frac{u(l_q) - u(l_p)}{h},$$

where

$$h = \sum_{k=p}^{q-1} e_k + 1;$$

5. Compute the utility value $u(l_k)$ for each level k, with $k = 1, ..., t$ as follows:

$$u(l_k) = \begin{cases} u(l_p) - \alpha \left(\sum_{j=k}^{p-1} e_j + 1 \right), & \text{for } k = 1, ..., p-1 \\ u(l_p) + \alpha \left(\sum_{j=p}^{k-1} e_j + 1 \right), & \text{for } k = p+1, ..., q-1, ..., q+1, ...t \end{cases}$$

Nevertheless, this procedure is suitable for discrete scales, which is not the case of any indicator of this dissertation's continuous-valued database. Ergo, to break down the intervals among 5 equidistant discrete levels corresponding to 0%, 25%, 50%, 75%, and 100% of each indicator was opted, whether the corresponding original value was a percentage or not (if not, it was converted to a percentage). Then, the utilities for levels 25%, 50%, and 75% were computed using the procedure stated above, assuming 0% and 100% as the two reference levels - the preference order of the discrete scale E of each indicator g_j varied according to the objective of minimising or maximising a given indicator.

4.2.2.4 Assigning utility values to the numerical scale levels

In line with the aforementioned procedure, computing the remaining utility values was pretty straightforward. As a matter of fact, the utility values assigned to the intermediate performances of criterion g_j in a given subinterval $g_j^l < g_j < g_j^u$ are defined by a linear interpolation. Equation (4.21) was the selected formulation to be applied.

$$u_j(g_j) = u_j(g_j^l) + \frac{g_j - g_j^l}{g_j^u - g_j^l} \left(u_j(g_j^u) - u_j(g_j^l) \right) \tag{4.21}$$

Further details on the befitting usage of the linear interpolation for assigning the utility values to the numerical scale levels will be addressed in Subsection 5.5.2.

4.3 Summary

This chapter addresses the basis of the PRICDEA model, its implementation, and built-in procedures. The two phases of PRICDEA include a Möbius coefficients' inference step and a performance evaluation step. Chapter 5 puts this model to the test in the framework of the problem described in Chapter 2.

Chapter 5

Case study

In the fifth chapter, the empirical application of the PRICDEA methodology is evaluated. Following the steps of Chapter 2, Chapter 5 provides an overview of the specific issue under analysis in Section 5.1, details the stakeholders in Section 5.2, and describes the process of data gathering and processing in Section 5.3. After the presentation of the variables of this case study (Section 5.4), the main results and discussion of the operationalisation of this performance assessment tool are explained in the penultimate section (Section 5.5).

5.1 Overview

As declared in Section 2.2, the Portuguese SNS was constituted in 1979 after almost fifty years of a dictatorial regime. By nature, it was based on the philosophy of a Beveridge system: primary and secondary healthcare providers are public entities governed by a Central Government that taxes citizens in exchange for funds to be distributed by different ministries (Lameire, Joffe, & Wiedemann, 1999), of which the Ministry of Health is a part of. Another characteristic of the Portuguese SNS is its social purpose, since it intends to provide a suitable and equitable care to the universe of beneficiaries, whether or not they are able to afford it. Ergo, the financial sustainability of each institution is, once more, proven to be vital to the operation of the whole system. However, recently, expenditures in health have been ascending as a result of, essentially, demographic shifts and technological advances, so the introduction of health reforms was imperious, not only in the name of reducing costs and waste of public funds, but also with a view to improving the efficiency and effectiveness of healthcare providers (Ferreira & Marques, 2017). These rehabilitation proposals included corporatisation (?, ?), vertical and horizontal merging of public healthcare providers, and creating public-private partnerships (Cruz & Marques, 2013; ?, ?).

Nowadays, corporate public sector institutions of the SNS are annually financed by prospective budgets brought to terms with the Ministry of Health taking into account the quantity and quality of provided services. Nevertheless, regardless of the objections surrounding the success of the aforementioned reforms, a study aimed at performance assessment in this sector is inescapable - and this is precisely where this dissertation emerges from.

5.2 Stakeholders and their representatives

In public policy-making, it is the actors and their behaviours that represent the heart of a model (Boerboom & Ferretti, 2014; Dente, 2014). These individuals or organisations are the ones who influence decisional outcomes since they go after goals regarding the problem and its possible solution (Dente, 2014). Thus, any actor with an inalienable interest in the decision process is named a stakeholder. The first fundamental step of a decision-making process developed to support public policy formulation is to identify the stakeholders and their objectives (Ferretti, 2016). In this dissertation, the stakeholders' preference information was incorporated in PRICDEA directly, from a participative approach developed through the focus group technique elaborated by Bottero *et al.* (2018), and indirectly, using a version of Marichal & Roubens' (2000) linear program for determining the weights of interacting criteria from a reference set.

Under the authority of fusing the knowledge and know-how of healthcare trivia and performance assessment, a single *virtuoso* Ministry of Health expert, proficient in the fields of health administration and health policy-making, was more than enough to be involved in the decision-making process. Nonetheless, the occasional contributions of an additional adroit academia expert, skilled in the area of performance assessment, were invaluable to counterbalance any possible ambiguity or bias and reach a fair consensus.

5.3 Data and sample

It should be clear by now that the main goal of the case study engaged in this dissertation is to evaluate the performance of secondary healthcare providers of the Portuguese SNS. Thus, indicators and respective values concerning the selected institutions were collected from the official benchmarking database, maintained by the Portuguese Central Health System Administration (ACSS, from the Portuguese abbreviation of *Administração Central do Sistema de Saúde*) at `http://benchmarking.acss.min-saude.pt/` - a public institute integrated in the indirect health State administration with administrative and financial autonomy -, to create the more applicable database mentioned in Subsection 4.2.2.1 and represented in Table 4.3.

5.3.1 Sample reduction

In line with what was affirmed above, defining the time interval was established as the first order of business. For that reason, a sole month could be the subject of this analysis, but why would a cross-section of an entire year be singled out when the accumulated results for the year itself could be evaluated? Moreover, seeing that 2017 was just a few months ago, the data validation process was yet to be completed, so 2016 was the year of choice.

Then, the sample (secondary healthcare providers) and respective variables (indicators) were attended to. Specifically, the ACSS benchmark bundles 42 institutions in 6 groups using hierarchical clustering (GROUP A to GROUP F) after variable standardisation and principal components analysis. However, on one hand, data from GROUP A institutions were not available, so they were not even included in the analysis. Moreover, in what concerns institutions: GROUP B's *Figueira da Foz District Hospital, Corporate Public Entity (CPE)* and *Santa Maria Maior Hospital, CPE* did not make the cut, since they both had an incomplete data set; the three institutions of GROUP F were removed from the sample, because IPOs have their specific technology of production (Ferreira, Nunes, & Marques, 2018); local health units are the result of vertical integration between one hospital and various primary healthcare centres, which means that comparing them to hospitals and hospital centres would be dishonest and biased; institutions managed as public-private partnerships have a considerable lack of data, so they had no use for the analysis. This lead to a total of 25 institutions (5 hospitals and 20 hospital centres) distributed across 4 groups (GROUP B to GROUP E). Yet, on another hand, the 35 indicators were categorised into 6 individual benchmark dimensions, *viz.*: ACCESS, PERFORMANCE ASSISTANCE, SAFETY, VOLUME AND USAGE, PRODUCTIVITY, and ECONOMIC-FINANCIAL. Finally, two indicators were eliminated from the SAFETY dimension (*Pressure ulcer rate* and *Central venous catheter-related bloodstream infections rate*), due to their values being almost null, as well as all of the VOLUME AND USAGE dimension indicators and the last two ECONOMIC-FINANCIAL dimension indicators (*Percentage of overtime costs and supplements in total personnel costs* and *Percentage of costs with services in total costs with personnel*) given their lack of meaning in such an analysis, resulting in 5 dimensions containing a sum of 25 indicators.

Be that as it may, such an amount of indicators was still considerable for a MCDM analysis, so a statistical correlation test was performed in MATLAB in pursuance of eliminating highly correlated variables taking advantage of the MATLAB native function `[R, P] = corrcoef(M)`. `corrcoef` returns the matrix of correlation coefficients `R` and the matrix of p-values `P`, where `M` is the 25×25 matrix with the values obtained from the elimination steps mentioned above, to test the hypothesis that there is no relationship between the indicators. If an off-diagonal element of `P` is smaller than the significance level (the default p-value, 0.05, was considered), then the corresponding correlation in `R` is considered significant. Thus, for each row of `P` where $P_{i,j} < 0.05$, for $i = 1, ..., 25$ and $j = 1, ..., 25$, this top-to-bottom elimination procedure was applied and yielded 9 indicators.

Nonetheless, *Costs with clinical consumption material per standard patient* was removed from the equation since they were already included in *Operating costs per standard patient*, and *Percentage of instrumented vaginal deliveries with 3^rd^ and 4^th^ degree lacerations* was deemed as non-significant for the analysis by the decision-maker. In addition, the decision-maker was adamant in the inclusion of the indicator *Average waiting time before surgery*, due to its compelling influence in the SNS performance. Since this indicator was one of the least correlated ones, the decision-maker's preference information was decided to be also incorporated in this stage.

Bottom line, the 200-entry sample is composed of 25 hospitals operating in 2016 according to 8 indicators. Note that pooling data over time assumes no technology change over the considered time period (?, ?; Chowdhury & Zelenyuk, 2016), but, since the time period of this case study is merely twelve months and no significant alterations on hospital administrations occurred to the extent of producing a relevant technology and productivity disparity, there is no margin for refutation. In brief, the database creation procedure mentioned in Subsection 4.2.2.1 and described in the present subsection intends to build a bullet-proof database of reduced sample size and improved results' resolution.

5.3.2 Decision-making units (DMUs)

In Chapter 3, all the details regarding the concepts and issues related with DMUs and indicators in the DEA framework were sweated over, as well as how they were chosen in a framework suchlike this in Chapter 4 and Subsection 5.3.1. At this moment, in line with the previous section, it can be asserted with certainty that twenty-five institutions among the original thirty-five were considered as the case study's DMUs. In particular, the twenty-five DMUs under consideration a_m, for $m = 1, ..., 25$, spread over four groups are:

1. GROUP B:
    1.1. *Póvoa de Varzim/Vila do Conde Hospital Centre, CPE*, a_1;
    1.2. *Médio Ave Hospital Centre, CPE*, a_2;
2. GROUP C:
    2.1. *Leiria Hospital Centre, CPE*, a_3;
    2.2. *Cova da Beira Hospital Centre, CPE*, a_4;
    2.3. *Barreiro/Montijo Hospital Centre, CPE*, a_5;
    2.4. *Setúbal Hospital Centre, CPE* (a_6);
    2.5. *Senhora da Oliveira (Guimarães) Hospital, CPE*, a_7;
    2.6. *Baixo Vouga Hospital Centre, CPE*, a_8;
    2.7. *Santarém District Hospital, CPE*, a_9;
    2.8. *Médio Tejo Hospital Centre, CPE*, a_{10};
    2.9. *Tâmega e Sousa Hospital Centre, CPE*, a_{11};
    2.10. *Entre Douro e Vouga Hospital Centre, CPE*, a_{12};
3. GROUP D:
    3.1. *Garcia de Orta Hospital, CPE*, a_{13};
    3.2. *Espírito Santo de Évora Hospital, CPE*, a_{14};
    3.3. *Trás-os-Montes e Alto Douro Hospital Centre, CPE*, a_{15};
    3.4. *Algarve University Hospital Centre, CPE*, a_{16};
    3.5. *Tondela-Viseu Hospital Centre, CPE*, a_{17};

3.6. *Vila Nova de Gaia/Espinho Hospital Centre, CPE*, a_{18};

3.7. *Hospital Fernando Fonseca, CPE*, a_{19};

4. GROUP E:

4.1. *Lisboa Central Hospital Centre, CPE*, a_{20};

4.2. *Lisboa Ocidental Hospital Centre, CPE*, a_{21};

4.3. *São João Hospital Centre, CPE*, a_{22};

4.4. *Porto Hospital Centre, CPE*, a_{23};

4.5. *Lisboa Norte Hospital Centre, CPE*, a_{24};

4.6. *Coimbra University Hospital Centre, CPE*, a_{25}.

5.4 Variables

Based on the ACSS retrieved and processed data, a set of indicators was defined in the next subsection, alongside its performance table. With that in mind, a set of criteria was erected and their basic statistics were also analysed.

5.4.1 Indicators

In consonance with the process described in Subsection 5.3.1, the eight yielded indicators g_n, for $n = 1, ..., 8$, can be enumerated as follows:

1. ACCESS:

1.1. *Number of non-urgent first medical appointments performed in adequate time per 100 first medical appointments*, g_1;

2. PERFORMANCE ASSISTANCE:

2.1. *Number of outpatient surgeries per 100 potential outpatient procedures*, g_2;

2.2. *Number of readmissions in 30 days after discharge per 100 inpatients*, g_3;

2.3. *Number of long-stay inpatients per 100 admissions*, g_4;

2.4. *Number of hip surgeries performed in the first 48 hours per 100 hip surgeries*, g_5;

3. PRODUCTIVITY:

3.1. *Annual occupancy rate*, g_6;

3.2. *Average waiting time before surgery*, g_7;

4. ECONOMIC-FINANCIAL:

4.1. *Operating costs per standard patient*, g_8.

Summarily, *Number of non-urgent first medical appointments performed in adequate time per 100 first medical appointments* deals with the quantity of first medical appointments that occur in tolerable time compared to the total number of first medical appointments, *Number of outpatient surgeries per 100 potential outpatient procedures* concerns the amount of outpatient surgeries within the total of outpatient procedures, *Number of readmissions in 30 days after discharge per 100 inpatients* attends the ratio between patients readmitted within 30 days after discharge and the total number of inpatient episodes, *Number of long-stay inpatients per 100 admissions* handles the fraction between inpatient admissions longer than 30 days and the total number of inpatient episodes, *Number of hip surgeries performed in the first 48 hours per 100 hip surgeries* assesses the percentage of elderly patients with hip surgeries within 48 hours after fracture compared to the global set of elderly patients with hip surgeries after fracture, *Annual occupancy rate* associates the number of acute admissions and the number of acute inpatient beds over time, *Average waiting time before surgery* relates the number of days until a surgical episode occurs with the total number of scheduled surgical episodes, and *Operating costs per standard patient* is rather self-explanatory.

5.4.2 Criteria

To support the decision-making problem of this case study and allow the comparative assessment between the DMUs itemised in Subsection 5.3.2, a family of five criteria and respective subcriteria, operationalised by the aforementioned indicators (Ehrgott, Naujoks, Stewart, & Wallenius, 2010), was built in line with the work of Ferreira & Marques (2017) and is specified in Table 5.1.

Table 5.1: Criteria, subcriteria, and corresponding indicators.

Criteria	Subcriteria	Indicators
Timeliness of services	Timeliness of medical appointments	*Number of non-urgent first medical appointments performed in adequate time per 100 first medical appointments*, g_1
	Timeliness of surgeries	*Number of hip surgeries performed in the first 48 hours per 100 hip surgeries*, g_5
	Waiting time before surgery	*Average waiting time before surgery*, g_7
Service availability	Occupancy	*Annual occupancy rate*, g_6
Care appropriateness	Outpatient surgeries adequacy	*Number of outpatient surgeries per 100 potential outpatient procedures*, g_2
	Readmissions	*Number of readmissions in 30 days after discharge per 100 inpatients*, g_3
	Large delay of care	*Number of long-stay inpatients per 100 admissions*, g_4
Economic-financial	Technical efficiency	*Operating costs per standard patient*, g_8

As almost every concept in healthcare, quality is complex and non-consensual. However, Donabedian (1988, 2005) categorised it into three interrelated categories: structural quality, process quality, and outcomes. In spite of the usual criticism (Davies & Crombie, 1995; Mant & Hicks, 1995), these definitions have been regularly applied (Ludwig, Van Merode, & Groot, 2010) and, in consonance with

Ferreira, Marques, Nunes, & Figueira (2017), it is safe to affirm that CARE APPROPRIATENESS is a quality-related criterion whose subcriteria correspond to a specification of process quality. Moreover, access is also a quite bewildering concept.

Nevertheless, Gulliford *et al.* (2002) identified four dimensions of access: service availability, personal barriers, financial barriers, and organisational barriers. Ergo, following the previous reasoning, TIMELINESS OF SERVICES and SERVICES AVAILABILITY criteria and associated subcriteria answer to the service availability and organisation barriers classes, respectively.

All of the aforementioned criteria and subcriteria are considered to be outputs of traditional DEA models, but, since the model developed in this dissertation is not conventional, the ECONOMIC-FINANCIAL criterion (typically an input) is presented and computed alongside his fellow criteria.

