Do PPP hospitals outperform the corporatized ones?

The Portuguese experience

Gonçalo Mateus da Silva Santos

Thesis to obtain the Master of Science Degree in

Industrial Engineering and Management

Supervisors: Prof. Maria Isabel Craveiro Pedro
Prof. Diogo Filipe da Cunha Ferreira

Examination Committee

Chairperson: Prof. Mónica Duarte Correia de Oliveira
Supervisor: Prof. Maria Isabel Craveiro Pedro
Member of the Committee: Prof. Andreia dos Santos Ramos Afonso de Matos Morais Nunes

October 2020

Declaration

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

Declaração

Declaro que o presente documento é um trabalho original da minha autoria e que cumpre todos os requisitos do Código de Conduta e Boas Práticas da Universidade de Lisboa.

Acknowledgments

With this thesis, I finish the Master in Industrial Engineering and Management, many years of dedication and perseverance that contributed to making this work possible. It is the end of a cycle that would not have been possible without the support of teachers, family and friends.

I want to acknowledge my dissertation supervisors Prof. Diogo Cunha Ferreira and Prof. Maria Isabel Pedro, for their insight, support, constant availability, and sharing of knowledge throughout the whole time I was writing my thesis. Their help and guidance made this thesis possible. I would like to thank FCT for its support to this thesis, under the project hSNS (PTDC/EGE-OGE/30546/2017).

I would also like to thank my family for their support and caring over all these years, for always being there for me and making it possible for me to pursue and accomplish my study objectives.

Last and surely not least, to all my friends that helped me grow as a person, and were always there for me during the good and bad times in my life. Particularly to Leonor for all her support and motivation to never give up throughout all these years without which this project would not be possible. Thank you.

To all of you – Thank you.

Abstract

In Portugal, public hospitals provide universal, general, and tendentiously free access to all Portuguese citizens, having only a handful of Public-Private Partnerships (PPPs). Since some significant reforms, which started in 2002, four hospitals were created under the PPP regime. The creation of those four hospitals presupposed a choice made in favour of another, being the other the establishment of regular public hospitals (EPE hospitals). Because of the entry, in Portugal, of private parties in the public health sector, the discussion on which one of these two models improves hospital performance increased. Consequently, a comparison between EPE and PPP hospitals is needed. The present document starts by presenting the Portuguese health care system. While focusing on secondary health care, the concept of PPP is brought up, and the four Portuguese PPPs are considered. A robust benchmarking tool is needed to compare the two groups of hospitals. This study uses Data Envelopment Analysis (DEA), which can be used to empirically measure hospital technical efficiency, and Malmquist index, a robust non-parametric index that can be applied to measure group performance, alongside recent data about Portuguese hospitals (FY2015-FY2019). The sample contains information about 28 hospitals, from which four are Public-Private Partnership (PPP)s. Four models were applied differentiated in two dimensions, desirability of variables and availability of information about PPP hospitals. The Data Envelopment Analysis (DEA) results show that PPP hospitals have the best average performance.

Keywords

Public-Private Partnerships; Public Hospitals; Healthcare Management; Data Envelopment Analysis; Benefit of Doubt; Malmquist index for group comparison.

Resumo

O serviço nacional de saúde português carateriza-se por ter acesso universal, geral e tendencialmente gratuito para todos os seus cidadãos, tendo no seu universo de hospitais apenas 4 em regime de Parceria Pública-Privada (PPP). A criação destes quatro hospitais ocorre no seguimento de reformas no sistema de saúde português influenciadas pela introdução do "*New Public Management*". A utilização deste modelo de gestão pressupõe uma preferência do mesmo sobre o modelo de gestão público (Entidade Pública Empresarial (EPE)). Com a criação destes hospitais em regime PPP em Portugal, surge a discussão sobre qual dos dois modelos beneficia o desempenho hospitalar. Consequentemente, é necessária uma comparação entre hospitais EPE e hospitais PPP. O presente trabalho começa por fazer uma apresentação do sistema de saúde português, focando os cuidados de saúde secundários. De seguida, aprofunda o conceito das PPPs, mencionando os quatro hospitais PPP existentes. De forma a fazer a comparação entre os dois modelos de hospitais é utilizada uma ferramenta de *benchmarking*. Este estudo usa *Data Envelopment Analysis* (DEA), usado para medir a eficiência de hospitais, e o índice de Malmquist, índice não-paramétrico que pode ser aplicado para medir o desempenho de grupos, juntamente com dados recentes de 28 hospitais portugueses (2015-2019), dos quais quatro são PPPs. Os resultados do DEA exprimem um melhor desempenho médio dos hospitais PPP.

Palavras Chave

Parceria Pública-Privada; Hospitais Públicos; Gestão dos Cuidados de Saúde; *Data Envelopment Analysis*; *Benefit of Doubt*; Malmquist para comparação de grupos.

