Preference information incorporation for decision-making in a DEA model using the Choquet integral: Performance assessment of secondary healthcare providers

Miguel Pereira
miguelalvespereira@tecnico.ulisboa.pt

Instituto Superior Técnico, Lisboa, Portugal

May 2018

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 (MCDM) 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

1. Introduction

New challenges present themselves to the healthcare industry on a daily basis and require a proficient management of the entirety of its dedicated gross domestic product (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). It is precisely this that the healthcare industry is desperate for nowadays, especially nation-wide institutions as, *e.g.*, Portugal's National Health Service (SNS, from the Portuguese abbreviation of *Serviço Nacional de Saúde*).

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. Nowadays, its total healthcare spending reached over 16€ billion in 2015. Effectively, according to Portugal's National Statistics' Institute, 8.9% of Portugal's 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 SNS 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 research places emphasis on the application of a well-known performance measurement technique, *sc.* 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.

Hence, this paper is composed of five sections aligned with the objectives and methodology described above. In this first section, an introduction is presented with a focus on the motivation that led to the research, alongside a compelling description of the problem at hand by aiming attention at the global health environment before particularising the Portuguese situation. An appraisal of the knowledge gap is the second and rational section before developing the cherished model. With the fundamentals settled, a methodology section is essential in furtherance of describing the methods in which the soon-to-be-named PRICDEA model is based on: the 'two-phase method' built on the additive (ADD) DEA model (Gouveia, Dias, & Antunes, 2008), the Choquet multiple criteria preference aggregation model (Bottero, Ferretti, Figueira, Greco, & Roy, 2018), and the linear program for inferring the weights of interacting criteria from a reference set (Marichal & Roubens, 2000). 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 fourth section 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 subsections of this section, where the PRICDEA model is tested in three different scenarios discussed with the decision-maker. The achievements and limitations of this research are presented in the last section of this paper.

2. Knowledge gap

2.1. DEA in healthcare

Challenges such as those mentioned in Section 1 were overflowing healthcare organisations in the 1980s just as healthcare managers were pursuing for efficiency improvements, particularly in the United States of America. A few years later, governmental initiatives extended the fixed pricing mechanism of diagnostic-related groupings of the previous decade to physicians' services through resource based relative schedule. Following the reasoning of Ozcan (2008), albeit suchlike pricing mechanisms influenced the control of the amounts paid to healthcare organisations and healthcare professionals, a deeper insight on 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, Cooper, & Rhodes (1978).

In the context of healthcare, the measurement of routine nursing service efficiency was DEA's first incursion in this domain and, from than onward, this frontier technique has become the dominant approach to efficiency measurement in healthcare, among other areas. *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.

2.2. 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 to 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, which began with the work of Thompson, Singleton, Thrall, & Smith (1986).

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 MCDM literature are planted in the aforementioned interactive multiobjective linear programming (MOLP) procedure by Golany (1988).

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. 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 research dwells.

3. Methodology

3.1. Where does PRICDEA come from?

In healthcare, DEA is considered to be the leading performance evaluation approach. Indeed, as a method that evaluates the relative efficiency of DMUs using multiple inputs to produce multiple outputs, there are several types of models 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 variable returns-to-scale) based on the 'two-phase method' developed by Gouveia *et al.* (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 *et al.* (2018), and indirectly, adapted from the linear program developed by Marichal & Roubens (2000).

The concept behind the PRICDEA model is quite transparent. 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.

Foremost, let A denote a set containing m actions $a_1, ..., a_k, ..., a_m$, G the criteria set with n criteria, $g_1, ..., g_i, ..., g_n$, and O the set of pairs of criteria g_i and g_j. 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 $u_i(g_i(a_k))$. Then, the Choquet capacity is defined as 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: $\forall S \subseteq T \subseteq G : \mu(S) \leqslant \mu(T)$.

However, it is usual in Choquet integral literature to formulate this multiple criteria preference aggregator 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 $S \subseteq G$,

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

Consequently, redefining the properties above and the coefficients required to define the fuzzy measure, for all $S \subseteq G$, one can reformulate the Choquet integral as

$$\begin{aligned} C_\mu(a_k) = & \sum_{g_i \in G} m(g_i) u_i(g_i(a_k)) + \\ & + \sum_{g_i, g_j \in O} (m(g_i, g_j) * \\ & * \min\{u_i(g_i(a_k)), u_j(g_j(a_k))\}) \end{aligned} \tag{2}$$