5.4.3 Performance table and basic statistics

On account of both the DMUs and criteria having been declared in the former subsections, the performances of those DMUs according to the envisaged indicators grouped by criterion are presented in Table 5.2.

Table 5.2: DMU performance by indicator.

		TIMELINESS OF SERVICES			SERVICE AVAILABILITY	PERFORMANCE ASSISTANCE			ECONOMIC-FINANCIAL
		g_1	g_5	g_7	g_6	g_2	g_3	g_4	g_8
GROUP B	a_1	95.13	90.98	0.55	82.97	65.39	6.78	1.17	3172
	a_2	74.43	34.21	0.74	84.57	80.84	7.29	2.84	3291
GROUP C	a_3	82.96	34.12	0.69	77.39	76.16	9.69	1.72	2753
	a_4	82.15	62.00	0.87	79.96	62.48	9.39	2.99	3616
	a_5	71.69	46.30	0.93	88.69	73.88	8.02	4.95	3369
	a_6	71.42	67.92	1.02	84.33	84.93	7.90	3.37	3635
	a_7	63.76	53.97	0.64	112.76	76.80	8.49	3.61	2859
	a_8	58.80	65.27	0.49	84.66	81.76	7.76	2.78	3067
	a_9	57.54	36.65	1.19	100.00	92.77	11.70	3.44	3223
	a_{10}	53.68	29.10	0.65	90.56	79.58	9.45	2.79	3447
	a_{11}	50.84	49.28	0.88	93.35	82.55	6.74	2.83	2621
	a_{12}	50.75	38.93	0.56	83.84	79.93	6.83	2.14	2824
GROUP D	a_{13}	83.66	42.77	1.01	83.10	87.79	7.51	3.84	2738
	a_{14}	66.93	19.19	0.49	75.72	73.02	6.28	2.33	3110
	a_{15}	65.86	80.27	0.98	87.95	86.74	10.88	3.10	2972
	a_{16}	65.75	18.15	1.54	87.86	79.03	7.77	5.65	3418
	a_{17}	64.70	34.16	1.46	87.98	91.06	8.66	3.75	2707
	a_{18}	62.12	62.50	0.84	88.19	79.02	8.47	3.67	2832
	a_{19}	54.42	25.07	0.72	93.78	75.11	7.81	5.14	2868
GROUP E	a_{20}	95.92	22.58	1.27	90.78	85.11	7.60	5.20	3125
	a_{21}	82.00	44.44	1.63	78.90	80.52	7.85	4.88	3075
	a_{22}	76.40	63.10	1.02	87.99	72.82	5.81	3.38	2638
	a_{23}	71.67	67.38	0.67	91.24	76.44	6.94	3.67	2880
	a_{24}	70.04	45.80	0.92	87.26	76.31	10.60	4.55	3009
	a_{25}	67.74	43.60	1.26	75.85	77.36	9.32	4.00	2861

Furthermore, main basic statistics on the indicators are presented in Table 5.3. At first glance, it is possible to detect that the coefficient of variation of g_1, g_2, g_3, g_6, and g_8 is below the reference literature level (25%), which means that those indicators do not demonstrate significant heterogeneity; however, such data homogeneity is understandable, since the physical meaning behind the preceding indicators usually revolves around neighbouring values. Additionally, the interquartile range indicates that, apart from g_5 and g_7, the sample distribution variability of each indicator is not critical, so their values are focused surrounding a certain performance. Besides, one should not dismiss the fact that these indicators were the result of a statistical correlation test (see Subsection 5.3.1), so there is no data overlay.

Table 5.3: Indicators' basic statistics.

Indicator	$\overline{x}$[1]	σ[2]	CV[3]	MIN[4]	MAX[5]	Q1[6]	Q2[7]	Q3[8]	IQR[9]
g_1	69.61	12.45	17.88	50.75	95.92	62.12	67.74	76.40	14.28
g_2	79.10	7.00	8.84	62.48	92.77	76.16	79.03	82.55	6.40
g_3	8.22	1.46	17.81	5.81	11.70	7.29	7.85	9.32	2.03
g_4	3.51	1.12	31.85	1.17	5.65	2.83	3.44	4.00	1.18
g_5	47.11	19.05	40.45	18.15	90.98	34.16	44.44	62.50	28.34
g_6	87.15	7.75	8.89	75.72	111.76	83.10	87.86	90.56	7.46
g_7	0.92	0.32	35.17	0.49	1.63	0.67	0.88	1.02	0.35
g_8	3044.40	293.49	9.64	2621	3635	2832	3009	3223	391

[1] Arithmetic average.
[2] Standard deviation.
[3] Coefficient of variation (computed by dividing the standard deviation by the arithmetic average and multiplying by 100).
[4] Minimum.
[5] Maximum.
[6] First quartile.
[7] Second quartile
[8] Third quartile.
[9] Interquartile range.

5.5 Results and discussion

Solving the PRICDEA model using appropriate software and discussing the results yielded by the model in an array of settings is the purpose of this section.

5.5.1 Model solving

PRICDEA was implemented using the high performance numerical calculation software MATLAB version 2018a and the optimisation package IBM ILOG CPLEX Optimisation Studio version 12.8. Nevertheless, in Subsection 4.2.1, a difference was established between the call for the usage of both the optimal Möbius coefficients obtained in PRICDEA's Phase 1 and the decision-maker's preference-information-incorporated Möbius coefficients in order to legitimise the benchmarks and slacks of PRICDEA's Phase 2. In fact, the latter used MATLAB to run the algorithm and summoned CPLEX to solve the linear program problem (the procedure employed for doing so is delineated in Appendix A), but the former, due to its complexity, was unable to use the aforestated CPLEX version and purely run on MATLAB. This reasoning is made crystal clear in Table 5.4.

Table 5.4: PRICDEA model solving by software.

PRICDEA model feature		MATLAB	CPLEX
Optimal Möbius coefficients	Phase 1	X	
	Phase 2	X	
Analogous Möbius coefficients	Phase 1	X	
	Phase 2		X

5.5.2 Findings

The PRICDEA model methodology and related procedures were applied to produce a comparative performance assessment of secondary healthcare providers. On account of the decision maker's preference information incorporation, some circumstances that arose from the questioning procedures were effortful and produced valuable judgments in contemplation of considering different perspectives of the analysis. Hence, three different scenarios were envisaged not only to take into consideration all of the decision-maker's contributions, but also to assess the robustness of the model using different approaches, *viz.*:

- SCENARIO 1 emphasizes indicator g_8, because operating costs are the cornerstone of the SNS sustainability;
- SCENARIO 2 appraises the politically correct view of the SNS, thus making g_8 the least important indicator;
- SCENARIO 3 disregards operating costs completely, as if g_8 (and, consequently, g_{38}) had no influence in performance assessment of secondary healthcare providers.

These scenarios are displayed in the chart of Figure 5.1 in a symbolical and non-scientifical in the interest of graphically illustrating the relative positioning of each one regarding the importance and inclusion of indicator g_8 in the analysis, as explained above.

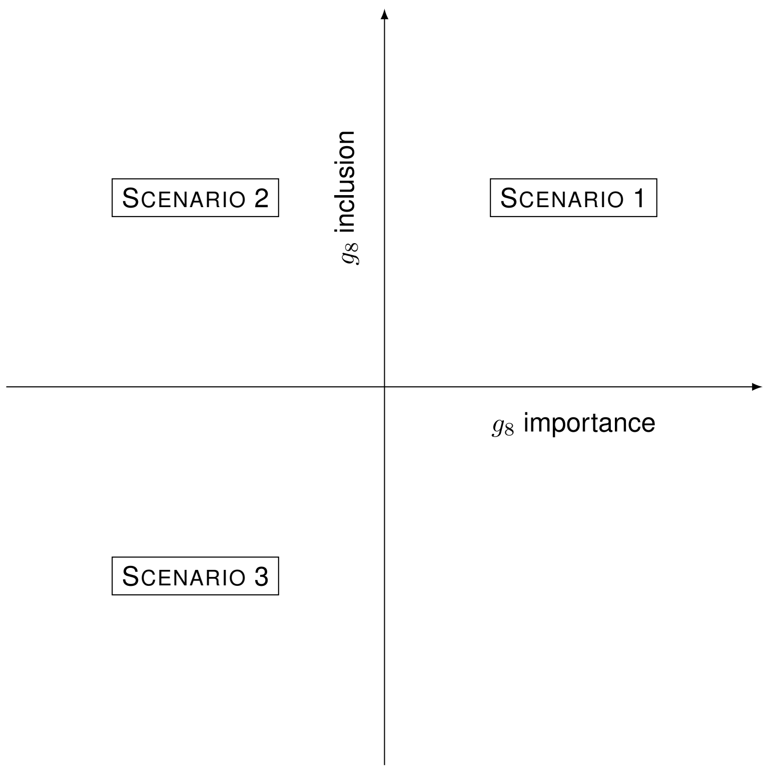

Figure 5.1: Case study scenarios.

It should be evident that the three scenarios use the same utilities (even though SCENARIO 3 does not use the utilities inherent to indicator g_8), which were computed from the data organised on the created database. Ergo, in spite of the methodological order exhibited in Section 4.2, the utilities' calculations will be addressed immediately, for the reason of being common among all scenarios. Since the Möbius coefficients and the comparative operationalisation of PRICDEA depend on the decision-maker's preferences for each scenario, all that will be discussed in their respective paragraphs.

The construction of the utilities' interval scales followed the procedure described in Subsection 4.2.2.3. Creating a discrete levelled scale to accommodate the original continuous one was helpful for the decision-maker to provide more accurate information regarding his preferences on the indicators. Nonetheless, according to the expert, some of the indicators possessed values that were too close to one another and too far from some of the created discrete levels for a firm decision to be made. Henceforth, a feature scaling method was used to standardise the range of data of indicators g_3, g_4, g_6, g_7, and g_8. Besides, g_6 is not a typical indicator, in the sense that it should be minimised or maximised - in fact, *Annual occupancy rate* has a reference level between 80% and 90% according to the initial ACSS database. Therefore, the decision-maker determined that 88% was the optimal value for this indicator and, as a deduction, all values above it were mirrored below it before being normalised. Thereby, the particular case for the computation of g_6's utilities is presented in Table 5.5 and, in a way, serves as an example for the course of action taken in the remaining indicators.

Table 5.5: Utility conversion procedure for indicator g_6.

Institutions	Modified performance[1]	Scaled performance[2]	Blank cards	Number of units between levels	Unit valuation	Interval scale	Linear interpolation	Utilities
a_7	64.24%	0.00%	0	8	0.1250	0.0000		0.0000
1 of 4[3]	70.18%	25.00%	0			0.1250		0.1250
a_{14}	75.72%	48.31%					0.2416	0.2416
a_{25}	75.85%	48.89%					0.2444	0.2444
a_9	76.00%	49.49%					0.2474	0.2474
2 of 4[4]	76.12%	50.00%	1			0.2500		0.2500
a_3	77.39%	55.37%					0.3037	0.3037
a_{21}	78.90%	61.72%					0.3672	0.3672
a_4	79.96%	66.19%					0.4119	0.4119
3 of 4[5]	82.06%	75.00%	3			0.5000		0.5000
a_{19}	82.22%	75.68%					0.5136	0.5136
a_{11}	82.65%	77.47%					0.5495	0.5495
a_1	82.97%	78.83%					0.5767	0.5767
a_{13}	83.10%	79.39%					0.5878	0.5878
a_{12}	83.84%	82.51%					0.6503	0.6503
a_6	84.33%	84.54%					0.6908	0.6908
a_2	84.57%	85.55%					0.7110	0.7110
a_8	84.66%	85.93%					0.7185	0.7185
a_{23}	84.76%	86.37%					0.7274	0.7274
a_{20}	85.22%	88.29%					0.7657	0.7657
a_{10}	85.44%	89.23%					0.7846	0.7846
a_{24}	87.26%	96.91%					0.9381	0.9381
a_5	87.31%	97.10%					0.9419	0.9419
a_{18}	87.81%	99.19%					0.9838	0.9838
a_{16}	87.86%	99.41%					0.9881	0.9881
a_{15}	87.95%	99.77%					0.9955	0.9955
a_{17}	87.98%	99.90%					0.9980	0.9980
a_{22}	87.99%	99.98%					0.9995	0.9995
4 of 4[6]	88.00%	100.00%				1.0000		1.0000

[1] Performance values after mirroring.
[2] Performance values after feature scaling.
[3] Level corresponding to 25% of g_6's overall performance.
[4] Level corresponding to 50% of g_6's overall performance.
[5] Level corresponding to 75% of g_6's overall performance.
[6] Level corresponding to 100% of g_6's overall performance.

Table B.1 in Appendix B.1 discloses every single utility value for each DMU per indicator. In the end, the lingering utility values were computed by simply applying the linear interpolation of Equation (4.21) laid out in Subsection 4.2.2.4.

On top of everything, graphical representations of the utilities for g_i, with $i = 1, ..., n$, as a function of the indicators' values are illustrated in Appendix B.2. From Figure B.1 to Figure B.8, it is visible not only the decision-maker's preferences in defining the levels with the blank cards (Figure B.6 is an explicit example), but also the monotonicity of the eight functions: Figure B.1, Figure B.2, Figure B.5, and Figure B.6 are monotonically increasing functions; Figure B.3, Figure B.4, Figure B.7, and Figure B.8 are monotonically decreasing functions. This concept is particularly useful in asserting whether or not an indicator should be maximised or minimised, respectively.

5.5.2.1 SCENARIO 1: the importance of operating costs

In consonance with the previous paragraph and with the phases outlined in Subsection 4.2.2.2, determining the Möbius coefficients was the last step before assessing the performance of the chosen secondary healthcare providers.

Firstly, the decision-maker pointed out three interactions among the elected criteria:

- Mutually-weakening effect between g_1 and g_5, g_{15};
- Mutually-weakening effect between g_2 and g_5, g_{25};
- Mutually-weakening effect between g_3 and g_8, g_{38}.

In this manner, the final amount of Möbius coefficients to be determined was eleven and the Möbius-transformed Choquet integral could be represented as

$$\begin{aligned}C_\mu(a_k) = {} & (m(g_1)u_1(g_1(a_k)) + m(g_2)u_2(g_2(a_k)) + m(g_3)u_3(g_3(a_k)) + m(g_4)u_4(g_4(a_k)) + \\ & + m(g_5)u_5(g_5(a_k)) + m(g_6)u_6(g_6(a_k)) + m(g_7)u_7(g_7(a_k)) + m(g_8)u_8(g_8(a_k))) + \\ & + (m(g_1, g_5)\min\{u_1(g_1(a_k)), u_5(g_5(a_k))\} + m(g_2, g_5)\min\{u_2(g_2(a_k)), u_5(g_5(a_k))\} + \\ & + m(g_3, g_8)\min\{u_3(g_3(a_k)), u_8(g_8(a_k))\})\end{aligned}$$

for a given DMU a_k, with $k = 1, ..., m$.

Then, following the step-by-step procedure enunciated in Subsection 4.2.2.2, the decision-maker ranked the finite set of reference projects $P = p_1, p_2, p_3, p_4, p_5, p_6, p_7, p_8, p_{15}, p_{25}, p_{38}$ and introduced a number of blank cards between them, not to mention revealing the value of the ratio-z. Table 5.6 reveals the decision-maker's choices.

Table 5.6: Ranking of projects and blank cards for SCENARIO 1.

Ranks and blank cards	
R_1	p_8
e_1	3
R_2	p_{15}, p_{25}, p_{38}
e_2	1
R_3	p_1, p_2, p_5
e_3	2
R_4	p_3, p_4, p_6, p_7
Ratio-z	3

With these preferences in mind, both PRICDEA's Phase 1 and DecSpace's SRF software were properly used to compute the desired coefficients. Wherein the former's linear program was already broadly described, the latter required the intermediate computation of the non-normalised $(w(p_k))$ and normalised weights before pursuing the attainment of the modified values $\overline{w}(p_k)$ and, consequently, the Möbius coefficients m_k and the Choquet capacities μ_k. These weights, coefficients, and capacities are unfolded in Table 5.7 and Table 5.8.

Table 5.7: Optimal Möbius coefficients for SCENARIO 1.

Indicators	**Möbius coefficients** m_k
g_1	0.200000000
g_2	0.200000000
g_3	0.100000000
g_4	0.100000000
g_5	0.200000000
g_6	0.100000000
g_7	0.100000000
g_8	0.400000000
g_{15}	-0.133333333
g_{25}	-0.133333333
g_{38}	-0.133333333
Total	1

Table 5.8: Non-normalised, normalised, and modified values, and analogous Möbius coefficients and Choquet capacities for SCENARIO 1.