Contents

List of Figures

List of Tables

Acronyms

ACSS Administração Central do Sistema de Saúde

BoD Benefit of Doubt

CMI Case-mix Index

CRS Constant Returns to Scale

CV Coefficient of Variation

DBFO Design, Built, Finance, Operate

DBFOT Design, Built, Finance, Operate, Transfer

DEA Data Envelopment Analysis

DMU Decision Making Unit

EPE Entidade Pública Empresarial (Public Enterprise)

ESS Espírito Santo Saúde

FTE Full Time Equivalent

GDP Gross Domestic Product

GPS Grupo Português de Saúde

HDG Homogeneous Diagnostic Groups

HPP Hospitais Privados Portugueses

hSNS Medição do desempenho dos hospitais portugueses através de métodos de análise multicritério

JMS José de Mello Saúde

KPI Key Performance Indicator

LSR Least Squares Regression

LSS List of Surgery Subscribers

MRTG Maximum Response Time Guaranteed

NHS National Health Service

NPM New Public Management

OLS Ordinary Least Squares

PCA Principal Component Analysis

PPP Public-Private Partnership

RHA Regional Health Administration

SA Sociedade Anónima (Public Limited Company)

SFA Stochastic Frontier Analysis

SPA Sector Público Administrativo

TFP Total Factor Productivity

UK United Kingdom

VRS Variable Returns to Scale

1 | Introduction

The Portuguese National Health Service (NHS) was created in 1979, bringing universal and free health care services to all citizens. This NHS is a Beveridge system, which means that health care is provided and financed by the government through tax payments [1]. It is safe to say that the health budget suffers much scrutiny since it is ”paid” by the citizens. However, investment in health cannot stop due to its importance in society. Expenditures related to the health care industry represent a substantial proportion of the Gross Domestic Product (GDP) [2].

In 2015 expenditures with health care represented approximately 9% of the Portuguese GDP, while in 2009 represented almost 10% (see Figure 1.1), being in that time above the European Union average of 9.5% [3]. The growth of expenditures in health, in the last decades, can be associated with the inefficiency of management in the health institutions [4], and also with the ageing of the population and technological developments. For that reason, there has been a growing interest in hospital efficiency analysis.

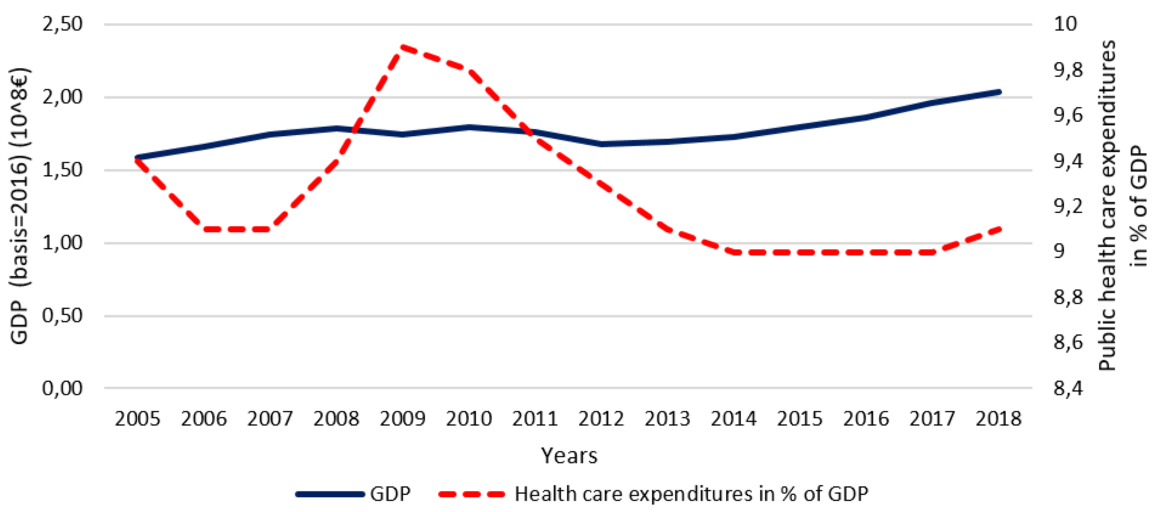

Figure 1.1: GDP and public health care expenditures in % of GDP evolution, FY2005 - FY2018 (Portugal) Source: PORDATA [3]

Throughout the years, traces of inefficiency and inadequate management have been constant in the NHS. To fight those traces, since the 80s the Portuguese NHS went through several periods of reform always focusing on improving hospital efficiency, introducing incentives to have better service quality and improving customer satisfaction [5].

The most notable reform period was the one that started in 2002 with the conversion of 31 hospitals belonging to the Administrative Public Sector to hospitals enterprises. The latter become under private law. Three years later, in 2005, those hospitals were all transformed into corporate public entities [1]. These transformations were driven by New Public Management (NPM), which was an effort to improve efficiency in the public sector by applying tools and ideas from private management [6]. After the implementation of NPM and the effects of *Governance*, which refers to a wide range of decisions carried out by the government or decision-makers to achieve national health policy objectives, the concept of Public-Private Partnership (PPP) emerged. Its popularity has been rising throughout the world ever since [7,8]. A PPP can be briefly described as a contract between a public entity and a private one who provides a service (being rewarded by the former), imposed by a set of conditions specified in a contract. These types of contracts were used in the health sector and, consequently, PPP hospitals were created.

The problem regarding PPPs in health, as Barros [9] referred, is the fact that both good results and problems have arisen. As a way of guaranteeing the provision of public services hospitals PPPs have led to several reflections on the conditions that favour their success. However, it is not clear what set of conditions is necessary or sufficient for this model to be successful.

In Portugal, PPPs in health care have emerged in four major hospitals: Hospital de Cascais, Hospital de Braga, Hospital de Vila Franca de Xira and Hospital Beatriz Ângelo - Loures. Knowing that the application of a PPP model can be seen as a double-edged sword, it is compulsory to study these particular examples. Try to compare them with classical public management to understand if it was the right decision using PPPs in health care and if it is a preferable option in the future.

1.1 Motivation

PPPs in the health care sector in the last years have experienced considerable growth [10]. In Portugal, four PPP hospitals were launched in what is called the first health care PPP wave. The Portuguese government opted for this NPM instrument to guarantee the provision of health services [11]. The choice for this form of execution presupposes a choice made in favour of another, being the other the traditional provision of the service. Some of the PPPs are facing the end of their contracts, and the discussion in whether or not to renew those contracts or even to create new ones is in the spotlight. Mainly because new polls are coming up and every political party defends different points of view about it.