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 using a variant of the method employed by Bottero *et al.* (2018), 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

$$\begin{aligned} \min \quad & d \\ \text{s.t.} \quad & C_\mu(a_x) - C_\mu(a_y) \leqslant d \\ & \text{Ranking of criteria:} \\ & \bullet\ m(g_i) - m(g_j) \geqslant \varepsilon, \text{ if } g_i \succ_G g_j \\ & \bullet\ m(g_i) = m(g_j), \text{ if } g_i \sim_G g_j \\ & \text{Ranking of pairs of criteria:} \\ & \bullet\ m(g_{ij}) - m(g_{kl}) \geqslant \varepsilon, \\ & \quad \text{if } g_{ij} \succ_P g_{kl} \\ & \bullet\ m(g_{ij}) = m(g_{kl}), \text{ if } g_{ij} \sim_P g_{kl} \\ & \text{Boundary and monotonicity} \\ & \text{conditions:} \\ & \bullet\ \sum_{g_i \in G} m(g_i) + \sum_{g_i, g_j \subseteq G} m(g_i, g_j) = \\ & \quad = 1 \\ & \bullet\ m(g_i) \geqslant 0, \forall g_i \in G \\ & \bullet\ m(g_i) + \sum_{j \in O} m(g_i, g_j) \geqslant 0, \\ & \quad \forall g_i \in G, \forall O \subseteq G \setminus g_i \end{aligned} \tag{3}$$

for $x = 1, ..., m$ and $y = 1, ..., m$, and a preset $\varepsilon = 0.001$, where a_x and a_y 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, $\succeq_G$ is the partial preorder on the set of criteria G, and $\succeq_P$ is the partial preorder on the set of pairs of interacting criteria P.

Phase 2 Solve the modified rendition of the 'weighted ADD model'

$$\begin{aligned} \min \quad & z_k = -\sum_{j=1}^{n} M_j s_j \\ \text{s.t.} \quad & \sum_{i=1}^{m} U_{ij}\lambda_i - s_j = U_{kj} \leqslant 0 \\ & \sum_{i=1}^{m} \lambda_i = 1 \\ & \lambda_i \geqslant 0,\ s_j \geqslant 0\,. \end{aligned} \tag{4}$$

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

3.2. 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 3.1, 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 section, 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 inferred optimal Möbius coefficients per DMU with the linear program in Equation (3), which will be used in the next phase of the PRICDEA model.

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) 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) 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 this questioning procedure that is used in the ranking step of PRICDEA's Phase 1 to build the constraints of Equation (3) and infer the optimal Möbius coefficients.

4. Case study

4.1. Overview

As declared in Section 1, 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, 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 for 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 improve the efficiency and effectiveness of healthcare providers (Ferreira & Marques, 2017).

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 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 paper emerges from.

4.2. Stakeholders and their representatives

In this research, 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.

4.3. Data and sample

It should be clear by now that the main goal of the case study engaged in this work 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 a more applicable database.

Therefore, at this moment, it can be asserted with certainty that twenty-five institutions 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 (Group B to Group E, according to ACSS nomenclature). However, due to understandable privacy issues, their names are in no position to be discretised.

4.4. Variables

In consonance with what was stated in the previous subsection, the eight indicators g_n, for $n = 1, ..., 8$, yielded from the process of constructing a more fitting database for the case study 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.

To support the decision-making problem of this case study and allow the comparison between the DMUs itemised in Subsection 4.3 and operationalised by the aforementioned indicators, a family of five criteria and respective subcriteria was built in line with the work of Ferreira & Marques (2017) and is specified in Table 1.

As almost every concept in healthcare, quality is complex and non-consensual. However, Donabedian (1988) categorised it into three interrelated categories: structural quality, process quality, and outcomes. In spite of the usual criticism, these definitions have been regularly applied and, in consonance with Ferreira, Marques, Nunes, & Figueira (2017), it is safe to affirm that Care appropriateness is a quality-related criteria 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 the service availability and organisation barriers classes, respectively.

4.5. 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 3.2, 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, but the former, due to its complexity, was unable to use the aforestated CPLEX version and purely run on MATLAB.

4.6. Findings and discussion

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

Table 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

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.

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, 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 by Bottero *et al.* (2018). 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. In the end, the lingering utility values were computed by simply applying the linear interpolation

$$u_j(g_j) = u_j(g_j^l) + \frac{g_j - g_j^l}{g_j^u - g_j^l}\left(u_j(g_j^u) - u_j(g_j^l)\right) \quad (5)$$

4.6.1 Scenario 1: the importance of operating costs

In consonance with the previous paragraph, 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}.