Indicators	Non-normalised value $w(p_k)$	Normalised value	Modified value $\overline{w}(p_k)$	Möbius coefficients m_k	Choquet capacities μ_k
g_1	1.67	9.11	1.67	0.218015666	0.218015666
g_2	1.67	9.11	1.67	0.218015666	0.218015666
g_3	1	5.45	1	0.130548303	0.130548303
g_4	1	5.45	1	0.130548303	0.130548303
g_5	1.67	9.11	1.67	0.218015666	0.218015666
g_6	1	5.45	1	0.130548303	0.130548303
g_7	1	5.45	1	0.130548303	0.130548303
g_8	3	16.36	3	0.391644909	0.391644909
g_{15}	2.11	11.5	-1.23	-0.160574413	0.275456919
g_{25}	2.11	11.5	-1.23	-0.160574413	0.275456919
g_{38}	2.11	11.51	-1.89	-0.246736292	0.275456919
Total	-	100	7.66	1	1

Last, but not least, the 25×11 utility matrix and the 11×1 Möbius coefficients' vectors were inputted one at a time in the PRICDEA model's Phase 2 and a 25×25 matrix with the efficient frontier and the benchmark values for the inefficient DMUs, and a 25×11 matrix with the slack values for the inneficient DMUs were outputted. Subsequently, one can declare that inputting the two different Möbius coefficients' vectors resulted in the same efficient frontier, comprised of a_1, a_3, a_6, a_8, a_9, a_{11}, a_{12}, a_{13}, a_{14}, a_{15}, a_{17}, a_{18}, a_{20}, a_{22}, and a_{23}, being a_2, a_4, a_5, a_7, a_{10}, a_{16}, a_{19}, a_{21}, a_{24}, and a_{25} the inefficients DMUs. The relationship between the cited benchmarks and the DMUs below the frontier is represented in Table C.1 and Table C.2 (see Appendix C.1 in Appendix C). These results demonstrate that, regardless of the origin of the used Möbius coefficients', the inefficient secondary healthcare providers aimed for a reference in the same manifold of efficient DMUs (a_1, a_3, a_8, a_{12}, a_{13}, a_{15}, a_{17}, a_{18}, a_{22}, and a_{23}), using between three and six of them in the two situations. For instance, a_5 has different benchmarks depending on the usage of optimal or analogous Möbius coefficients, as depicted in the excerpts Table 5.9 and Table 5.10.

Table 5.9: Excerpt of Table C.1 for a_5.

	GROUP B	GROUP C				GROUP D			GROUP E	
	a_1	a_3	a_8	a_{12}	a_{13}	a_{15}	a_{17}	a_{18}	a_{22}	a_{23}
(...)	(...)	(...)	(...)	(...)	(...)	(...)	(...)	(...)	(...)	(...)
a_5	0.11969164	0	0	0	0	0	0.1914839	0.42720705	0.26161742	0
(...)	(...)	(...)	(...)	(...)	(...)	(...)	(...)	(...)	(...)	(...)

Table 5.10: Excerpt of Table C.2 for a_5.

	GROUP B	GROUP C				GROUP D			GROUP E	
	a_1	a_3	a_8	a_{12}	a_{13}	a_{15}	a_{17}	a_{18}	a_{22}	a_{23}
(...)	(...)	(...)	(...)	(...)	(...)	(...)	(...)	(...)	(...)	(...)
a_5	0.12354058	0	0	0	0	0.32435097	0.12320636	0.24744125	0.18146084	0
(...)	(...)	(...)	(...)	(...)	(...)	(...)	(...)	(...)	(...)	(...)

On the contrary, each member of the set of benchmark DMUs was referred to at least once at up to seven or ten times, depending on the aforementioned vectors, respectively. For example, Table 5.11 and Table 5.12 illustrate this aspect regarding a_{22}.

Table 5.11: Excerpt of Table C.1 for a_{22}.

		(...)	a_{22}	(...)
GROUP B	a_2	(...)	0.03413266	(...)
GROUP C	a_4	(...)	0	(...)
	a_5	(...)	0.26161742	(...)
	a_7	(...)	0	(...)
	a_{10}	(...)	0	(...)
GROUP D	a_{16}	(...)	0.25095884	(...)
	a_{19}	(...)	0.20971234	(...)
GROUP E	a_{21}	(...)	0.16287653	(...)
	a_{24}	(...)	0.13032539	(...)
	a_{25}	(...)	0.00516484	(...)

Table 5.12: Excerpt of Table C.2 for a_{22}.

		(...)	a_{22}	(...)
GROUP B	a_2	(...)	0.03413266	(...)
GROUP C	a_4	(...)	0	(...)
	a_5	(...)	0.18146084	(...)
	a_7	(...)	0	(...)
	a_{10}	(...)	0	(...)
GROUP D	a_{16}	(...)	0.23422989	(...)
	a_{19}	(...)	0.14697602	(...)
GROUP E	a_{21}	(...)	0	(...)
	a_{24}	(...)	0	(...)
	a_{25}	(...)	0.00297375	(...)

In what concerns slacks, Table C.7 and Table C.8 in Appendix C.2 array the values that the inefficient DMUs should improve in relation to the computed indicators' utilities. In the case of the optimal Möbius coefficients, on one hand, a_2 is the DMU with the least number of slacks to enhance and, on the opposite side, a_4, a_5, a_{10} and a_{16} appear to be the ones with the most to improve (see Table 5.13 and then Table 5.14).

Table 5.13: Excerpt of Table C.7 for a_{16}.

Institutions	$u_1(g_1(a_k))$	$u_2(g_2(a_k))$	$u_3(g_3(a_k))$	$u_4(g_4(a_k))$	$u_5(g_5(a_k))$	$u_6(g_6(a_k))$	$u_7(g_7(a_k))$	$u_8(g_8(a_k))$	$\min\{u_1(g_1(a_k)), u_5(g_5(a_k))\}$	$\min\{u_2(g_2(a_k)), u_5(g_5(a_k))\}$	$\min\{u_3(g_3(a_k)), u_8(g_8(a_k))\}$
(...)	(...)	(...)	(...)	(...)	(...)	(...)	(...)	(...)	(...)	(...)	(...)
a_{16}	0.0273	0.1310	0.0000	0.2218	0.1893	0.0000	0.1096	0.7518	0.1893	0.1893	0.2198
(...)	(...)	(...)	(...)	(...)	(...)	(...)	(...)	(...)	(...)	(...)	(...)

Table 5.14: Excerpt of Table C.8 for a_{16}.

Institutions	$u_1(g_1(a_k))$	$u_2(g_2(a_k))$	$u_3(g_3(a_k))$	$u_4(g_4(a_k))$	$u_5(g_5(a_k))$	$u_6(g_6(a_k))$	$u_7(g_7(a_k))$	$u_8(g_8(a_k))$	$\min\{u_1(g_1(a_k)), u_5(g_5(a_k))\}$	$\min\{u_2(g_2(a_k)), u_5(g_5(a_k))\}$	$\min\{u_3(g_3(a_k)), u_8(g_8(a_k))\}$
(...)	(...)	(...)	(...)	(...)	(...)	(...)	(...)	(...)	(...)	(...)	(...)
a_{16}	0.0325	0.1301	0.0000	0.2321	0.2004	0.0000	0.1158	0.7419	0.2004	0.1896	0.2167
(...)	(...)	(...)	(...)	(...)	(...)	(...)	(...)	(...)	(...)	(...)	(...)

On the other hand, it is visible that g_8 and g_4 are, by far, the indicators with the most severe slacks (in opposition to g_2 and g_1), in spite of g_5 and again g_4 being the ones that the majority of the institutions needs to upgrade, particularly when facing the indicators that showed the best results in this field - g_7 and g_6 (contemplate Table 5.15 and forthwith Table 5.16). As for the case of the analogous Möbius coefficients, the results are similar to the optimal ones both in what concerns DMUs and indicators, apart from a_{10}'s withdrawal from the set 'DMUs with the most improvements needed'.

Bottom line, the inference of the optimal Möbius coefficients was proved to be accurate and credible in this scenario, by meeting PRICDEA's results obtained using Möbius coefficients entirely computed

Table 5.15: Excerpt of Table C.7 for g_8.

Groups	**Institutions**	**(...)**	$u_8(g_8(a_k))$	**(...)**
GROUP B	a_2	(...)	0.4328	(...)
GROUP C	a_4	(...)	0.7558	(...)
	a_5	(...)	0.6392	(...)
	a_7	(...)	0.0000	(...)
	a_{10}	(...)	0.7013	(...)
GROUP D	a_{16}	(...)	0.7518	(...)
	a_{19}	(...)	0.0633	(...)
GROUP E	a_{21}	(...)	0.2365	(...)
	a_{24}	(...)	0.1487	(...)
	a_{25}	(...)	0.0012	(...)

Table 5.16: Excerpt of Table C.8 for g_8.

Groups	**Institutions**	**(...)**	$u_8(g_8(a_k))$	**(...)**
GROUP B	a_2	(...)	0.4328	(...)
GROUP C	a_4	(...)	0.7558	(...)
	a_5	(...)	0.5843	(...)
	a_7	(...)	0.0000	(...)
	a_{10}	(...)	0.5359	(...)
GROUP D	a_{16}	(...)	0.7419	(...)
	a_{19}	(...)	0.0000	(...)
GROUP E	a_{21}	(...)	0.1498	(...)
	a_{24}	(...)	0.0120	(...)
	a_{25}	(...)	0.0000	(...)

incorporating the decision-maker's preference information. The yielding of similar conclusions in terms of the efficient frontier, the benchmarks, and the slacks validate this point.

5.5.2.2 SCENARIO 2: the role of ethics

Similarly to SCENARIO 1, the decision-maker elected the same criteria interactions and ranked the members of the criteria family in an invariable manner, apart from the aforesaid g_8 indicator, which is now on the bottom of the ranking, but kept his judgment on appraising the ratio-z (see Table 5.17).

Table 5.17: Ranking of projects and blank cards for SCENARIO 2.

Ranks and blank cards	
R_1	p_{15}, p_{25}, p_{38}
e_1	1
R_2	p_1, p_2, p_5
e_2	2
R_3	p_3, p_4, p_6, p_7
e_3	3
R_4	p_8
Ratio-z	3

This means that the non-normalised and normalised weights would change, as well as the Möbius coefficients, although not significantly in comparison to the first scenario. Table 5.18 and Table 5.19 display all these values.

Table 5.18: Optimal Möbius coefficients for SCENARIO 2.

Indicators	Möbius coefficients m_k
g_1	0.300000000
g_2	0.300000000
g_3	0.200000000
g_4	0.200000000
g_5	0.300000000
g_6	0.200000000
g_7	0.200000000
g_8	0.100000000
g_{15}	-0.266666667
g_{25}	-0.266666667
g_{38}	-0.266666667
Total	1

Table 5.19: Non-normalised, normalised, and modified values, and optimal Möbius coefficients and Choquet capacities for SCENARIO 2.

Indicators	Non-normalised value $w(p_k)$	Normalised value	Modified value $\overline{w}(p_k)$	Möbius coefficients m_k	Choquet capacities μ_k
g_1	2.56	10.14	2.56	0.211395541	0.211395541
g_2	2.56	10.14	2.56	0.211395541	0.211395541
g_3	1.89	7.49	1.89	0.156069364	0.156069364
g_4	1.89	7.49	1.89	0.156069364	0.156069364
g_5	2.56	10.14	2.56	0.211395541	0.211395541
g_6	1.89	7.49	1.89	0.156069364	0.156069364
g_7	1.89	7.49	1.89	0.156069364	0.156069364
g_8	1	3.96	1	0.082576383	0.082576383
g_{15}	3	11.88	-2.12	-0.175061932	0.247729149
g_{25}	3	11.89	-2.12	-0.175061932	0.247729149
g_{38}	3	11.89	0.11	0.009083402	0.247729149
Total	-	100	12.11	1	1

Properly inputting these values in the PRICDEA algorithm's Phase 2 yielded again a 25×25 benchmarks' matrix and a 25×11 slacks' matrix. Once more, for both optimal and analogous Möbius coefficients' vectors, the efficient frontier is constituted by a_1, a_3, a_6, a_8, a_9, a_{11}, a_{12}, a_{13}, a_{14}, a_{15}, a_{17}, a_{18}, a_{20}, a_{22}, and a_{23}, being a_2, a_4, a_5, a_7, a_{10}, a_{16}, a_{19}, a_{21}, a_{24}, and a_{25} the inefficients DMUs. Additionally, the benchmarks for the inefficient DMUs are identical to the previous scenario, as Table C.3 and Table C.4 in Appendix C.1 demonstrate. Nonetheless, side by side, a_1, a_8, a_{12}, a_{13}, a_{17}, a_{22}, and a_{23} remain the same, except for the inclusion of a_{20} and the exclusion of a_3, a_{15}, and a_{18}, and each inefficient DMU used between one and six benchmarks in the two situations. a_{24}'s case is exemplified below in Table 5.20 and Table 5.21 to make a case for the benchmark differences between the using the optimal or the analogous

Möbius coefficients.

Table 5.20: Excerpt of Table C.3 for a_{24}.

	GROUP B	GROUP C				GROUP D				GROUP E	
	a_1	a_3	a_8	a_{12}	a_{13}	a_{15}	a_{17}	a_{18}	a_{20}	a_{22}	a_{23}
(...)	(...)	(...)	(...)	(...)	(...)	(...)	(...)	(...)	(...)	(...)	(...)
a_{24}	0.13705826	0	0	0	0	0.86294174	0	0	0	0	0
(...)	(...)	(...)	(...)	(...)	(...)	(...)	(...)	(...)	(...)	(...)	(...)

Table 5.21: Excerpt of Table C.4 for a_{24}.

	GROUP B	GROUP C		GROUP D			GROUP E	
	a_1	a_8	a_{12}	a_{13}	a_{17}	a_{20}	a_{22}	a_{23}
(...)	(...)	(...)	(...)	(...)	(...)	(...)	(...)	(...)
a_{24}	0	0.14390512	0.0597048	0	0.0758291	0	0.72056098	0
(...)	(...)	(...)	(...)	(...)	(...)	(...)	(...)	(...)

Come again, each member of the DMUs' benchmark set was referred to at least once and up to ten times (for the optimal coefficients) or nine times (for the analogous coefficients). In this scenario, there is the case of DMU a_4 uniquely benchmarked by one secondary healthcare unit to be reported. Additionally, Table 5.22 and Table 5.23 portray the situation of a_{12} benchmarking status variation depending on the employed coefficients.

Table 5.22: Excerpt of Table C.3 for a_{12}.

		(...)	a_{12}	(...)
GROUP B	a_2	(...)	0.28634371	(...)
GROUP C	a_4	(...)	0	(...)
	a_5	(...)	0	(...)
	a_7	(...)	0.28955189	(...)
	a_{10}	(...)	0	(...)
GROUP D	a_{16}	(...)	0	(...)
	a_{19}	(...)	0.42503723	(...)
GROUP E	a_{21}	(...)	0	(...)
	a_{24}	(...)	0	(...)
	a_{25}	(...)	0.01005647	(...)

Table 5.23: Excerpt of Table C.4 for a_{12}.

		(...)	a_{12}	(...)
GROUP B	a_2	(...)	0.28634371	(...)
GROUP C	a_4	(...)	0	(...)
	a_5	(...)	0.10988544	(...)
	a_7	(...)	0.47610969	(...)
	a_{10}	(...)	0.17616989	(...)
GROUP D	a_{16}	(...)	0	(...)
	a_{19}	(...)	0.4457919	(...)
GROUP E	a_{21}	(...)	0	(...)
	a_{24}	(...)	0.0597048	(...)
	a_{25}	(...)	0.56246073	(...)

Lastly, the slack values for the optimal and analogous Möbius coefficients are displayed in Table C.9

and Table C.10 (Appendix C.2), respectively. In the first set of circumstances, a_2 is, yet again, the DMU with the lowest number of slack values to improve, whereas a_4 needs a boost in every single indicator, despite a_{25} and a_{21} also showing poor results (study Table 5.24 and Table 5.25).

Table 5.24: Excerpt of Table C.9 for a_{21}.

Institutions	$u_1(g_1(a_k))$	$u_2(g_2(a_k))$	$u_3(g_3(a_k))$	$u_4(g_4(a_k))$	$u_5(g_5(a_k))$	$u_6(g_6(a_k))$	$u_7(g_7(a_k))$	$u_8(g_8(a_k))$	$\min\{u_1(g_1(a_k)), u_5(g_5(a_k))\}$	$\min\{u_2(g_2(a_k)), u_5(g_5(a_k))\}$	$\min\{u_3(g_3(a_k)), u_8(g_8(a_k))\}$
(...)	(...)	(...)	(...)	(...)	(...)	(...)	(...)	(...)	(...)	(...)	(...)
a_{21}	0.0137	0.1393	0.0743	0.3309	0.4179	0.0000	0.0058	0.0064	0.2011	0.3571	0.0558
(...)	(...)	(...)	(...)	(...)	(...)	(...)	(...)	(...)	(...)	(...)	(...)