However, empirical evidence of the benefits of using PPPs in health care is mixed [7]. This subject raises some concerns about having private companies operating and managing hospitals, which are structures with enormous social responsibility. Little attention has been paid in the literature regarding differences between EPE Hospitals and Public-Private Partnerships, especially in terms of the performance gap between them. The coexistence of both PPP and EPE hospital models demands in-depth analysis. Therefore, a reasoned decision needs to be reached about the suitability of each model. It is necessary to have unbiased information available and compare the performance of both models, to achieve such a decision. Hence, given the importance and the contemporary nature of the theme, it was considered relevant to evaluate if PPP Hospitals can outperform in terms of performance the corporatized ones.

1.2 Objectives

The primary goal of this master dissertation is to apply the most suitable benchmarking alternative to the most detailed set possible, in order to compare the two groups of hospitals and draw economic and financial lessons from it. The thesis objective can be simplified to the question made on the thesis title ”Do PPP hospitals outperform the corporatized ones?”.

Some intermediate steps must also be accomplished regarding the dissertation problem. For example, it is believed that there is a gap in the literature about the comparison of PPP and EPE hospitals' performance. Hence, this thesis will first, confirm or disprove the existence of such a literature gap. Then, will clarify some concepts that are directly related to the research problem such as the Portuguese NHS, PPPs and benchmarking models. The goal is to create a strong knowledge foundation for future work that is going to be done in the dissertation. The intermediate steps of this thesis can be stated as:

- Understand what the benefits and disadvantages that PPP can have when applied to the health sector are;
- Getting to know how the Portuguese NHS works, analyzing the four hospitals in Portugal that are PPPs;
- To do a deep literature review regarding PPPs in health care worldwide, and also in particular in Portugal ;
- To study potential and suitable benchmarking methodologies and determine the most adequate to use to compare the two groups of health care providers.
- Select the adequate variables to evaluate hospital performance, following criteria like availability, quality, and past similar researches;
- Determine which hospitals are going to comprise the sample, in order to have a robust database;
- Implement the methodology of this study with the creation of a suitable and adequate computational framework.

1.3 Thesis structure

The remainder of this thesis is structured as follows, visual aid provided by Figure 1.2:

Chapter 2 consists of the theoretical support, starting with the Portuguese NHS, discussing its history and evolution since its creation. The secondary health care is then approached, where the different types of hospital models are explained. The concept of PPP is also approached, discussing its benefits and risks, and different models. Then, the connection between the two topics is made.

Chapter 3 consists of a worldwide literature review of the main articles, papers, books and thesis that are relevant to the area of research of this thesis. A more thorough research is done in studies that consider the Portuguese reality.

Chapter 4 addresses the concept of benchmarking, studying the main models available to estimate the performance of homogeneous services. A more in-depth investigation is made on Data Envelopment Analysis (DEA) methodology and Malmquist indexes.

Chapter 5 is divided into two parts. The first part presents the data and variables of this study. It explains how data was obtained and treated and how the variables were selected. The second part explains the different models of this study, created to take full advantage of the database constructed, and how those models were implemented.

Chapter 6 presents the main results obtained in this study. It also provides a detailed analysis and discussion of the results.

Finally, Chapter 7 discusses and summarises the work made and the results obtained with this dissertation. Final remarks on the more relevant findings are provided, as well as limitations and future work.

Figure 1.2: Stages of the thesis
Source: Author

2 | The Portuguese National Health Service

Chapter 1 contextualizes the problem, presenting a brief description of the research motivations and what is aimed to be achieved. The present chapter introduces the reader to the Portuguese NHS, and is divided into six parts. The first part provides an overview of how the health care in Portugal is provided. The second part has a brief summary of the NHS history and explains its different levels of care. The following part describes the secondary care in Portugal, while also, explaining the different types of management of hospitals. In the forth part, an analysis of the concept of PPPs is given, studying its definition, pros and cons, and its different types of models. The fifth part links the concept of PPP with health care, describing the three models of PPPs hospitals. Lastly, the four Portuguese PPP hospitals, created in the first wave, are presented.

2.1 Health care in Portugal: Overview

Either public or private entities provide health care in Portugal. Public providers belong to the NHS, whereas private partners may have some partnerships with the public sector. Meanwhile, the financing of this system is supported by the State itself (through taxes), special social health insurance schemes (health subsystems, like Instituto Público de Gestão Participada - known as ADSE, that only apply to certain professions), private voluntary health insurance (such as Multicare, AdvanceCare or Medis), and by the citizen himself when he covers his expenses (*out of pocket*) [1].

The implementation and compliance with the policies proposed in the national health plan are carried out centrally by the Ministry of Health and the Central Administration of the Health System [12].

2.2 A brief summary of the NHS history

In 1974 a democratic revolution occurred in Portugal, putting an end to a dictatorship that lasted over 40 years. Two years later, the Portuguese Constitution was approved, being recognized citizens' right to health care by "the creation of a universal, free-of-charge National Health Service" [1].[1] On July 28, 1978, an order was published, known as "*Despacho Arnaut*" which was an anticipation of the NHS. "*Despacho Arnaut*" allowed access to all citizens to medical-social services, regardless of their ability to pay [13].