Then, following the step-by-step procedure (Bottero et al., 2018), the decision-maker ranked the finite set of reference projects $P = p_1, p_2, p_3, p_4, p_5, p_6, p_7, p_8, p_{15}, p_{25}, p_{38}$ as $p_8 \succ p_{15} \sim p_{25} \sim p_{38} \succ p_1 \sim p_2 \sim p_5 \succ p_3 \sim p_4 \sim p_6 \sim p_7$ and introduced 3, 1, and 2 blank cards between them, respectively, not to mention revealing the value of the ratio-z as 3. With these preferences in mind, both PRICDEA's Phase 1 and DecSpace's - a web application that makes use of MCDM methods to support a decision process by giving a possible solution to a given problem - SRF software, developed by Roy & Figueira (1998), were properly used to compute the desired coefficients. Wherein the former's linear program was already broadly described, the latter required intermediate computations before pursuing the attainment of the analogous Möbius coefficients. These coefficients are unfolded in Table 2.

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. As for the relationship between the cited benchmarks and the DMUs below the frontier, 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. 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.

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

4.6.2 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. Thus, $p_{15} \sim p_{25} \sim p_{38} \succ p_1 \sim p_2 \sim p_5 \succ p_3 \sim p_4 \sim p_6 \sim p_7 \succ p_8$, with 1, 2, and 3 blank cards between them, respectively, is now the decision-maker's ranking with a ratio-$z = 3$. This means that the Möbius coefficients would change, although not significantly in comparison to the first scenario. Table 3 displays all these values.

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. 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. 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.

Lastly, the slack values for the optimal and analogous Möbius coefficients are now addressed. 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. 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 encour-

Table 2: Möbius coefficients for SCENARIO 1.

Indicators	Optimal Möbius coefficients	Analogous Möbius coefficients
g_1	0.200000000	0.218015666
g_2	0.200000000	0.218015666
g_3	0.100000000	0.130548303
g_4	0.100000000	0.130548303
g_5	0.200000000	0.218015666
g_6	0.100000000	0.130548303
g_7	0.100000000	0.130548303
g_8	0.400000000	0.391644909
g_{15}	−0.133333333	−0.160574413
g_{25}	−0.133333333	−0.160574413
g_{38}	−0.133333333	−0.246736292

Table 3: Möbius coefficients for SCENARIO 2.

Indicators	Optimal Möbius coefficients	Analogous Möbius coefficients
g_1	0.300000000	0.211395541
g_2	0.300000000	0.211395541
g_3	0.200000000	0.156069364
g_4	0.200000000	0.156069364
g_5	0.300000000	0.211395541
g_6	0.200000000	0.156069364
g_7	0.200000000	0.156069364
g_8	0.100000000	0.082576383
g_{15}	−0.266666667	−0.175061932
g_{25}	−0.266666667	−0.175061932
g_{38}	−0.266666667	0.009083402

aging 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.

The sum and substance of SCENARIO 2 indicates 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.

4.6.3 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. Nevertheless, the expert recycled the other indicators and respective utilities, and ranked the former as $\succ p_{15} \sim p_{25} \succ p_1 \sim p_2 \sim p_5 \succ p_3 \sim p_4 \sim p_6 \sim p_7$ with 1 and 2 blank cards between them, respectively. This time, ratio-z was attributed a value of 2. Despite the Möbius coefficients 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 4 has these numbers arranged.

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. 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.

The slack results expose 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},

Table 4: Möbius coefficients for Scenario 3.

Indicators	Optimal Möbius coefficients	Analogous Möbius coefficients
g_1	0.250000000	0.25
g_2	0.250000000	0.25
g_3	0.150000000	0.15625
g_4	0.150000000	0.15625
g_5	0.250000000	0.25
g_6	0.150000000	0.15625
g_7	0.150000000	0.15625
g_{15}	−0.175000000	−0.1875
g_{25}	−0.175000000	−0.1875

which need to improve quite significantly in almost every single aspect. 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 the analogous Möbius coefficients, the results were one and the same, save for minor numerical differences.

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. Conclusions

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.

As for limitations, 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. 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 to similar results between scenarios lies.

References

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

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

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

Donabedian, A. (1988). The quality of care: How can it be assessed? *Journal of the Amer-*

ican Medical Association, *260*(12), 1743–1748. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/3045356

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

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 applications.* Norwell: Kluwer Academic Publishers.

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

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

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

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

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

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

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*.

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

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., 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

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