Table 5.25: Excerpt of Table C.10 for a_{21}.

Institutions	$u_1(g_1(a_k))$	$u_2(g_2(a_k))$	$u_3(g_3(a_k))$	$u_4(g_4(a_k))$	$u_5(g_5(a_k))$	$u_6(g_6(a_k))$	$u_7(g_7(a_k))$	$u_8(g_8(a_k))$	$\min\{u_1(g_1(a_k)), u_5(g_5(a_k))\}$	$\min\{u_2(g_2(a_k)), u_5(g_5(a_k))\}$	$\min\{u_3(g_3(a_k)), u_8(g_8(a_k))\}$
(...)	(...)	(...)	(...)	(...)	(...)	(...)	(...)	(...)	(...)	(...)	(...)
a_{21}	0.0219	0.0000	0.7856	0.2073	0.1530	0.0000	0.0000	0.2571	0.1321	0.1530	0.7776
(...)	(...)	(...)	(...)	(...)	(...)	(...)	(...)	(...)	(...)	(...)	(...)

At the level of indicators, g_2 and g_1 are once more the ones with the lowest values, inasmuch as g_4, g_8, and now g_5, which exhibit the most worrying outcomes, even though g_4 and g_5 are the indicators that most inefficient secondary healthcare providers need to raise, especially when compared to the most encouraging ones (g_6, g_2, and g_7). In the second outlook, approximate results were yielded, except from the fact that g_3 and g_{38} are now part of the group of indicators with the most serious slacks, and g_5, g_{15}, and g_{38} are at once the indicators that most DMUs need to revamp. A closer look at Table 5.26 and Table 5.27 demonstrate precisely this for g_5.

Table 5.26: Excerpt of Table C.9 for g_5.

Groups	**Institutions**	**(...)**	$u_5(g_5(a_k))$	**(...)**
GROUP B	a_2	(...)	0.1866	(...)
GROUP C	a_4	(...)	0.3768	(...)
	a_5	(...)	0.3519	(...)
	a_7	(...)	0.0111	(...)
	a_{10}	(...)	0.4657	(...)
GROUP D	a_{16}	(...)	0.2004	(...)
	a_{19}	(...)	0.1820	(...)
GROUP E	a_{21}	(...)	0.2038	(...)
	a_{24}	(...)	0.4179	(...)
	a_{25}	(...)	0.0319	(...)

Table 5.27: Excerpt of Table C.10 for g_5.

Groups	**Institutions**	**(...)**	$u_5(g_5(a_k))$	**(...)**
GROUP B	a_2	(...)	0.1866	(...)
GROUP C	a_4	(...)	0.3768	(...)
	a_5	(...)	0.1653	(...)
	a_7	(...)	0.0297	(...)
	a_{10}	(...)	0.2921	(...)
GROUP D	a_{16}	(...)	0.3189	(...)
	a_{19}	(...)	0.3316	(...)
GROUP E	a_{21}	(...)	0.1953	(...)
	a_{24}	(...)	0.1530	(...)
	a_{25}	(...)	0.1084	(...)

The sum and substance of SCENARIO 2 indicate that the new ranking provided by the decision-maker did not have a significant effect in the DEA analysis' benchmarks and slacks, due to their similarity to the ones of SCENARIO 1. In fact, the efficient frontier was the same and most of the benchmarks and slacks outcomes were in agreement, which might be explained due to the ratio-$z = 3$ and the small amount of blank cards in-between levels in both scenarios, thus minimising the effect on the Möbius coefficients of the rank drop of g_8. However, PRICDEA was anew to infer well-founded optimal Möbius coefficients that performed alike the analogous Möbius coefficients, given that the results using both vectors was identical, although there were some minor exceptions this time.

5.5.2.3 SCENARIO 3: the renunciation on operating costs

SCENARIO 3 has quite a distinct outline when compared to the other two scenarios. Here, given the decision-maker's decision to remove *Operating costs per standard patient* from the analysis, g_8 and g_{38} are the two variables taken out of the equation. Hence, nine Möbius coefficients were required for the calculation of the Möbius-transformed Choquet integral represented as

$$\begin{aligned} C_\mu(a_k) = (& m(g_1)u_1(g_1(a_k)) + m(g_2)u_2(g_2(a_k)) + m(g_3)u_3(g_3(a_k)) + m(g_4)u_4(g_4(a_k)) + \\ & + m(g_5)u_5(g_5(a_k)) + m(g_6)u_6(g_6(a_k)) + m(g_7)u_7(g_7(a_k))) + \\ & + (m(g_1, g_5)\min\{u_1(g_1(a_k)), u_5(g_5(a_k))\} + m(g_2, g_5)\min\{u_2(g_2(a_k)), u_5(g_5(a_k))\} \end{aligned}$$

for a given DMU a_k, with $k = 1, ..., m$. Nevertheless, the expert recycled the other indicators and respective utilities, and ranked the former as presented in Table 5.28. This time, ratio-z was attributed a value of 2.

Table 5.28: Ranking of projects and blank cards for SCENARIO 3.

Ranks and blank cards	
R_1	p_{15}, p_{25}
e_1	1
R_2	p_1, p_2, p_5
e_2	2
R_3	p_3, p_4, p_6, p_7
Ratio-z	2

Despite the non-normalised, normalised, and modified values, the Möbius coefficients, and the Choquet capacities values changed, they did not change significantly, due to the assignment of the same number of blank cards between the ranks and the similar value of ratio-z. Table 5.29 and Table 5.30 have these numbers arranged.

Table 5.29: Optimal Möbius coefficients for SCENARIO 3.

Indicators	**Möbius coefficients** m_k
g_1	0.250000000
g_2	0.250000000
g_3	0.150000000
g_4	0.150000000
g_5	0.250000000
g_6	0.150000000
g_7	0.150000000
g_{15}	-0.175000000
g_{25}	-0.175000000
Total	1

Table 5.30: Non-normalised, normalised, and modified values, and analogous Möbius coefficients and Choquet capacities for SCENARIO 3.

Indicators	**Non-normalised value** $w(p_k)$	**Normalised value**	**Modified value** $\overline{w}(p_k)$	**Möbius coefficients** m_k	**Choquet capacities** μ_k
g_1	1.6	12.5	1.6	0.25	0.25
g_2	1.6	12.5	1.6	0.25	0.25
g_3	1	7.81	1	0.15625	0.15625
g_4	1	7.81	1	0.15625	0.15625
g_5	1.6	12.5	1.6	0.25	0.25
g_6	1	7.81	1	0.15625	0.15625
g_7	1	7.81	1	0.15625	0.15625
g_{15}	2	15.63	-1.2	-0.1875	0.3125
g_{25}	2	15.63	-1.2	-0.1875	0.3125
Total	-	100	6.4	1	1

Inserting the new utilities and Möbius coefficients in the PRICDEA model's Phase 2, it returned a 25×25 matrix of benchmarks and a 25×9 matrix of slacks. Another time, a_1, a_3, a_6, a_8, a_9, a_{11}, a_{12}, a_{13}, a_{14}, a_{15}, a_{17}, a_{18}, a_{20}, a_{22}, and a_{23} comprise the efficient frontier, being a_2, a_4, a_5, a_7, a_{10}, a_{16}, a_{19}, a_{21}, a_{24}, and a_{25} the inefficients DMUs. The benchmarks in this scenario are rather equal to the previous scenarios (a_1, a_8, a_{12}, a_{13}, a_{17}, and a_{22}), apart from the set being larger at this moment (a_{15} was reintegrated) - the inefficient secondary healthcare providers used between one and five of these benchmarks (regard Table 5.31 and Table 5.32 for the example of a_{19}).

Table 5.31: Excerpt of Table C.5 for a_{19}.

	GROUP B	**GROUP C**		**GROUP D**				**GROUP E**
	a_1	a_8	a_{12}	a_{13}	a_{15}	a_{17}	a_{20}	a_{22}
(...)	(...)	(...)	(...)	(...)	(...)	(...)	(...)	(...)
a_{19}	0.63403657	0	0	0.36596343	0	0	0	0
(...)	(...)	(...)	(...)	(...)	(...)	(...)	(...)	(...)

Table 5.32: Excerpt of Table C.6 for a_{19}.

	Group B	**Group C**		**Group D**				**Group E**
	a_1	a_8	a_{12}	a_{13}	a_{15}	a_{17}	a_{20}	a_{22}
(...)	(...)	(...)	(...)	(...)	(...)	(...)	(...)	(...)
a_{19}	0.6197795	0	0	0.27418726	0	0	0.10603325	0
(...)	(...)	(...)	(...)	(...)	(...)	(...)	(...)	(...)

As for the benchmarks themselves, they were referred to at least once and up to ten times whether using the optimal or the analogous Möbius coefficients. Table 5.33 and Table 5.34 manifest this condition below for a_{20}.

Table 5.33: Excerpt of Table C.5 for a_{20}.

		(...)	a_{20}	(...)
Group B	a_2	(...)	0.30438876	(...)
Group C	a_4	(...)	0	(...)
	a_5	(...)	0	(...)
	a_7	(...)	0	(...)
	a_{10}	(...)	0	(...)
Group D	a_{16}	(...)	0	(...)
	a_{19}	(...)	0	(...)
Group E	a_{21}	(...)	0	(...)
	a_{24}	(...)	0	(...)
	a_{25}	(...)	0	(...)

Table 5.34: Excerpt of Table C.6 for a_{20}.

		(...)	a_{20}	(...)
Group B	a_2	(...)	0.30438876	(...)
Group C	a_4	(...)	0	(...)
	a_5	(...)	0	(...)
	a_7	(...)	0	(...)
	a_{10}	(...)	0	(...)
Group D	a_{16}	(...)	0	(...)
	a_{19}	(...)	0.10603325	(...)
Group E	a_{21}	(...)	0.18401362	(...)
	a_{24}	(...)	0	(...)
	a_{25}	(...)	0.33661917	(...)

These results are displayed in Appendix C.1 (Table C.5 and Table C.6).

Table C.11 in Appendix C.2 exposes that the vast majority of the inefficient DMUs need to improve in almost every aspect, in spite a_2 turns up again as the best among the worst, opposed to a_4, a_7, a_{19}, and a_{25}, which need to improve quite significantly in almost every single aspect (scrutinise Table 5.35 and Table 5.36 in advance for a_{19}).

Table 5.35: Excerpt of Table C.11 for a_{19}.

Institutions	$u_1(g_1(a_k))$	$u_2(g_2(a_k))$	$u_3(g_3(a_k))$	$u_4(g_4(a_k))$	$u_5(g_5(a_k))$	$u_6(g_6(a_k))$	$u_7(g_7(a_k))$	$\min\{u_1(g_1(a_k)), u_5(g_5(a_k))\}$	$\min\{u_2(g_2(a_k)), u_5(g_5(a_k))\}$
(...)	(...)	(...)	(...)	(...)	(...)	(...)	(...)	(...)	(...)
a_{19}	0.5084	0.0000	0.1930	0.6635	0.5169	0.0672	0.0294	0.5169	0.2357
(...)	(...)	(...)	(...)	(...)	(...)	(...)	(...)	(...)	(...)

Table 5.36: Excerpt of Table C.12 for a_{19}.

Institutions	$u_1(g_1(a_k))$	$u_2(g_2(a_k))$	$u_3(g_3(a_k))$	$u_4(g_4(a_k))$	$u_5(g_5(a_k))$	$u_6(g_6(a_k))$	$u_7(g_7(a_k))$	$\min\{u_1(g_1(a_k)), u_5(g_5(a_k))\}$	$\min\{u_2(g_2(a_k)), u_5(g_5(a_k))\}$
(...)	(...)	(...)	(...)	(...)	(...)	(...)	(...)	(...)	(...)
a_{19}	0.5279	0.0000	0.1881	0.6358	0.4911	0.0862	0.0000	0.4911	0.2162
(...)	(...)	(...)	(...)	(...)	(...)	(...)	(...)	(...)	(...)

In what concerns indicators, it is clear that g_4, g_3, g_5, and g_{15} are the ones with the highest slack values, while g_2 remains the indicator with the lowest slack values. However, g_2 is the indicator that the least DMUs need to meliorate, in contrast to virtually all of the remaining indicators short of g_6 and g_7. In comparison with Table C.12 (see Appendix C.2), the results with the analogous Möbius coefficients were one and the same, save for minor numerical differences. Such similarities are brought to mind in Table 5.37 and Table 5.38 for the predicament of g_4.

Table 5.37: Excerpt of Table C.11 for g_4.

Groups	**Institutions**	**(...)**	$u_4(g_4(a_k))$	**(...)**
GROUP B	a_2	(...)	0.0246	(...)
GROUP C	a_4	(...)	0.6567	(...)
	a_5	(...)	0.2904	(...)
	a_7	(...)	0.4019	(...)
	a_{10}	(...)	0.0743	(...)
GROUP D	a_{16}	(...)	0.2565	(...)
	a_{19}	(...)	0.6635	(...)
GROUP E	a_{21}	(...)	0.4245	(...)
	a_{24}	(...)	0.2757	(...)
	a_{25}	(...)	0.4490	(...)

Table 5.38: Excerpt of Table C.12 for g_4.

Groups	**Institutions**	**(...)**	$u_4(g_4(a_k))$	**(...)**
GROUP B	a_2	(...)	0.0246	(...)
GROUP C	a_4	(...)	0.6567	(...)
	a_5	(...)	0.2904	(...)
	a_7	(...)	0.4019	(...)
	a_{10}	(...)	0.0743	(...)
GROUP D	a_{16}	(...)	0.2565	(...)
	a_{19}	(...)	0.6358	(...)
GROUP E	a_{21}	(...)	0.3764	(...)
	a_{24}	(...)	0.2757	(...)
	a_{25}	(...)	0.3610	(...)

Finally, PRICDEA proved *de novo* to be able to infer legitimate Möbius coefficients that yield positive results when compared to the Möbius coefficients obtained with the direct guidance of the decision-maker. In SCENARIO 3, a system quite distinct from SCENARIO 1 and SCENARIO 2, where operating costs had no effect, was tested, but the outcomes were similar to those schemes. This might be, once again, due to the low value of ratio-z and the small number of blank cards in-between levels judged by the decision-maker, which has significant impacts on the optimal and analogous Möbius coefficients.

5.6 Summary

In this chapter, treating the data was the first and fundamental step for obtaining sturdy results, aside from incorporating the invaluable perceptions of the decision-maker to acquire both the inferred and the analogous Möbius coefficients. Before solving the case study using PRICDEA, a comprehensive scrutiny of the variables was conducted and several statistics were computed. Finally, discussing the results obtained in the three scenarios yielded a favourable image of PRICDEA's performance not only on inferring the Möbius coefficients, but also on assessing the efficient frontier and the benchmarks and slacks of the inefficient DMUs. The following chapter contains the essential conclusions of this work, discusses its limitations, and proposes forthcoming considerations regarding the future of PRICDEA.

Chapter 6

Conclusions and future remarks

6.1 Achievements

The central objective of this dissertation, manifested in Chapter 1, was to suggest a leading-edge technique focused on the enduring optimisation situation in Portuguese healthcare (Chapter 2) that combined interactive variables and incorporated preference information of a decision-maker with a view to reforming institutional performance using a revamped DEA model. To do so, an extensive literature review was conducted in Chapter 3 in pursuance of a well-founded base of knowledge, so that the creation of the so-called PRICDEA model could be materialised and based on pre-existing methodologies, *viz.*, the 'two-phase method' envisaged by Gouveia *et al.* (2008), the Choquet multi-criteria preference aggregation model developed by Bottero *et al.* (2018), and the theoretical concepts on determination of weights of interacting criteria from a reference set matured by Marichal & Roubens (2000), as detailed in Chapter 4. This model was applied to a case study (Chapter 5), where three different settings were established with the cooperation of a decision-maker to the extent of assessing not only the robustness of the PRICDEA model, but also to yield further results on the importance and flexibility of certain criteria, which can be implemented in the real world scenario of the Portuguese SNS secondary healthcare providers.