[1] Decree number 10/04 of 1976, article 64

Following the constitution and ”*Despacho Arnaut*”, in 1979, the law that enabled the right to health care was approved and, accordingly, the NHS was created.[2] The NHS guarantees universal, general and free health care services to all citizens. Universal access means access to all citizens regardless of their economic or social status (ability or willingness to pay), and general means to all areas and needs [14]. Although both of these last constitutionally principles (universality and generality) have prevailed to this day, the ”free-of-charge” principle did not. A revision of the constitution was made in 1989,[3] making the access ”tendentiously free” instead of free, which means the users of the NHS now have to pay fees [15].

According to the Ministry of Health [16], the NHS comprises health promotion and surveillance, disease prevention, patient diagnosis and treatment, and medical and social rehabilitation. Its objective is to protect the citizens and society from diseases. The NHS enjoys administrative and financial autonomy, and it is structured as a dispersing and decentralized system, comprising central, regional and local bodies.

The decentralization of the NHS occurred after some reform trends of other European countries [1], influenced by the management model of the NPM [15]. NPM regarded decentralization as an effective way of promoting efficiency, effectiveness, improving service delivery, having a better allocation of resources according to needs, involving the community in health decision-making, and reducing inequities in health [17]. This decentralization process started in the 90's with the publication of a law[4] referring that health authorities are now established at the national, regional and municipal level to meet the needs of citizens.

However, such a decentralization phenomenon was only put into practice in 1993[5] through the creation of five Regional Health Administration (RHA)s [17]:

- **North**, based in Porto - districts of Braga, Bragança, Porto, Viana do Castelo, and Vila Real;
- **Center**, based in Coimbra - districts of Aveiro, Castelo Branco, Coimbra, Guarda, Leiria and Viseu;
- **Lisbon and Tagus Valley**, based in Lisbon - district of Lisbon, Santarém and Setúbal;
- **Alentejo**, based in Évora - districts of Beja, Évora and Portalegre;
- **Algarve**, based in Faro - district of Faro.

According to that law, a RHA has *”legal personality, administrative and financial autonomy and own assets”*, and has the functions of *”planning, resource allocation, guidance and coordination of activities, human resources management, technical and administrative support and evaluation of the functioning of health care institutions and services”* .[6] A Board of Directors per RHA is appointed by the Minister of Health. The RHA proposes programs on which the financial plans of the health units are going to be

[2] Law number 56/79, 15th September
[3] Law number 1/89, 8th July
[4] Law number 48/90, 24th August
[5] Decree-Law number 11/93, 15th January
[6] See footnote 1

based while having into consideration the government's budgetary contribution. Thus, the creation of the RHAs gives them financial responsibility, as each one receives a budget to be used for the health assistance of its target population [15].

The NHS provides health care services at different levels, such as primary care, secondary care (or differentiated care), post-hospital care and rehabilitation (continued care), and palliative care (end-of-life care) [1].

Primary health care in Portugal consists of a network created by all RHAs. RHAs aids, simultaneously, in health and disease prevention, management of severe health situations according to physical, psychological, social and cultural dimensions [14]. Examples of primary care activities include prenatal care, childcare, medical care for the adult population, family planning and perinatal care, first aid, among many others. To sum up, primary activities cover all ambulatory health care that is given outside of hospitals [15].

Secondary health care, also known as hospital care since it is provided by hospitals and hospital centres, is the most specialized type of care in the NHS [14]. The public sector in Portugal provides secondary care through public hospitals. Services include hospitalization, emergency, surgery, other complementary means of diagnosis and therapy, medical appointments, day hospital treatments, and diagnosis [15]. Hospitals are unevenly distributed throughout the country, having into consideration population, health needs, and the presence of medical professionals [18].

The post-hospital rehabilitation care, known as "continued care", is a network that aims at trying to stabilize and guarantee the full physical recovery of a patient after hospitalization.[7] Patients may stay in these units from 30 days (short term) to more than 90 days (long term) [15].

Palliative care is specialized medical care to handle people living with serious illnesses. Its main goal is to improve the quality of life for both the patient and the family. Palliative treatments have no intention to cure the patient. Instead, they aim to decrease the effects of the disease on the overall well-being of the patient [1].

2.3 Secondary health care in Portugal: Hospitals

Secondary care, as said before, is mainly provided by the hospital's clinical staff. The report "*Estatísticas da Saúde* 2017" defines hospital as a health facility that provides medical and surgical treatment and nursing care for sick or injured people, and may contribute to disease prevention, scientific research, and teaching [19].

Hospitals are relevant in the health care system as their budgets represent a considerable share of the overall health budget. Furthermore, the policies adopted in the management of the hospital have a

[7] Decree-Law number 101/2006, 6th June

material impact on the NHS.

It is possible to differentiate hospitals into two different types, public and private. In a public hospital, its owner and main supporter is the state, which can provide universal or restricted access. A private hospital has a private entity as its owner and main supporter, for-profit or not, and may be of universal or restricted access.

In 2017, there were 225 hospitals in Portugal, which represents an increase of 27 hospitals when compared to 2007. The existing hospitals in 2017 were divided into 114 private hospitals (15 more than in 2007), 107 EPE hospitals and 4 PPP hospitals. Given that all PPP hospitals were also of universal access, the number of universal access hospitals per 100,000 inhabitants was 1.0 in 2017 [19].

2.3.1 Hospital management models

The Portuguese hospital system has suffered several changes and reforms throughout the years, especially in the last couple of decades, starting with the change of government after the elections of 2002. This new political cycle introduced corporate management practices into the public sector, aiming at increasing hospital efficiency as well as service quality [10].