Therefore, the application of PRICDEA to the gathered and treated ACSS data sample yielded, first of all, inferred Möbius coefficients quite similar to the strictly-decision-maker-preference-information-computed ones in Phase 1, which attests the validity of the algorithm in returning legitimate and genuine numbers. Thus, the computations behind this complex first linear program, with all their manifold of underlying constraints, proved to be both accurate and sensitive to the inputted information. Secondly, a_1, a_3, a_6, a_8, a_9, a_{11}, a_{12}, a_{13}, a_{14}, a_{15}, a_{17}, a_{18}, a_{20}, a_{22}, and a_{23} were the DMUs that constituted the production frontier common to SCENARIO 1, SCENARIO 2, and SCENARIO 3. Moreover, the remaining results of Phase 2's linear program (benchmarks and slacks) were sound and substantiated lawful indications on who the inefficient secondary healthcare providers should look up to in order to improve their performance. Besides, similar results were obtained using the analogous Möbius coefficients, ergo, once more, PRICDEA's robustness is vindicated. Globally, if, on one hand, these conclusions demonstrate the preponderance of the decision-maker's judgment, for good or for bad, in the analysis, on the other hand, they corroborated PRICDEA as a credible approach not only to fill in the knowledge gap in literature, but also to provide healthcare managers and policy makers authentic data that can support decision-making.

Bottom line, one can declare that the goals set forth at the beginning of this dissertation were entirely reached, since:

- The Portuguese healthcare background was profusely described and related to the case study investigated in this research;
- DEA's features were acknowledged in the interest of this technique being applied in the healthcare setting;
- The Choquet multiple criteria preference aggregation model was successfully employed as an auspicious approach for incorporating the information of decision-makers' preferences and dealing with interactive criteria;
- PRICDEA, a breaking new performance assessment methodology for criteria weights' inference, interactive variable embodiment, and preference information incorporation for decision-making, was created;
- The results of the case study in the healthcare sector were corroborated with outcomes from Möbius coefficients uniquely computed stemming from the decision-maker's judgment for a matter of validity and were triumphantly positive, not only for the model itself, but also for future aiding in health policy-making.

6.2 Recommendations

In line with the preceding section, it is safe to say that 40% of the Portuguese secondary healthcare providers that underwent the PRICDEA methodological testing are inefficient, regardless of the purpose of each analysed scenario. In essence, this means that, despite the current SNS position in the aforementioned world ranking, there is still plenty to be accomplished in the health sector in Portugal across diverse levels.

Particularly, institutions a_2, a_4, a_5, a_7, a_{10}, a_{16}, a_{19}, a_{21}, a_{24}, and a_{25} should panegyrise their respective benchmarks (see Appendix C.1) and mind their own slack values (in Appendix C.2) if they intend to improve their individual performance. Above all, such healthcare providers should seek out for advice in consulting professionals who have the social, economic, and political expertise to boost their organisations. At first, an investment must be made in pursuance of prosperity, but, in the long term, that expenditure will not seem that significant, especially because, in the end, it is all about the people.

6.3 Limitations

The major limitations of the present work can be epitomised in a few remarks.

Firstly, PRICDEA's optimisation inference algorithm relies on profuse constraints which affect not only computational performance, but also potential results, since multiple optima can be obtained and it may be the case that plenty of null outputs are generated. This last point was explained by Marichal & Roubens (2000) by placing responsibility on the completeness and coherence of the information set, and the possible incompatibility between that same information and the assumption of a 2-order fuzzy measure.

Next in order, the decision-maker's resolutions of placing multiple criteria in the same level, attributing a small number of blank cards in-between levels, and valuing the ratio-z as a modest number on all scenarios undoubtedly impacted the results. This is where the explanation of similar results between scenarios lies. Naturally, if different inputs were provided, different outputs would have been generated, but the decision-maker was firm in each decision, because of the extreme importance of all criteria in such a delicate and intricate context.

Additionally, and according to Ozcan (2008), finding the appropriate variables and their measurements is frequently difficult. However, assuming this was not the case, Jacobs *et al.* (2006) make a valuable point in indicating the environmental determinants of performance detailed in Subsection 3.2.2.4. Such constraints were not considered in the case study of this dissertation, but it would be a *rococo* envisioning, as Section 6.4 indicates.

6.4 Future work

As prospective developments, in an academia standpoint, it would not be unreasonable to consolidate PRICDEA by including the aforementioned environmental constraints, or apply it to other depths in healthcare or even other spheres of influence. Nonetheless, the perspective of adapting PRICDEA to a network-type performance assessment emerges as an exceptionally interesting outlook on hospital services, with a view to improving performance as system-by-system bottom-up approach, taking into account each one's internal structure. For now, a paper in an international peer-reviewed scientific journal is being prepared for submission. In an industry point of view, evaluating the eventual real-world execution of the results obtained in this dissertation would be an extraordinary honour, and efforts are being made in that direction.

References

Ali, A. I., Lerme, C. S., & Seiford, L. M. (1995). Components of efficiency evaluation in data envelopment analysis. *European Journal of Operational Research*, *80*(3), 462–473. doi: 10.1016/0377-2217(94)00131-U

Athanassopoulos, A. D., & Podinovski, V. V. (1997). Dominance and potential optimality multiple criteria models decision analysis with imprecise information. *Journal of the Operational Research Society*, *48*, 142–150.

Atkinson, A. B. (2005). *The Atkinson review: final report. Measurement of government output and productivity for the national accounts*. Retrieved from http://eprints.lse.ac.uk/33553/ doi: 10.1787/9789264009011-en

Banker, R. D., Charnes, A., & Cooper, W. W. (1984, sep). Some Models for Estimating Technical and Scale Inefficiencies in Data Envelopment Analysis. *Management Science*, *30*(9), 1078–1092. Retrieved from http://pubsonline.informs.org/doi/abs/10.1287/mnsc.30.9.1078 doi: 10.1287/mnsc.30.9.1078

Barros, P. P., Machado, S. R., & Simões, J. d. A. (2011). Portugal. Health system review. *Health systems in transition*, *13*(4), 1–156. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/22222781

Boerboom, L., & Ferretti, V. (2014, nov). Actor-Network-Theory perspective on a forestry decision support system design. *Scandinavian Journal of Forest Research*, *29*(1), 84–95. Retrieved from http://www.tandfonline.com/doi/abs/10.1080/02827581.2014.946960 doi: 10.1080/02827581.2014.946960

Bottero, M., Ferretti, V., Figueira, J. R., Greco, S., & Roy, B. (2018, apr). On the Choquet multiple criteria preference aggregation model: Theoretical and practical insights from a real-world application. *European Journal of Operational Research*, *37*. Retrieved from http://linkinghub.elsevier.com/retrieve/pii/S0377221718303254 doi: 10.1016/j.ejor.2018.04.022

Bouyssou, D. (1999). Using DEA as a tool for MCDM: Some remarks. *Journal of the Operational Research Society*, *50*(9), 974–978. doi: 10.1057/palgrave.jors.2600800

Charnes, A., Cooper, W. W., Lewin, A. Y., & Seiford, L. M. (1994). *Data Envelopment Analysis: Theory, Methodology, and Applications*. Boston: Kluwer Academic Publishers. Retrieved from http://link.springer.com/10.1007/978-94-011-0637-5 doi: 10.1007/978-94-011-0637-5

Charnes, A., Cooper, W. W., & Rhodes, E. (1978, nov). Measuring the efficiency of decision making units. *European Journal of Operational Research*, *2*(6), 429–444. Retrieved from http://linkinghub.elsevier.com/retrieve/pii/0377221778901388 doi: 10.1016/0377-2217(78)90138-8

Charnes, A., Cooper, W. W., Wei, Q. L., & Huang, Z. M. (1989, jul). Cone ratio data envelopment analysis and multi-objective programming. *International Journal of Systems Science*, *20*(7), 1099–1118. Retrieved from http://www.tandfonline.com/doi/abs/10.1080/00207728908910197 doi: 10.1080/00207728908910197

Charnes, A., Cooper, W. W., Wei, Q. L., & Huang, Z. M. (1990, jul). Fundamental theorems of nondominated solutions associated with cones in normed linear spaces. *Journal of Mathematical Analysis and Applications*, *150*(1), 54–78. Retrieved from http://linkinghub.elsevier.com/retrieve/pii/0022247X9090195L doi: 10.1016/0022-247X(90)90195-L

Chilingerian, J. A. (1994). Exploring Why Some Physicians' Hospital Practices are More Efficient: Taking DEA Inside the Hospital. In *Data envelopment analysis: Theory, methodology, and applications* (pp. 167–193). Boston: Kluwer Academic Publishers. Retrieved from http://link.springer.com/10.1007/978-94-011-0637-5{_}9 doi: 10.1007/978-94-011-0637-5_9

Choquet, G. (1954). Theory of capacities. In *Annales de l'institut fourier* (Vol. 5, pp. 131–295). Retrieved from http://aif.cedram.org/item?id=AIF{_}1986{_}{_}36{_}2{_}167{_}0 doi: 10.5802/aif.896

Chowdhury, H., & Zelenyuk, V. (2016, sep). Performance of hospital services in Ontario: DEA with truncated regression approach. *Omega*, *63*, 111–122. Retrieved from http://linkinghub.elsevier.com/retrieve/pii/S030504831500211X doi: 10.1016/j.omega.2015.10.007

Coase, R. H. (1937, nov). The Nature of the Firm. *Economica*, *4*(16), 386. Retrieved from http://www.jstor.org/stable/2626876?origin=crossref doi: 10.2307/2626876

Coelli, T. (1996). A Guide to DEAP Version 2.1: A Data Envelopment Analysis (Computer) Program. *Working Paper 96/08, Centre for Efficiency and Productivity Analysis*(1994), 1–49.

Coelli, T., Rao, D., & Battese, G. (1998). *An Introduction to Efficiency and Productivity Analysis*. Boston: Kluwer Academic Publishers. Retrieved from http://link.springer.com/10.1007/b136381 doi: 10.1007/b136381

Cooper, W. W., Seiford, L. M., & Tone, K. (2000). *Data Envelopment Analysis: A Comprehensive Text with Models, Applications, References and DEA-Solver Software*. Boston, MA: Springer US. Retrieved from http://link.springer.com/10.1007/978-0-387-45283-8 doi: 10.1007/978-0-387-45283-8

Corrente, S., Figueira, J. R., & Greco, S. (2014, jun). Dealing with interaction between bipolar multiple criteria preferences in PROMETHEE methods. *Annals of Operations Research*, *217*(1), 137–164. Retrieved from http://link.springer.com/10.1007/s10479-014-1554-1 doi: 10.1007/s10479-014-1554-1

Cortes, M. (2016). Breve olhar sobre o estado da saúde em Portugal. *Sociologia, Problemas e Práticas*(80), 117–143. Retrieved from http://journals.openedition.org/spp/2117

Coulter, A., & Magee, H. (2003). *The European Patient of the Future*. Maidenhead, Berkshire: Open University Press.

Cruz, C., & Marques, R. D. R. d. C. (2013). Integrating infrastructure and clinical management in PPP for health care. *Journal of Management in Engineering.*, *29*(4), 471–481.

Daraio, C., & Simar, L. (2007). *Advanced Robust and Nonparametric Methods in Efficiency Analysis* (Vol. 4). Boston, MA: Springer US. Retrieved from http://link.springer.com/10.1007/978-0-387-35231-2 doi: 10.1007/978-0-387-35231-2

Davies, H. T., & Crombie, I. K. (1995, sep). Assessing the quality of care. *British Medical Journal*, *311*(7008), 766. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/7580432 http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=PMC2550785

Dente, B. (2014). *Understanding Policy Decisions* (PoliMI SpringerBriefs, Ed.). Milan, Italy: Springer International Publishing. Retrieved from http://link.springer.com/10.1007/978-3-319-02520-9 doi: 10.1007/978-3-319-02520-9

Donabedian, A. (1988). The quality of care: How can it be assessed? *Journal of the American Medical Association*, *260*(12), 1743–1748. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/3045356

Donabedian, A. (2005, dec). Evaluating the Quality of Medical Care. *Milbank Quarterly*, *83*(4), 691–729. Retrieved from http://doi.wiley.com/10.1111/j.1468-0009.2005.00397.x doi: 10.1111/j.1468-0009.2005.00397.x

Doyle, J., & Green, R. (1993, nov). Data envelopment analysis and multiple criteria decision making. *Omega*, *21*(6), 713–715. Retrieved from http://linkinghub.elsevier.com/retrieve/pii/030504839390013B doi: 10.1016/0305-0483(93)90013-B

Dyson, R. G., & Thanassoulis, E. (1988, jun). Reducing Weight Flexibility in Data Envelopment Analysis. *Journal of the Operational Research Society*, *39*(6), 563–576. Retrieved from https://www.tandfonline.com/doi/full/10.1057/jors.1988.96 doi: 10.1057/jors.1988.96

Ehrgott, M., Naujoks, B., Stewart, T. J., & Wallenius, J. (2010). Multiple Criteria Decision Making for Sustainable Energy and Transportation Systems. In *19th international conference on multiple criteria decision making* (Vol. 634). Berlin, Heidelberg: Springer Berlin Heidelberg. Retrieved from http://link.springer.com/10.1007/978-3-642-04045-0 doi: 10.1007/978-3-642-04045-0

Eken, M. H., & Kale, S. (2014). Bank Branch Efficiency with DEA. In *Handbook of research on strategic performance management and measurement using data envelopment analysis* (pp. 626–667). IGO Global. Retrieved from http://services.igi-global.com/resolvedoi/resolve.aspx?doi=10.4018/978-1-4666-4474-8.ch022 doi: 10.4018/978-1-4666-4474-8.ch022

EuroQol Group. (1990, dec). EuroQol - a new facility for the measurement of health-related quality of life. *Health Policy*, *16*(3), 199–208. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/10109801

Everitt, B. S., Landau, S., Leese, M., & Stahl, D. (2011). *Cluster Analysis* (5th ed.). Chichester, UK: John Wiley & Sons, Ltd. Retrieved from http://doi.wiley.com/10.1002/9780470977811 doi: 10.1002/9780470977811

Ferrando, A., Ivaldi, C., Buttiglieri, A., Pagano, E., Bonetto, C., Arione, R., ... Ciccone, G. (2005, aug). Guidelines for preoperative assessment: impact on clinical practice and costs. *International Journal for Quality in Health Care*, *17*(4), 323–329. Retrieved from http://academic.oup.com/intqhc/article/17/4/323/2886502/Guidelines-for-preoperative-assessment-impact-on doi: 10.1093/intqhc/mzi039

Ferreira, D. C., & Marques, R. D. R. d. C. (2017). *Do quality and access to hospital services impact on their technical efficiency?*

Ferreira, D. C., Marques, R. D. R. d. C., Nunes, A., & Figueira, J. R. (2017, sep). Patients' satisfaction: The medical appointments valence in Portuguese public hospitals. *Omega*, *000*, 1–19. Retrieved from http://linkinghub.elsevier.com/retrieve/pii/S0305048317302694 doi: 10.1016/j.omega.2017.08.009

Ferreira, D. C., Nunes, A., & Marques, R. D. R. d. C. (2018). Economies of scope in the health sector: The case of Portuguese hospitals. *European Journal of Operational Research.*, *266*(2), 716–735.

Ferretti, V. (2016, sep). From stakeholders analysis to cognitive mapping and Multi-Attribute Value Theory: An integrated approach for policy support. *European Journal of Operational Research*, *253*(2), 524–541. Retrieved from http://linkinghub.elsevier.com/retrieve/pii/S0377221716301072 doi: 10.1016/j.ejor.2016.02.054

Fetter, R. B., Shin, Y., Freeman, J. L., Averill, R. F., & Thompson, J. D. (1980, feb). Case mix definition by diagnosis-related groups. *Medical care*, *18*(2 Suppl), iii, 1–53. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/7188781

Figueira, J. R., Greco, S., & Roy, B. (2009, dec). ELECTRE methods with interaction between criteria: An extension of the concordance index. *European Journal of Operational Research*, *199*(2), 478–495. Retrieved from http://linkinghub.elsevier.com/retrieve/pii/S0377221708010138 doi: 10.1016/j.ejor.2008.11.025

Figueira, J. R., & Roy, B. (2002). Determining the weights of criteria in the ELECTRE type methods with a revised Simos' procedure. *European Journal of Operational Research*, *139*(2), 317–326. doi: 10.1016/S0377-2217(01)00370-8

Fisher, R. A. (1925). *Statistical Methods for Research Workers* (No. V). doi: 10.1056/NEJMc061160

Florez, D. C., Aguirre, S., Amaya, C.-A., & Velasco, N. (2008, apr). Optimization of the laundry service in a public hospital in Bogotá D.C., Colombia: A case of vehicle routing with split delivery. In *2008 ieee systems and information engineering design symposium* (pp. 106–111). IEEE. Retrieved from http://ieeexplore.ieee.org/document/4559694/ doi: 10.1109/SIEDS.2008.4559694

Gerdtham, U. G., Rehnberg, C., & Tambour, M. (1999, aug). The impact of internal markets on health care efficiency: evidence from health care reforms in Sweden. *Applied Economics*, *31*(8), 935–945. Retrieved from http://www.tandfonline.com/doi/abs/10.1080/000368499323652 doi: 10.1080/000368499323652

Golany, B. (1988, aug). An Interactive MOLP Procedure for the Extension of DEA to Effectiveness Analysis. *Journal of the Operational Research Society*, *39*(8), 725–734. Retrieved from https://www.tandfonline.com/doi/full/10.1057/jors.1988.127 doi: 10.1057/jors.1988.127

Golany, B., & Roll, Y. A. (1994). Incorporating standards via DEA. In A. Charnes, W. W. Cooper, A. Y. Lewin, & L. M. Seiford (Eds.), *Data envelopment analysis: Theory, methodology, and appli-*

cations. Norwell: Kluwer Academic Publishers.