Between 1979 and 2002, all public hospitals were under the management of the Administrative Public Sector (*Sector Público Administrativo*, SPA)[8] [6]. All hospitals were subjected to public/administrative law, being publicly managed and owned. Following the NPM ideology, which consisted in bringing and adopting management and organization principles from the private sector to the health system, and reinforcing the agreements with the private and social sectors [14], at the end of 2002, 31 out of the 34 traditional SPA hospitals have already been transformed into public limited companies or hospital enterprises (Sociedade Anónima (Public Limited Company) (SA)).[9] SA hospitals have limited liabilities and are subjected to commercial/private law [20].

In 2005, all 31 SA hospitals and 5 SPA hospitals were transformed into corporate public entities (Entidade Pública Empresarial (Public Enterprise) (EPE))[10] to maintain the unequivocal public nature of the hospitals, and enhancing the supervision and intervention of the Ministry of Health and Finance. Ten new EPE hospitals were created in 2007 [21]. SA and EPE management differs from public management because they have more autonomy in certain aspects. They do not follow a private management model because their autonomy is supervised. It is an intermediate situation regarding the type of management between a SPA hospital and a private hospital [22].

Meanwhile, it was announced by the government some hospital projects that would be launched under a PPP [23].[11] The Task Force for Partnerships in Health was created with its own specialized

[8] In order to simplify notation hospitals that belong to the public sector are going to be referred as Sector Público Administrativo (SPA) hospitals

[9] This occurred under the Law number 27/2002, 8th November

[10] Decree-Law number 93/2005, 7th June

[11] Decree-Law number 185/2002, 20th August

human resources as well as external consultants to be responsible for these PPPs. PPPs have private investment, public funding, private management, and public ownership. Under this PPP model, four hospitals were created, and some more were announced, but their creation was after dismissed [24].

With all the reforms from 2002 onwards, the health system integrated four different management models:

- SPA hospital
- SA hospital
- EPE hospital
- PPP hospital

A – SPA hospital Administrative Public Sector, also known as traditional management, was the most used management model in Portugal until 2002. All public hospitals that were not transformed into corporatized ones are SPA. The Ministry of Health supervises these entities in fiscal terms, having administrative authority over hospital management [6]. Public hospitals have limited administrative and financial autonomy, and their staff have state employee status.

According to Ferreira [5], SPA hospitals are characterized by an isolated management system where bad management is not condemned, creating a culture of indifference about the spending. Indeed, the lack of spending control is one of the central problems of the public sector. The lack of information systems and performance evaluation mechanisms show that hospitals are inadequate to meet the challenges. Hospitals need to be managed in terms of business management in order to achieve greater management efficiency [5].

B – SA hospital In companies operating under private/commercial law, the capital is shared among the shareholders. SA Hospitals are equivalent to private companies, although in this case, the state owns all the shares. That is why they are public enterprises with exclusively public capital. SA hospitals have financial and administrative autonomy. Contrary to SPA, they have more freedom in terms of contracting and acquisition of health equipment, drugs, and human resources.

Still, they are subjected to regulatory intervention by the Ministries of Health and Finance. They also have the advantage of being able to use private management tools [6]. This new model offers performance incentives and the possibility of indebtedness of up to 30% of its share capital [21].

C – EPE hospital According to the Observatório Português dos Sistemas de Saúde (Spring Report 2006) [23], there are numerous similarities and differences between the operating regimes of SA and EPE hospitals. In both cases, the capital is public and both hospitals are subject, in terms of supervision, to the Court of Auditors (*Tribunal de Contas*). The indebtedness for both schemes cannot exceed 30% of the share capital.

Corporate public entities were created to give lower autonomy to the EPE hospitals comparing with the SA ones [6]. EPEs are subjected to higher intervention and control from the government. For example, ministries of Health and Finance must approve activity reports and budgets, monitor the business plan, approve internal regulations, and deal with critical issues.

EPEs have a more entrepreneurial form of management that combines management efficiency with users' satisfaction. Pragmatically, as Ferreira and Cunha states [20], EPEs differ from SAs since SA hospitals are easier to be privatized. Therefore, in 2005, all SA hospitals were converted into EPE.

D – PPP hospital The concept of PPP will be discussed ahead in Section 2.4. Nonetheless, it is possible to understand the concept of a PPP hospital. According to "*Estatísticas da Saúde* 2017", the definition of PPP hospital is *"hospital where the state acts as the main supporter and whose management is controlled and carried out by a private entity through a contract established with the state, being of universal or private access"* [19].

PPP hospitals are privately managed while also having private ownership. However, when the contract ceases the hospital equipment reverts to the state. They are entities with financial, administrative and investment autonomy, resulting from contracts between the public and private administration [22].

The following Table 2.1 provides a summary of the four different types of hospitals.

Table 2.1: Comparative table between different types of hospitals[1].

	SPA Hospital	**SA/EPE Hospital**	**PPP Hospital**
Capital Structure	Public	State as only shareholder	Private[2]
Governance	Board of Directors	Board of Directors General Meeting Supervisory bodies	Board of Directors General Meeting Supervisory bodies
Funding	Public and Private	Public and Private	Public and Private
Internal Control	Meaningless	Business type	Business type
Indebtedness	No authorization for financial debt	Financial debt cannot exceed 30% of share capital	No restrictions
Investment Decisions	Limited autonomy	Extended autonomy	Total autonomy
Provisions	Limited autonomy	Extended autonomy	Total autonomy
Human Resources	Civil service regime (needs approval from authority) No incentives policy	Individual contracts Outsourcing Incentives policy	Outsourcing Incentives policy

1 - It is considered SA and EPE hospitals as one type because in the dimensions that they are compared they are the same
2 - Equipment reverts to the state at the end of the concession
Source: author

2.4 Public-Private Partnership

PPPs were created and used for the first time in the United Kingdom (UK) in the 1970s. They appeared as a way to undertake major public projects without the prerequisite of exclusive public funding, while also sharing the risk with the private sector. PPPs became very popular worldwide, especially in Europe, being used by the governments so they could have better funding to deliver better services and infrastructures in their respective countries [7].

There is not a textbook definition of PPP. Sometimes it is only referred to as a traditional project carried out by the public sector, as it can be defined as a simple contract between the private and public sector. However, as Reich explains in his book [25], a good PPP definition always has three points. First, the existence of at least one private for-profit organization and one not-for-profit or public organization. Second, both entities want the creation of social value. Third, both entities share efforts and gains. With this in mind, some PPP definitions follow:

- The World Bank Institute defines PPP as *"a long-term contract between a private party and a government agency, for providing a public asset or service, in which the private party bears significant risk and management responsibility"*. [26]
- According to the Green Paper of the European Community (COM327/2004) [27] PPPs can be defined as *"forms of cooperation between public authorities and the world of business which aim to ensure the funding, construction, renovation, management or maintenance of an infrastructure or the provision of a service"*.
- In Portugal, the government, created a decree-law[12] where it is possible to find the definition of PPP: *"contract or the union of contracts whereby private entities, designated as private partners, undertake, on a long-term basis, towards a public partner, to ensure the performance of an activity aimed at the satisfaction of a collective need, and assume responsibility, in whole or in part, for the financing and operation thereof"*.

It is possible to identify a common feature in the definitions of PPP above, which is the existence of a contract between the public partner and the private partner. The private sector provides a service, following a set of conditions imposed by the contract, for which it is rewarded by the contracting public entity. Hence, the State is designated as a grantor and the private company as dealer [28].

It is possible to say that PPPs follow five main principles [29]:

- Orientation towards the satisfaction of collective needs;
- Include long-term relationships;
- Involve full or partial project funding;
- Are results oriented;
- The private partner must assume a significant portion of the risks.

[12] Decree-Law number 86/2003, 26th April (amended by Decree-Law number 111/2012, 23th May)

2.4.1 Benefits and risks of PPPs

The development of PPPs was fueled by the recognition that, despite the need for the state intervention in the provision of public services, they can be effectively provided by private entities, thus exploiting the benefits of private management in the public goods or service sectors [10].

The option of going for a PPP over a public model is based on the expectancy of some potential advantages such as risk-sharing and transfer of responsibilities [1]. There is a higher flexibility when it comes to raising funds which leads to a higher number of infrastructures built [30]. The state can achieve lower overall costs resulting from more innovative solutions and efficiency and effectiveness gains achieved by giving the responsibility of the construction to the private company [30].

However, there are also some drawbacks when using a PPP such as the contracts must be very clearly defined so both parties have their role in the partnership well defined, higher starting costs (preparation, study, evaluation and negotiation phases cost money) [30]. Lack of know-how by the state can be reflected in the negotiation phase [24], which is also a consequence of deficient existence of literature that could explain how to proceed in those negotiations [7]. Another problem is the temptation of non-inclusion of some items of public expenditures in the Balance Sheets [24] (*"Desorçamentação"*), which can lead to future financial limitations, burdening the State budget and causing sustainability problems on public finances [28]. Although there is momentary relief in the state budget, it is "offset" in the future by a spending commitment.

Shortly, for every project, it has to be assessed whether the PPP brings real added value (value for money) over alternative public procurement options [30].

2.4.2 PPP models

It is possible to classify PPPs in several ways. The Green Paper of the European Community (COM327/2004) [27] labels them in contractual type and institutionalized type.

- **Contractual type**: the public-private partnership is exclusively based on contractual relations such as conception, financing, accomplishment, renovation, or exploitation of a work or service.
- **Institutionalized type**: involves public-private cooperation in a separate new entity, called Special Purpose Vehicle (SPV), to guarantee the delivery of work or the provision of service for the public benefit [29].

The contractual type has several different models that can be classified accordingly to their goal (see Figure 2.1). The following models are worth mentioning: BOM (build, own, maintain), BOO (build, own, operate), BDO (build, develop, operate), DCMF (design, construct, manage, finance), DBO (design, build, operate), DBFO (design, build, finance, operate), DBFOT (design, build, finance, operate, transfer), BBO (buy, built, operate), LDO (lease, develop, operate), BOT (build, operate, transfer), BOOT (build, own, operate, transfer), BROT (build, rent, own, transfer), BTO (built, transfer, operate).

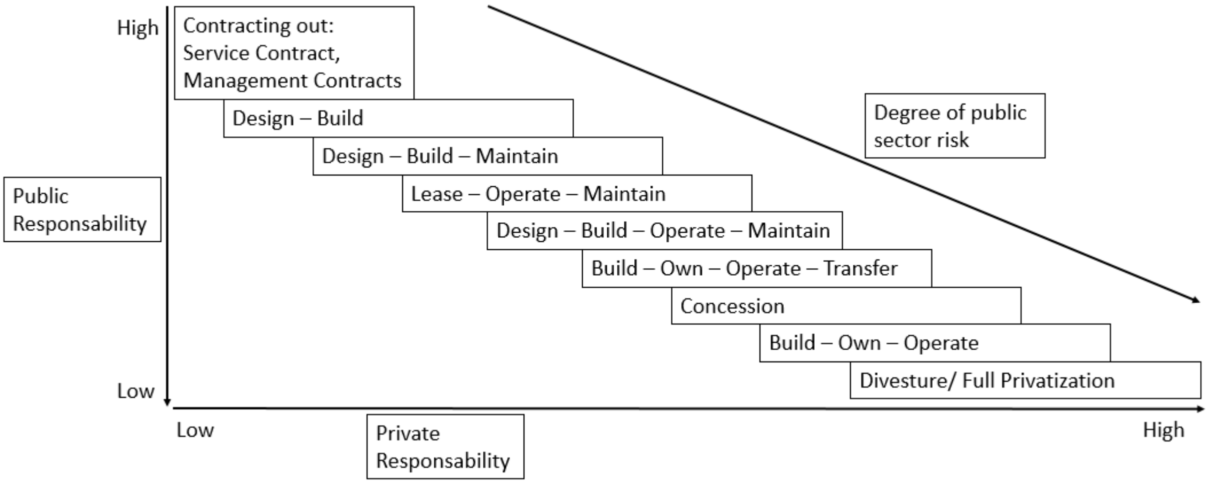

Figure 2.1: Scale and scope of private and public responsability
Source: Roehrich *et al.* [7]