Goldstein, H., & Spiegelhalter, D. J. (1996). League Tables and Their Limitations: Statistical Issues in Comparisons of Institutional Performance. *Journal of the Royal Statistical Society. Series A (Statistics in Society)*, *159*(3), 385–443. doi: 10.2307/2983325

Gouveia, M. C., Dias, L. C., & Antunes, C. H. (2008). Additive DEA based on MCDA with imprecise information. *Journal of the Operational Research Society*, *59*(1), 54–63. doi: 10.1057/palgrave.jors.2602317

Grabisch, M. (1996a, mar). The application of fuzzy integrals in multicriteria decision making. *European Journal of Operational Research*, *89*(3), 445–456. Retrieved from `http://linkinghub.elsevier.com/retrieve/pii/037722179500176X` doi: 10.1016/0377-2217(95)00176-X

Grabisch, M. (1996b, may). The representation of importance and interaction of features by fuzzy measures. *Pattern Recognition Letters*, *17*(6), 567–575. Retrieved from `http://linkinghub.elsevier.com/retrieve/pii/0167865596000207` doi: 10.1016/0167-8655(96)00020-7

Grabisch, M. (1997, dec). k-order additive discrete fuzzy measures and their representation. *Fuzzy Sets and Systems*, *92*(2), 167–189. Retrieved from `http://linkinghub.elsevier.com/retrieve/pii/S0165011497001681` doi: 10.1016/S0165-0114(97)00168-1

Grabisch, M., & Roubens, M. (1999, nov). An axiomatic approach to the concept of interaction among players in cooperative games. *International Journal of Game Theory*, *28*(4), 547–565. Retrieved from `http://link.springer.com/10.1007/s001820050125` doi: 10.1007/s001820050125

Grosskopf, S., & Valdmanis, V. (1987, jun). Measuring hospital performance: A non-parametric approach. *Journal of Health Economics*, *6*(2), 89–107. Retrieved from `http://linkinghub.elsevier.com/retrieve/pii/0167629687900014` doi: 10.1016/0167-6296(87)90001-4

Gulliford, M., Figueroa-Munoz, J., Morgan, M., Hughes, D., Gibson, B., Beech, R., & Hudson, M. (2002, jul). What does 'access to health care' mean? *Journal of Health Services Research & Policy*, *7*(3), 186–188. Retrieved from `http://journals.sagepub.com/doi/10.1258/135581902760082517` doi: 10.1258/135581902760082517

Guterman, S., & Dobson, A. (1986). Impact of the Medicare prospective payment system for hospitals. *Health Care Financing Review*, *7*(3), 97–114. Retrieved from `http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=4191526{&}tool=pmcentrez{&}rendertype=abstract` doi: hcfr-7-3-97 [pii]

Halme, M., Joro, T., Korhonen, P. J., Salo, S., & Wallenius, J. (1999, jan). A Value Efficiency Approach to Incorporating Preference Information in Data Envelopment Analysis. *Management Science*, *45*(1), 103–115. Retrieved from `http://pubsonline.informs.org/doi/abs/10.1287/mnsc.45.1.103` doi: 10.1287/mnsc.45.1.103

Harris, J. E. (1977). The Internal Organization of Hospitals: Some Economic Implications. *The Bell Journal of Economics*, *8*(2), 467–482. Retrieved from `http://www.jstor.org/stable/3003297` doi: 10.2307/3003297

Helling, D. K., Nelson, K. M., Ramirez, J. E., & Humphries, T. L. (2006). Kaiser Permanente Colorado Region Pharmacy Department: innovative leader in pharmacy practice. *Journal of the American*

Pharmacists Association : JAPhA, *46*(1), 67–76. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/16529342

Hollingsworth, B., & Parkin, D. (2001, mar). The efficiency of the delivery of neonatal care in the UK. *Journal of Public Health Medicine*, *23*(1), 47–50. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/11315693

Hollingsworth, B., & Smith, P. C. (2003, sep). Use of ratios in data envelopment analysis. *Applied Economics Letters*, *10*(11), 733–735. Retrieved from http://www.tandfonline.com/doi/abs/10.1080/1350485032000133381 doi: 10.1080/1350485032000133381

Iezzoni, L. I. (2003). *Risk Adjustment for Measuring Healthcare Outcomes* (3rd ed.). Baltimore: Health Administration Press.

Instituto Nacional de Estatística. (2015). *Despesas de Saúde em Portugal* (Tech. Rep.). Retrieved from https://www.ine.pt/xurl/doc/263919711

Jacobs, R., Smith, P. C., & Street, A. (2006). *Measuring Efficiency in Health Care: Analytic Techniques and Health Policy*. Cambridge: Cambridge University Press. Retrieved from http://ebooks.cambridge.org/ref/id/CBO9780511617492 doi: 10.1017/CBO9780511617492

Ji, A., Liu, H., Qiu, H.-j., & Lin, H. (2015, nov). Data envelopment analysis with interactive variables. *Management Decision*, *53*(10), 2390–2406. Retrieved from http://www.emeraldinsight.com/doi/10.1108/MD-11-2014-0631 doi: 10.1108/MD-11-2014-0631

Joro, T., & Korhonen, P. J. (2015). *Extension of Data Envelopment Analysis with Preference Information* (Vol. 218). Boston, MA: Springer US. Retrieved from http://link.springer.com/10.1007/978-1-4899-7528-7 doi: 10.1007/978-1-4899-7528-7

Joro, T., Korhonen, P. J., & Wallenius, J. (1998, jul). Structural Comparison of Data Envelopment Analysis and Multiple Objective Linear Programming. *Management Science*, *44*(7), 962–970. Retrieved from http://pubsonline.informs.org/doi/abs/10.1287/mnsc.44.7.962 doi: 10.1287/mnsc.44.7.962

Keeney, R. L., Raiffa, H., & Rajala, D. W. (1979). *Decisions with Multiple Objectives: Preferences and Value Trade-Offs* (Vol. 9) (No. 7). New York: Wiley. Retrieved from http://ieeexplore.ieee.org/document/4310245/ doi: 10.1109/TSMC.1979.4310245

Kim, S. Y. (2013). Efficacy versus Effectiveness. *Korean Journal of Family Medicine*, *34*(4), 227. Retrieved from https://www.ncbi.nlm.nih.gov/pmc/articles/PMC3726789/ doi: 10.4082/kjfm.2013.34.4.227

Kooreman, P. (1994, oct). Nursing home care in The Netherlands: a nonparametric efficiency analysis. *Journal of Health Economics*, *13*(3), 301–16; discussion 345–6. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/10138856

Korhonen, P. J., Silvennoinen, K., Wallenius, J., & Öörni, A. (2012, jun). Can a linear value function explain choices? An experimental study. *European Journal of Operational Research*, *219*(2), 360–367. Retrieved from http://linkinghub.elsevier.com/retrieve/pii/S0377221711011349 doi: 10.1016/j.ejor.2011.12.040

Korhonen, P. J., & Wallenius, J. (1989). A Careful Look at Efficiency in Multiple Objective Linear

Programming. *Asia-Pacific Journal of Operational Research*, *6*, 46–62. Retrieved from http://link.springer.com/10.1007/978-3-642-49298-3{_}31 doi: 10.1007/978-3-642-49298-3_31

Lameire, N., Joffe, P., & Wiedemann, M. (1999). Healthcare systems - an international review: an overview. *Nephrology Dialysis Transplantation*, *14*(6), 3–9.

Ludwig, M., Van Merode, F., & Groot, W. (2010, jun). Principal agent relationships and the efficiency of hospitals. *The European Journal of Health Economics*, *11*(3), 291–304. Retrieved from http://link.springer.com/10.1007/s10198-009-0176-z doi: 10.1007/s10198-009-0176-z

Mant, J., & Hicks, N. (1995, sep). Detecting differences in quality of care: the sensitivity of measures of process and outcome in treating acute myocardial infarction. *British Medical Journal*, *311*(7008), 793–796. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/7580444 http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=PMC2550793

Marichal, J.-L., & Roubens, M. (2000, aug). Determination of weights of interacting criteria from a reference set. *European Journal of Operational Research*, *124*(3), 641–650. Retrieved from http://linkinghub.elsevier.com/retrieve/pii/S0377221799001824 doi: 10.1016/S0377-2217(99)00182-4

Meigs, J., & Beilinson, J. (2009, aug). Inventor Dean Kamen Says Healthcare Debate “Backward Looking”. *Popular Mechanics*. Retrieved from http://www.popularmechanics.com/science/health/a5101/4327012/

Mobley, L. R., & Magnussen, J. (2002). The impact of managed care penetration and hospital quality on efficiency in hospital staffing. *Journal of Health Care Finance*, *28*(4), 24–42. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/12148662

Murofushi, T., & Soneda, S. (1993). Techniques for reading fuzzy measures (III): Interaction index. In *Proceedings of the 9th fuzzy systems symposium, sapporo, japan* (pp. 693–696).

National Center for Health Statistics. (2017). *Health, United States, 2016: With Chartbook on Long-term Trends in Health* (Tech. Rep.). Retrieved from https://www.cdc.gov/nchs/data/hus/hus16.pdf{#}056{%}0Ahttps://www.cdc.gov/nchs/data/hus/hus16.pdf{#}listtables{%}0Ahttps://www.cdc.gov/nchs/data/hus/hus16.pdf{#}019

Nunamaker, T. R. (1983). Measuring routine nursing service efficiency: a comparison of cost per patient day and data envelopment analysis models. *Health Services Research*, *18*(2 Pt 1), 183–208. Retrieved from https://www.ncbi.nlm.nih.gov/pmc/articles/PMC1068745/

Othman, F. M., Mohd-Zamil, N. A., Rasid, S. Z. A., Vakilbashi, A., & Mokhber, M. (2016). Data Envelopment Analysis : A Tool of Measuring Efficiency in Banking Sector. *International Journal of Economics and Financial Issues*, *6*(3), 911–916.

Ozcan, Y. A. (2005). *Quantitative Methods in Health Care Management: Techniques and Applications*. Jossey-Bass. doi: 10.1007/978-3-7908-2151-2

Ozcan, Y. A. (2008). *Health Care Benchmarking and Performance Evaluation: An Assessment using Data Envelopment Analysis (DEA)* (Vol. 210). Boston, MA: Springer US. Retrieved from http://link.springer.com/10.1007/978-1-4899-7472-3 doi: 10.1007/978-1-4899-7472-3

Ozcan, Y. A., & Cotter, J. J. (1994, jun). An assessment of efficiency of area agencies on aging in

Virginia through data envelopment analysis. *The Gerontologist*, *34*(3), 363–70. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/8076878

Paddock, S. M. (2014, jun). Statistical Benchmarks for Health Care Provider Performance Assessment: A Comparison of Standard Approaches to a Hierarchical Bayesian Histogram-Based Method. *Health Services Research*, *49*(3), 1056–1073. Retrieved from http://doi.wiley.com/10.1111/1475-6773.12149 doi: 10.1111/1475-6773.12149

Palmer, S., & Torgerson, D. J. (1999). Economics notes: Definitions of efficiency. *British Medical Journal*, *318*(October 2006), 1999. Retrieved from https://www.ncbi.nlm.nih.gov/pmc/articles/PMC1115526/

Parkin, D., & Hollingsworth, B. (1997, nov). Measuring production efficiency of acute hospitals in Scotland, 1991-94: validity issues in data envelopment analysis. *Applied Economics*, *29*(11), 1425–1433. Retrieved from https://www.tandfonline.com/doi/full/10.1080/000368497326255 doi: 10.1080/000368497326255

Pedraja-Chaparro, F., Salinas-Jiménez, J., & Smith, P. C. (1999, jun). On the quality of the data envelopment analysis model. *Journal of the Operational Research Society*, *50*(6), 636–644. Retrieved from http://link.springer.com/10.1057/palgrave.jors.2600741 doi: 10.1057/palgrave.jors.2600741

Puig-Junoy, J. (1998a, apr). Measuring health production performance in the OECD. *Applied Economics Letters*, *5*(4), 255–259. Retrieved from http://www.tandfonline.com/doi/abs/10.1080/135048598354933 doi: 10.1080/135048598354933

Puig-Junoy, J. (1998b, may). Technical efficiency in the clinical management of critically ill patients. *Health economics*, *7*(3), 263–77. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/9639338

Roy, B., & Figueira, J. R. (1998). Détermination des poids des critères dans les méthodes du type ELECTRE avec la technique de Simos revisée. *Université Paris - Dauphine*, *Document d*.

Sakellarides, C., Castelo-Branco, L., Barbosa, P., & Azevedo, H. (2014). The impact of the financial crisis on the health system and health in Portugal. *European Observatory on Health Systems and Policies*, 1–56.

Santana, P. (2002, jul). Poverty, social exclusion and health in Portugal. *Social Science & Medicine*, *55*(1), 33–45. Retrieved from http://linkinghub.elsevier.com/retrieve/pii/S0277953601002180 doi: 10.1016/S0277-9536(01)00218-0

Shanmugam, R. (2014). Data Envelopment Analysis for Operational Efficiency. In *Encyclopedia of business analytics and optimization* (pp. 580–590). IGI Global. Retrieved from http://services.igi-global.com/resolvedoi/resolve.aspx?doi=10.4018/978-1-4666-5202-6.ch053 doi: 10.4018/978-1-4666-5202-6.ch053

Sherman, H. D. (1984, oct). Hospital efficiency measurement and evaluation. Empirical test of a new technique. *Medical care*, *22*(10), 922–38. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/6436590

Sherman, H. D., & Zhu, J. (2006). *Service Productivity Management: Improving Service Performance* *using Data Envelopment Analysis (DEA)*. New York: Springer US. Retrieved from http://link.springer.com/10.1007/0-387-33231-6 doi: 10.1007/0-387-33231-6

Simos, J. (1990a). Evaluer l'impact sur l'environnement: Une approche originale par l'analyse multicritère et la négociation. *Presses Polytechniques et Universitaires Romandes*.

Simos, J. (1990b). *L'évaluation environnementale: Un processus cognitif négocié* (Unpublished doctoral dissertation). EPFL.

Singaroyan, R., Seed, C. A., & Egdell, R. M. (2006, mar). Is a target culture in health care always compatible with efficient use of resources? A cost-effectiveness analysis of an intervention to achieve thrombolysis targets. *Journal of Public Health*, *28*(1), 31–34. Retrieved from http://academic.oup.com/jpubhealth/article/28/1/31/1553680/Is-a-target-culture-in-health-care-always doi: 10.1093/pubmed/fdi072

Steuer, R. A. (1986). *Multiple criteria optimization: theory, computation and application*. New York: Wiley.

Stewart, T. J. (1996). Relationships between data envelopment analysis and multiple criteria decision analysis. *Journal of the Operational Research Society*, *47*, 654–665.

Sugeno, M. (1974). *Theory of fuzzy integrals and its applications* (Unpublished doctoral dissertation). Tokyo Institute of Technology.

Sugeno, M. (1977). Fuzzy measures and fuzzy integrals: A survey. In M. M. Gupta, G. N. Saridis, & B. R. Gaines (Eds.), *Fuzzy automata and decision processes* (pp. 89–102). North-Holland, Amsterdam: Elsevier. Retrieved from http://linkinghub.elsevier.com/retrieve/pii/B9781483214504500274 doi: 10.1016/B978-1-4832-1450-4.50027-4

Thanassoulis, E., & Dyson, R. G. (1992, jan). Estimating preferred target input-output levels using data envelopment analysis. *European Journal of Operational Research*, *56*(1), 80–97. Retrieved from http://linkinghub.elsevier.com/retrieve/pii/037722179290294J doi: 10.1016/0377-2217(92)90294-J

Thompson, R. G., Langemeier, L. N., Lee, C.-T., Lee, E., & Thrall, R. M. (1990, oct). The role of multiplier bounds in efficiency analysis with application to Kansas farming. *Journal of Econometrics*, *46*(1-2), 93–108. Retrieved from http://linkinghub.elsevier.com/retrieve/pii/030440769090049Y doi: 10.1016/0304-4076(90)90049-Y

Thompson, R. G., Singleton, F. D., Thrall, R. M., & Smith, B. A. (1986, dec). Comparative Site Evaluations for Locating a High-Energy Physics Lab in Texas. *Interfaces*, *16*(6), 35–49. Retrieved from http://pubsonline.informs.org/doi/abs/10.1287/inte.16.6.35 doi: 10.1287/inte.16.6.35

Toffler, A. (1991). *Power shift. Knowledge, wealth and violence at the edge of the 21st century*. New York: Bantam Books.