Standing out from all of these models is the DBFOT, notable for its popularity and well-known adherence. It is the predominant model in Portugal [31]. It is a contract between the State and a private entity, where the latter is responsible for the construction, financing, operation, maintenance of infrastructure and provision of health services.

The DBFOT agreement is a long term contract that usually lasts for 25 or 30 years. At the end of that period, ownership is fully transferred to the State (this is what differences it from the DBFO model).

Generically, this model has several advantages, like using the private sector investment capability to fulfil the public needs. Also, by including private operators from the beginning of the process of programming and construction of the units, positive effects on costs and execution time of the work will appear. Operating costs also become more predictable and controllable [32]. Another advantage is that most of the responsibilities and risks land on the private entity, especially the financing risk, which last until the end of the agreement [33]. On the opposite side, there may be a loss of public control over the quality of the service provided, due to inadequate accountability mechanisms of the private partner [34].

2.5 PPP in health care

In the provision of social services, such as health care, the public sector takes a predominant role. Although, when the private sector starts to have a share in that provision, it is commonly referred to as privatization. In privatization, there is an irreversible transfer of ownership to the private sector [35] having reserved the role to regulate the market to the public sector. That is not the case of PPP arrangements. In these cases, there is only a temporary or partial transfer of ownership. This model is the most used around the world, when the private sector is involved in the health care system. Despite the clinical and/or infrastructure management being a responsibility of the private entity, the customer only has to pay a fee, as if it was a typical public hospital (EPE hospital) [33].

PPP projects have flourished because they are an alternative solution to the problems of budgetary constraints in the public sector [29] since they are a practical way of securing significant investments. There is the belief that, by using this model, both efficiency and effectiveness can be improved, and overall costs in construction and operation of hospitals could decrease [36].

2.5.1 PPP models in health care

PPPs in health care can be used in the hospital sector since hospitals consume about half of the health sector budget. The different PPP projects in this sector are differentiated based on the group of activities included in the contract. Arrangements can go from just considering the management of the hospital infrastructure by a private entity to having a full-service provision at all levels of care [37]. As it can be seen in Figure 2.2, in the health sector there are essentially three models: United Kingdom (UK) model (model 1), Portuguese model (model 2) and Alzira model (model 3) [36].

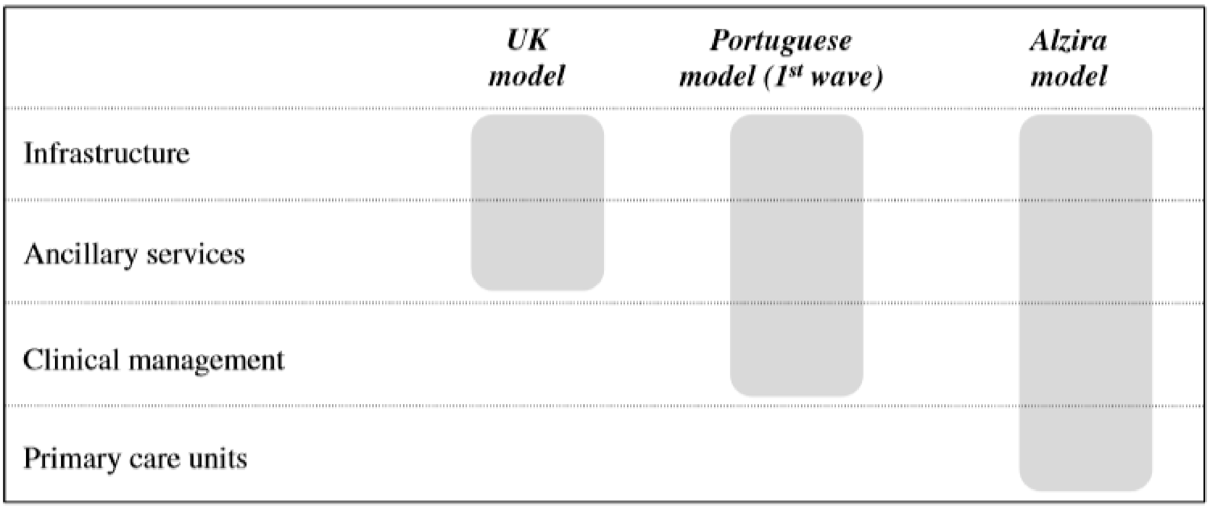

Figure 2.2: PPP models for health care.
Source: Cruz and Marques [36]

2.5.1.A UK model

The UK model is a partnership made to create/restructure and operate hospital infrastructure. It is not linked with hospital and clinical management. This model is a typical DBFOT contract and regards design, construction, financing, operation, and transfer.

In some contracts, new infrastructure is built, while in other refurbishment work and capacity increase is applied. The private partner is responsible for the management of the infrastructure (the building itself and all of the systems that guarantee basic conditions, such as air conditioning, elevators, ventilation, water and energy systems), and of the ancillary services [36]. In the contract, the provision of soft facilities (cleaning, security, parking, catering, maintenance and management of some sophisticated medical equipment) can also be included [37].