Tone, K. (2004). Malmquist Productivity Index. In *Handbook on data envelopment analysis* (pp. 203–227). Boston: Kluwer Academic Publishers. Retrieved from http://link.springer.com/10.1007/1-4020-7798-X{_}8 doi: 10.1007/1-4020-7798-X_8

Tsionas, E. G. (2003, jun). Combining DEA and stochastic frontier models: An empirical Bayes approach. *European Journal of Operational Research*, *147*(3), 499–510. Retrieved from http://

linkinghub.elsevier.com/retrieve/pii/S0377221702002485 doi: 10.1016/S0377-2217(02)00248-5

Üstün, T., Chatterji, S., Mechbal, A., Murray, C., & WHS Collaborating Groups. (2003). Chapter 58: The World Health Surveys. Geneva: World Health Organization. Retrieved from http://www.who.int/entity/healthinfo/survey/hspa58.pdf

Von Winterfeldt, D., & Edwards, W. (1986). *Decision analysis and behavioral research*. Cambridge: Cambridge University Press.

Ware, J. E., & Sherbourne, C. D. (1992, jun). The MOS 36-item Short Form Health Status Survey (SF-36). *Medical Care*, *30*(6), 473–83. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/1593914

Wierzbicki, A. P. (1986, jun). On the completeness and constructiveness of parametric characterizations to vector optimization problems. *Operations-Research-Spektrum*, *8*(2), 73–87. Retrieved from http://link.springer.com/10.1007/BF01719738 doi: 10.1007/BF01719738

Williamson, O. E. (1973). Markets and Hierarchies: Some Elementary Considerations. *The American Economic Review*, *63*(2), 10. Retrieved from http://www.jstor.org/stable/1817092

Wong, Y.-H. B., & Beasley, J. E. (1990, sep). Restricting Weight Flexibility in Data Envelopment Analysis. *The Journal of the Operational Research Society*, *41*(9), 829. Retrieved from http://www.jstor.org/stable/2583498?origin=crossref doi: 10.2307/2583498

World Health Organisation. (2000). *World Health Report 2000* (Tech. Rep.). Geneva: WHO.

Xia, M., & Chen, J.-X. (2017, dec). Data Envelopment Analysis Based on Choquet Integral. *International Journal of Intelligent Systems*, *32*(12), 1312–1331. Retrieved from http://doi.wiley.com/10.1002/int.21901 doi: 10.1002/int.21901

Zhu, J. (1996, jan). Data Envelopment Analysis with Preference Structure. *Journal of the Operational Research Society*, *47*(1), 136–150. Retrieved from https://www.tandfonline.com/doi/full/10.1057/jors.1996.12 doi: 10.1057/jors.1996.12

Appendix A

MATLAB and CPLEX integration

Admitting that the IBM ILOG CPLEX Optimisation Studio is successfully installed on the computer, the exploitation of its solver in the MATLAB environment required the adoption of the following steps:

1. With MATLAB open, the `pathtool` command must be typed;
2. After that, a window appears, and the 'Add Folder...' button must be pressed and the path `C:\Program Files\IBM\ILOG\CPLEX_Studio_Community1263\cplex\matlab\x64_win64` (for Microsoft Windows Operating System) or `Macintosh HD/Applications/CPLEX_Studio_Community128/cplex/matlab/x86-64_osx` (for Apple Macintosh Operating System) should be chosen;
3. Saving and closing are the final steps to link the IBM ILOG CPLEX Optimisation Studio solver to MATLAB.

Appendix B

Utilities

B.1 Values

Table B.1: Utility values for each DMU according per indicator.

		Timeliness of services			Service availability	Care appropriateness			Economic-financial
		g_1	g_5	g_7	g_6	g_2	g_3	g_4	g_8
Group B	a_1	0.9166	0.8711	0.9115	0.5767	0.4275	0.6947	1.0000	0.5599
	a_2	0.5649	0.2218	0.6464	0.7110	0.6594	0.5360	0.3778	0.3839
Group C	a_3	0.7078	0.2210	0.7202	0.3037	0.5761	0.1330	0.7556	0.9349
	a_4	0.6940	0.4943	0.4934	0.4119	0.3886	0.1643	0.3433	0.0187
	a_5	0.5336	0.3254	0.4382	0.9419	0.5407	0.3833	0.0519	0.2623
	a_6	0.5305	0.5619	0.3423	0.6908	0.7321	0.4094	0.2582	0.0000
	a_7	0.4430	0.4026	0.7850	0.0000	0.5876	0.2864	0.2203	0.8903
	a_8	0.3863	0.5317	0.9891	0.7185	0.6757	0.4379	0.3906	0.7152
	a_9	0.3719	0.2427	0.2086	0.2444	0.8714	0.0000	0.2451	0.4845
	a_{10}	0.3278	0.1780	0.7761	0.7846	0.6370	0.1582	0.3886	0.1854
	a_{11}	0.2953	0.3510	0.4864	0.5495	0.6899	0.7090	0.3798	1.0000
	a_{12}	0.2943	0.2623	0.8930	0.6503	0.6433	0.6808	0.5647	0.9248
Group D	a_{13}	0.7199	0.2952	0.3524	0.5878	0.7830	0.4911	0.1852	0.9423
	a_{14}	0.4792	0.1097	1.0000	0.1250	0.5291	0.8519	0.4914	0.6516
	a_{15}	0.4669	0.7181	0.3820	0.9955	0.7643	0.0424	0.3196	0.7788
	a_{16}	0.4657	0.1037	0.0324	0.9881	0.6272	0.4359	0.0000	0.2140
	a_{17}	0.4537	0.2214	0.0610	0.9980	0.8411	0.2492	0.1995	0.9576
	a_{18}	0.4242	0.5000	0.5318	0.9838	0.6270	0.2892	0.2116	0.9169
	a_{19}	0.3362	0.1434	0.6775	0.5136	0.5576	0.4272	0.0383	0.8814
Group E	a_{20}	0.9300	0.1290	0.1504	0.7657	0.7352	0.4724	0.0335	0.6294
	a_{21}	0.6914	0.3095	0.0000	0.3672	0.6536	0.4204	0.0573	0.7034
	a_{22}	0.5955	0.5069	0.3449	0.9995	0.5265	1.0000	0.2557	0.9916
	a_{23}	0.5334	0.5558	0.7399	0.7274	0.5811	0.6452	0.2110	0.8698
	a_{24}	0.5148	0.3212	0.4488	0.9381	0.5788	0.0575	0.0820	0.7424
	a_{25}	0.4884	0.3023	0.1565	0.2416	0.5975	0.1713	0.1614	0.8883

B.2 Plots

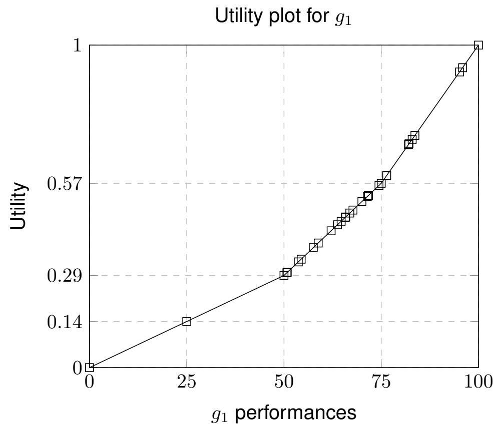

Figure B.1: Utility plot for g_1.

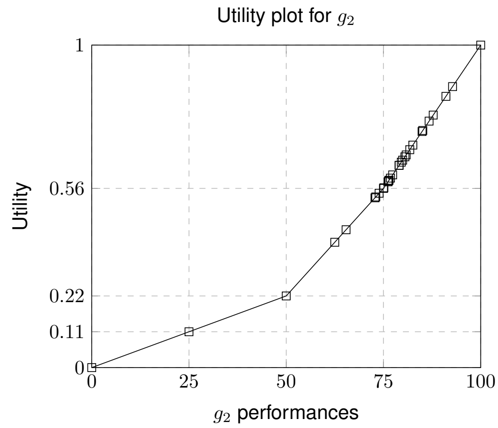

Figure B.2: Utility plot for g_2.

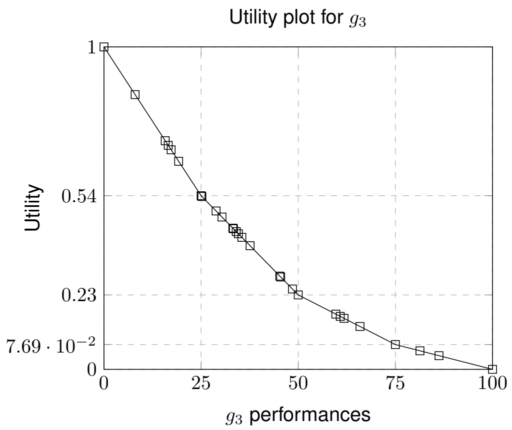

Figure B.3: Utility plot for g_3.

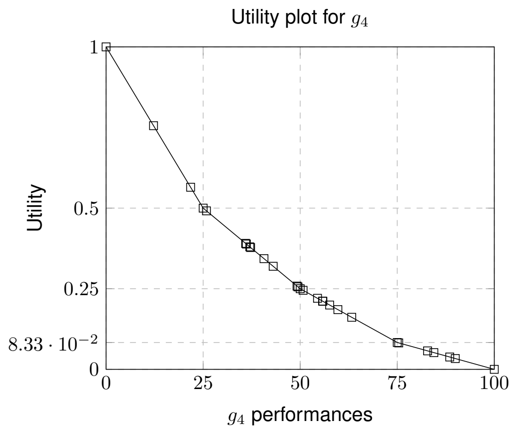

Figure B.4: Utility plot for g_4.

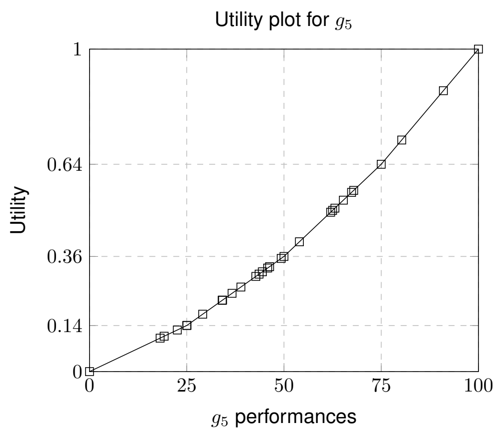

Figure B.5: Utility plot for g_5.

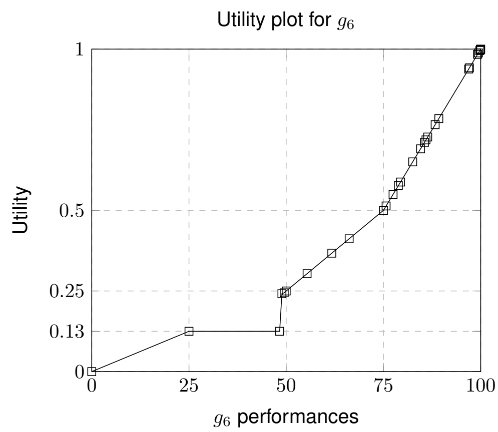

Figure B.6: Utility plot for g_6.

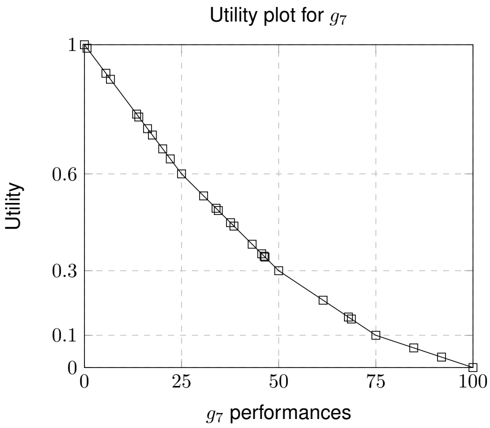

Figure B.7: Utility plot for g_7.

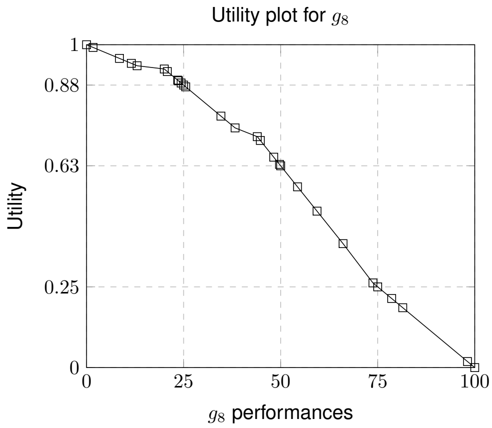

Figure B.8: Utility plot for g_8.

Appendix C

Case study results

C.1 Benchmarks

C.2 Slacks

Table C.1: Optimal Möbius coefficients' benchmarks for the inefficient DMUs of SCENARIO 1.

		GROUP B	**GROUP C**			**GROUP D**				**GROUP E**	
		a_1	a_3	a_8	a_{12}	a_{13}	a_{15}	a_{17}	a_{18}	a_{22}	a_{23}
GROUP B	a_2	0.219836762	0	0.187433764	0.194656301	0.181287758	0	0.182652753	0	0.034132663	0
GROUP C	a_4	0.347464499	0.459720009	0	0	0	0.192815492	0	0	0	0
	a_5	0.119691635	0	0	0	0	0	0.191483897	0.42720705	0.261617418	0
	a_7	0.025121	0.194045007	0.001129727	0.289551894	0	0	0	0	0	0.490152372
	a_{10}	0	0	0.16800413	0.463653686	0	0	0	0.368342184	0	0
GROUP D	a_{16}	0	0.01480115	0	0	0	0	0.734240015	0	0.250958835	0
	a_{19}	0	0.581925148	0	0.208362511	0	0	0	0	0.209712342	0
GROUP E	a_{21}	0.011039141	0.402770485	0	0	0.403819568	0.01949428	0	0	0.162876526	0
	a_{24}	0.125148812	0	0	0	0	0	0.225491057	0.519034745	0.130325386	0
	a_{25}	0.070751922	0.757305366	0	0	0.041835364	0.124942487	0	0	0.005164861	0

Table C.2: Analogous Möbius coefficients' benchmarks for the inefficient DMUs of SCENARIO 1.

		GROUP B	**GROUP C**			**GROUP D**				**GROUP E**	
		a_1	a_3	a_8	a_{12}	a_{13}	a_{15}	a_{17}	a_{18}	a_{22}	a_{23}
GROUP B	a_2	0.219836762	0	0.187433764	0.194656301	0.181287758	0	0.182652753	0	0.034132663	0
GROUP C	a_4	0.347464499	0.459720009	0	0	0	0.192815492	0	0	0	0
	a_5	0.123540584	0	0	0	0	0.32435097	0.123206357	0.247441245	0.181460843	0
	a_7	0.025121	0.194045007	0.001129727	0.289551894	0	0	0	0	0	0.490152372
	a_{10}	0.214506121	0	0.43189239	0	0	0.231324194	0	0.122277295	0	0
GROUP D	a_{16}	0.024332649	0	0	0	0	0	0.741437457	0	0.234229895	0
	a_{19}	0.160082097	0.509895465	0	0.181643529	0	0	0.001402887	0	0.146976022	0
GROUP E	a_{21}	0.226630711	0.235636279	0	0	0.533474393	0.004258616	0	0	0	0
	a_{24}	0.119284137	0.010760051	0	0	0	0.869955812	0	0	0	0
	a_{25}	0.073652192	0.755056974	0	0	0.04357956	0.124737527	0	0	0.002973746	0

Table C.3: Optimal Möbius coefficients' benchmarks for the inefficient DMUs of SCENARIO 2.

		GROUP B	**GROUP C**			**GROUP D**				**GROUP E**		
		a_1	a_3	a_8	a_{12}	a_{13}	a_{15}	a_{17}	a_{18}	a_{20}	a_{22}	a_{23}
GROUP B	a_2	0.202986937	0	0.150890625	0.286343712	0.086021847	0	0.168462848	0	0.10529403	0	0
GROUP C	a_4	1	0	0	0	0	0	0	0	0	0	0
	a_5	0.130573173	0	0	0	0	0.594784204	0.00968388	0	0	0.264958744	0
	a_7	0.025121	0.194045007	0.001129727	0.289551894	0	0	0	0	0	0	0.490152372
	a_{10}	0.214506121	0	0.43189239	0	0	0.231324194	0	0.122277295	0	0	0
GROUP D	a_{16}	0.024332649	0	0	0	0	0	0.741437457	0	0	0.234229895	0
	a_{19}	0.133497798	0.403819577	0	0.425037231	0	0	0.037645394	0	0	0	0
GROUP E	a_{21}	0.431833174	0	0	0	0.153077441	0	0.415089386	0	0	0	0
	a_{24}	0.137058262	0	0	0	0	0.862941738	0	0	0	0	0
	a_{25}	0.071933048	0.753913932	0	0.010056467	0.036434012	0.127028264	0	0	0	0.000634277	0

Table C.4: Analogous Möbius coefficients' benchmarks for the inefficient DMUs of SCENARIO 2.

		GROUP B	**GROUP C**		**GROUP D**		**GROUP E**		
		a_1	a_8	a_{12}	a_{13}	a_{17}	a_{20}	a_{22}	a_{23}
GROUP B	a_2	0.202986937	0.150890625	0.286343712	0.086021847	0.168462848	0.10529403	0	0
GROUP C	a_4	1	0	0	0	0	0	0	0
	a_5	0.027339114	0.027291908	0.109885442	0	0	0	0.835483536	0
	a_7	0.067011734	0	0.476109687	0	0	0	0.099472422	0.357406157
	a_{10}	0.013194557	0.525771748	0.176169886	0	0.040637234	0	0.244226575	0
GROUP D	a_{16}	0.025798696	0	0	0	0.328207473	0	0.645993831	0
	a_{19}	0.211802967	0	0.445791899	0	0	0	0.342405134	0
GROUP E	a_{21}	0.363994374	0	0	0.636005626	0	0	0	0
	a_{24}	0	0.14390512	0.059704804	0	0.0758291	0	0.720560975	0
	a_{25}	0.161848352	0	0.562460728	0.083148436	0	0	0.192542485	0

Table C.5: Optimal Möbius coefficients' benchmarks for the inefficient DMUs of SCENARIO 3.

		GROUP B	**GROUP C**		**GROUP D**			**GROUP E**	
		a_1	a_8	a_{12}	a_{13}	a_{15}	a_{17}	a_{20}	a_{22}
GROUP B	a_2	0.142948885	0.243600624	0.253033508	0	0	0.056028221	0.304388762	0
GROUP C	a_4	1	0	0	0	0	0	0	0
	a_5	0.101025394	0.051992374	0	0	0.069151605	0	0	0.777830627
	a_7	0.468834739	0.267746972	0	0.263418288	0	0	0	0
	a_{10}	0.167525769	0.509649388	0	0	0.214656697	0	0	0.108168146
GROUP D	a_{16}	0.025798696	0	0	0	0	0.328207473	0	0.645993831
	a_{19}	0.634036568	0	0	0.365963432	0	0	0	0
GROUP E	a_{21}	0.363994374	0	0	0.636005626	0	0	0	0
	a_{24}	0.107578843	0.053347587	0	0	0.231248299	0	0	0.607825271
	a_{25}	0.521800281	0	0	0.478199719	0	0	0	0

Table C.6: Analogous Möbius coefficients' benchmarks for the inefficient DMUs of SCENARIO 3.

		GROUP B	**GROUP C**		**GROUP D**			**GROUP E**	
		a_1	a_8	a_{12}	a_{13}	a_{15}	a_{17}	a_{20}	a_{22}
GROUP B	a_2	0.142948885	0.243600624	0.253033508	0	0	0.056028221	0.304388762	0
GROUP C	a_4	1	0	0	0	0	0	0	0
	a_5	0.101025394	0.051992374	0	0	0.069151605	0	0	0.777830627
	a_7	0.468834739	0.267746972	0	0.263418288	0	0	0	0
	a_{10}	0.167525769	0.509649388	0	0	0.214656697	0	0	0.108168146
GROUP D	a_{16}	0.025798696	0	0	0	0	0.328207473	0	0.645993831
	a_{19}	0.619779496	0	0	0.274187256	0	0	0.106033248	0
GROUP E	a_{21}	0.339252178	0	0	0.476734206	0	0	0.184013617	0
	a_{24}	0.107578843	0.053347587	0	0	0.231248299	0	0	0.607825271
	a_{25}	0.47653897	0	0	0.186841862	0	0	0.336619168	0

Table C.7: Optimal Möbius coefficients' slacks of the inefficient DMUs of SCENARIO 1.

Groups	Institutions	$u_1(g_1(a_k))$	$u_2(g_2(a_k))$	$u_3(g_3(a_k))$	$u_4(g_4(a_k))$	$u_5(g_5(a_k))$	$u_6(g_6(a_k))$	$u_7(g_7(a_k))$	$u_8(g_8(a_k))$	$\min\{u_1(g_1(a_k)), u_5(g_5(a_k))\}$	$\min\{u_2(g_2(a_k)), u_5(g_5(a_k))\}$	$\min\{u_3(g_3(a_k)), u_8(g_8(a_k))\}$
GROUP B	a_2	0.0000	0.0000	0.0000	0.1039	0.2317	0.0000	0.0000	0.4328	0.2044	0.1342	0.1222
GROUP C	a_4	0.0399	0.1722	0.1464	0.4132	0.0484	0.1200	0.2281	0.7558	0.0000	0.0000	0.2452
	a_5	0.0000	0.0771	0.1327	0.2633	0.1675	0.0000	0.0000	0.6392	0.1351	0.1144	0.2354
	a_7	0.0645	0.0068	0.2707	0.2188	0.0111	0.6191	0.0000	0.0000	0.0000	0.0000	0.2673
	a_{10}	0.0298	0.0057	0.3375	0.0168	0.2171	0.0000	0.0000	0.7013	0.1648	0.2171	0.3375
GROUP D	a_{16}	0.0273	0.1310	0.0000	0.2218	0.1893	0.0000	0.1096	0.7518	0.1893	0.1893	0.2198
	a_{19}	0.2619	0.0221	0.0018	0.5727	0.1462	0.0082	0.0000	0.0633	0.1462	0.1462	0.0000
GROUP E	a_{21}	0.0006	0.0000	0.0029	0.3807	0.0049	0.1811	0.5061	0.2365	0.0000	0.0000	0.0000
	a_{24}	0.0000	0.0584	0.3661	0.2313	0.1633	0.0000	0.0000	0.1487	0.1240	0.1078	0.3481
	a_{25}	0.2040	0.0000	0.0096	0.5306	0.0314	0.1833	0.5177	0.0012	0.0000	0.0000	0.0000

Table C.8: Analogous Möbius coefficients' slacks of the inefficient DMUs of SCENARIO 1.

Groups	Institutions	$u_1(g_1(a_k))$	$u_2(g_2(a_k))$	$u_3(g_3(a_k))$	$u_4(g_4(a_k))$	$u_5(g_5(a_k))$	$u_6(g_6(a_k))$	$u_7(g_7(a_k))$	$u_8(g_8(a_k))$	$\min\{u_1(g_1(a_k)), u_5(g_5(a_k))\}$	$\min\{u_2(g_2(a_k)), u_5(g_5(a_k))\}$	$\min\{u_3(g_3(a_k)), u_8(g_8(a_k))\}$
GROUP B	a_2	0.0000	0.0000	0.0000	0.1039	0.2317	0.0000	0.0000	0.4328	0.2044	0.1342	0.1222
GROUP C	a_4	0.0399	0.1722	0.1464	0.4132	0.0484	0.1200	0.2281	0.7558	0.0000	0.0000	0.2452
	a_5	0.0000	0.1143	0.0000	0.2986	0.2581	0.0000	0.0000	0.5843	0.1579	0.2033	0.1028
	a_7	0.0645	0.0068	0.2707	0.2188	0.0111	0.6191	0.0000	0.0000	0.0000	0.0000	0.2673
	a_{10}	0.1955	0.0000	0.2251	0.0944	0.4657	0.0000	0.0000	0.5359	0.3356	0.3706	0.1962
GROUP D	a_{16}	0.0325	0.1301	0.0000	0.2321	0.2004	0.0000	0.1158	0.7419	0.2004	0.1896	0.2167
	a_{19}	0.3131	0.0000	0.0228	0.6475	0.2312	0.0000	0.0486	0.0000	0.2312	0.1602	0.0000
GROUP E	a_{21}	0.0691	0.0000	0.0305	0.4475	0.1005	0.1529	0.5659	0.1498	0.0995	0.0000	0.0000
	a_{24}	0.0083	0.1433	0.0637	0.3235	0.4098	0.0000	0.0000	0.0120	0.1913	0.3569	0.0476
	a_{25}	0.2049	0.0000	0.0100	0.5315	0.0327	0.1830	0.5185	0.0000	0.0013	0.0000	0.0000

Table C.9: Optimal Möbius coefficients' slacks of the inefficient DMUs of SCENARIO 2.

Groups	**Institutions**	$u_1(g_1(a_k))$	$u_2(g_2(a_k))$	$u_3(g_3(a_k))$	$u_4(g_4(a_k))$	$u_5(g_5(a_k))$	$u_6(g_6(a_k))$	$u_7(g_7(a_k))$	$u_8(g_8(a_k))$	$\min\{u_1(g_1(a_k)), u_5(g_5(a_k))\}$	$\min\{u_2(g_2(a_k)), u_5(g_5(a_k))\}$	$\min\{u_3(g_3(a_k)), u_8(g_8(a_k))\}$
GROUP B	a_2	0.0000	0.0000	0.0000	0.0989	0.1866	0.0000	0.0000	0.4111	0.1647	0.0966	0.1247
GROUP C	a_4	0.2226	0.0389	0.5304	0.6567	0.3768	0.1648	0.4181	0.5412	0.3768	0.0389	0.5412
	a_5	0.0260	0.1174	0.0000	0.3384	0.3519	0.0000	0.0000	0.5460	0.2025	0.2940	0.1012
	a_7	0.0645	0.0068	0.2707	0.2188	0.0111	0.6191	0.0000	0.0000	0.0000	0.0000	0.2673
	a_{10}	0.1955	0.0000	0.2251	0.0944	0.4657	0.0000	0.0000	0.5359	0.3356	0.3706	0.1962
GROUP D	a_{16}	0.0325	0.1301	0.0000	0.2321	0.2004	0.0000	0.1158	0.7419	0.2004	0.1896	0.2167
	a_{19}	0.2142	0.0372	0.0180	0.6479	0.1820	0.0000	0.1169	0.0000	0.1820	0.1227	0.0000
GROUP E	a_{21}	0.0029	0.0000	0.0582	0.4857	0.2038	0.3861	0.4729	0.0801	0.2038	0.0122	0.0000
	a_{24}	0.0137	0.1393	0.0743	0.3309	0.4179	0.0000	0.0058	0.0064	0.2011	0.3571	0.0558
	a_{25}	0.2000	0.0000	0.0097	0.5334	0.0319	0.1839	0.5226	0.0000	0.0000	0.0000	0.0000

Table C.10: Analogous Möbius coefficients' slacks of the inefficient DMUs of SCENARIO 2.

Groups	**Institutions**	$u_1(g_1(a_k))$	$u_2(g_2(a_k))$	$u_3(g_3(a_k))$	$u_4(g_4(a_k))$	$u_5(g_5(a_k))$	$u_6(g_6(a_k))$	$u_7(g_7(a_k))$	$u_8(g_8(a_k))$	$\min\{u_1(g_1(a_k)), u_5(g_5(a_k))\}$	$\min\{u_2(g_2(a_k)), u_5(g_5(a_k))\}$	$\min\{u_3(g_3(a_k)), u_8(g_8(a_k))\}$
GROUP B	a_2	0.0000	0.0000	0.0000	0.0989	0.1866	0.0000	0.0000	0.4594	0.1647	0.0966	0.2263
GROUP C	a_4	0.2226	0.0389	0.5304	0.6567	0.3768	0.1648	0.4181	0.4211	0.3768	0.0389	0.4211
	a_5	0.0319	0.0000	0.5579	0.2618	0.1653	0.0000	0.0000	0.7531	0.1613	0.1531	0.7289
	a_7	0.0084	0.0074	0.4144	0.2164	0.0297	0.7077	0.0000	0.0000	0.0217	0.0000	0.3959
	a_{10}	0.1031	0.0000	0.4555	0.0000	0.2921	0.0000	0.0000	0.6105	0.2156	0.2862	0.4691
GROUP D	a_{16}	0.0915	0.0000	0.3098	0.2565	0.3189	0.0000	0.2339	0.7999	0.3189	0.3074	0.5719
	a_{19}	0.1930	0.0000	0.3658	0.5128	0.3316	0.2407	0.0317	0.0000	0.3316	0.2376	0.3071
GROUP E	a_{21}	0.1001	0.0000	0.1448	0.4245	0.1953	0.2166	0.5559	0.1655	0.1953	0.0339	0.0503
	a_{24}	0.0219	0.0000	0.7856	0.2073	0.1530	0.0000	0.0000	0.2571	0.1321	0.1530	0.7776
	a_{25}	0.0000	0.0000	0.5574	0.3827	0.1084	0.4588	0.5890	0.0000	0.1084	0.0366	0.5133

Table C.11: Optimal Möbius coefficients' slacks of the inefficient DMUs of SCENARIO 3.

Groups	**Institutions**	$u_1(g_1(a_k))$	$u_2(g_2(a_k))$	$u_3(g_3(a_k))$	$u_4(g_4(a_k))$	$u_5(g_5(a_k))$	$u_6(g_6(a_k))$	$u_7(g_7(a_k))$	$\min\{u_1(g_1(a_k)), u_5(g_5(a_k))\}$	$\min\{u_2(g_2(a_k)), u_5(g_5(a_k))\}$
GROUP B	a_2	0.0432	0.0000	0.0000	0.0246	0.1503	0.0000	0.0000	0.1149	0.0869
GROUP C	a_4	0.2226	0.0389	0.5304	0.6567	0.3768	0.1648	0.4181	0.3768	0.0389
	a_5	0.0746	0.0000	0.4904	0.2904	0.2342	0.0000	0.0000	0.2093	0.1894
	a_7	0.2798	0.0000	0.2859	0.4019	0.2259	0.6176	0.0000	0.1870	0.0179
	a_{10}	0.1873	0.0000	0.2986	0.0743	0.4479	0.0000	0.0000	0.3199	0.3736
GROUP D	a_{16}	0.0915	0.0000	0.3098	0.2565	0.3189	0.0000	0.2339	0.3189	0.3074
	a_{19}	0.5084	0.0000	0.1930	0.6635	0.5169	0.0672	0.0294	0.5169	0.2357
GROUP E	a_{21}	0.1001	0.0000	0.1448	0.4245	0.1953	0.2166	0.5559	0.1953	0.0339
	a_{24}	0.0743	0.0000	0.6582	0.2757	0.2750	0.0000	0.0000	0.2092	0.2273
	a_{25}	0.3341	0.0000	0.4260	0.4490	0.2934	0.3404	0.4876	0.2934	0.0619

Table C.12: Analogous Möbius coefficients' slacks of the inefficient DMUs of SCENARIO 3.

Groups	**Institutions**	$u_1(g_1(a_k))$	$u_2(g_2(a_k))$	$u_3(g_3(a_k))$	$u_4(g_4(a_k))$	$u_5(g_5(a_k))$	$u_6(g_6(a_k))$	$u_7(g_7(a_k))$	$\min\{u_1(g_1(a_k)), u_5(g_5(a_k))\}$	$\min\{u_2(g_2(a_k)), u_5(g_5(a_k))\}$
GROUP B	a_2	0.0432	0.0000	0.0000	0.0246	0.1503	0.0000	0.0000	0.1149	0.0869
GROUP C	a_4	0.2226	0.0389	0.5304	0.6567	0.3768	0.1648	0.4181	0.3768	0.0389
	a_5	0.0746	0.0000	0.4904	0.2904	0.2342	0.0000	0.0000	0.2093	0.1894
	a_7	0.2798	0.0000	0.2859	0.4019	0.2259	0.6176	0.0000	0.1870	0.0179
	a_{10}	0.1873	0.0000	0.2986	0.0743	0.4479	0.0000	0.0000	0.3199	0.3736
GROUP D	a_{16}	0.0915	0.0000	0.3098	0.2565	0.3189	0.0000	0.2339	0.3189	0.3074
	a_{19}	0.5279	0.0000	0.1881	0.6358	0.4911	0.0862	0.0000	0.4911	0.2162
GROUP E	a_{21}	0.1339	0.0000	0.1363	0.3764	0.1505	0.2496	0.5049	0.1505	0.0000
	a_{24}	0.0743	0.0000	0.6582	0.2757	0.2750	0.0000	0.0000	0.2092	0.2273
	a_{25}	0.3960	0.0000	0.4105	0.3610	0.2114	0.4008	0.3943	0.2114	0.0000