As said before, this model can build or restructure the hospital depending on the option chosen. Based on this, there are two different classifications [33]: greenfield, if the structure is built from scratch, and brownfield, for infrastructure restructuring, rehabilitation and/or expansion. At the end of the contract period, usually 25 to 30 years, the assets are reversible to the State. The State finances the partnership during the contract according to the availability of the contracted services and performance.

The adoption of this model has been a common practice in many European Union countries, including France, Germany, Italy, Greece, Ireland and, of course, the UK, and also in Canada, Japan, South Africa and Australia [33]. In Portugal, it is also predicted that the second wave of PPP hospitals will follow the UK model. The wide acceptance of this model has to do with not including clinical management by the private sector, leaving the hospital's core business to the public sector [34].

2.5.1.B Portuguese model

Contracts following the Portuguese model include infrastructure and soft facilities services as the UK one, but also clinical and medical activities. Therefore, it is the responsibility of the private contractor to recruit, train and manage clinical staff,[13] as well as to maintain, operate and purchase medical supplies for the [10].

The concessionaire designated by the PPP is responsible for the management of the hospital [33]. The concessionaire is usually a consortium of companies because each company will be responsible for different project components. For example, a construction company to build the infrastructure, health services management companies to manage the hospital, and banks to finance the investments.

The integration of infrastructure, management, and clinical services in the same institution is expected to increase efficiency [36]. Having in the consortium, both the infrastructure planner and the designer of future clinical services gives more flexibility to the architectural and engineering infrastructure designs.

The Portuguese model (model 2) was applied in all of the four hospitals of the first health care PPP wave.[14]

[13] Although, if this model is applied to a running hospital (substitution hospital), the clinical staff as to be maintained under the new management (brownfield).

[14] Currently some of them have or are scheduled to change the model type

2.5.1.C Alzira model

The Alzira model is characterized by having full-service provision at all levels of care. It is not limited to the hospital perimeter, as it also extends to primary care centres [37].

This model is based on the idea that it is possible to control the flow of patients from primary care facilities to hospitals. This flow can be responsible for inefficiency in the health network if it is not controlled and coordinated.

The goal of this model is to have the concessionaire managing the hospital's resources, making sure that only predetermined cases are directed to the hospital, and less urgent ones are treated at health centres. By doing so, there is an increase in efficiency in using specialized facilities.

An example of this model is the Hospital de La Ribera, Valencia (Alzira), Spain [38].

2.6 PPP hospitals in Portugal

The first experience of implementing a private management model in a public hospital began in 1995 with the signing of a contract for the private management of a general hospital, Hospital Fernando Fonseca [29], in Amadora, a district hospital with 670 beds and integrated into the NHS. This contract was signed by the RHA of Lisbon and Tagus Valley and the private consortium - Hospital Amadora/Sintra, Sociedade Gestora, S.A. [34]. By government decision, this partnership ended in 2008 and was turned into an EPE hospital. Since 2001, this kind of contracts in health care became more frequent, being announced in that year the first wave of hospitals in a PPP regime. Also, in 2001, the Task Force for Health Partnerships was created, to control and be responsible for PPPs.

Health PPPs are divided into the first wave and second wave PPP programs. According to the Ministry of Finance, there are two main differences between them, while the first wave concerns the construction and maintenance of infrastructure, as well as the management of the health facility, the second wave covers only the construction project and the maintenance of the building [39].

The first wave of hospitals consists of two new hospitals (Sintra and Loures) and three replacement hospitals (Cascais, Braga and Vila Franca de Xira) [24], nonetheless the Sintra Hospital never came to fruition. A new hospital means the PPP aims to build a hospital unit from scratch, targeting a population that was integrated into the area of influence of another hospital unit, without closing any physical structure. A replacement hospital closes existing physical units, replacing them with new infrastructures [33].

The first wave is based on a management contract with two managing entities, one for building management and another for the clinical services component. Table 2.2 briefly presents the main features of the first wave of hospital PPPs.

Two contracts with different durations have been signed in the first wave. When those contracts come to an end, a new tender is launched to create a new one. The clinical services contract can be

Table 2.2: Main features of the Portuguese first wave of hospital PPPs.

	Private partner responsible for the infrastructure management	**Private partner responsible for the clinical management**
Model	DBFOT, including clinical services	
Contract Issues	To design, build, and preserve (manage) the infrastructure and hard facilities (ancillary services)	To manage clinical staff and deliver health care to all citizens
Duration	30 years	10 years
Responsible	Hospital building management company	Hospital facility management company
Activity	Design, construction and maintenance of hospital facilities, hard facilities management, and heavy fixed equipment	Hospital management, clinical service management, soft facilities management, and mobile equipment
Payment	Based on services availability (penalties for service failures)	Based on contracted production (penalties for low quality)

Source: author. Adapted from Tribunal de Contas [24] and Cruz and Marques [29, 36]

renegotiated every ten years, which means that the private sector cannot adopt a quiet life during the contract period, bringing the benefit of competition since poor performers will be replaced [9, 36].

The first wave had both clinical exploration and construction/maintenance of infrastructure, included in the partnership. However, those kinds of partnerships ceased in 2008. The second wave has only the private partner allocated to the construction/maintenance of infrastructures [33]. Hospital de Braga, since September 2019, no longer has the private entity responsible for clinical exploration and starting in 2020, Hospital de Vila Franca de Xira will no longer have it as well.

Launched in 2006, the second wave of PPPs is similar to model 1 (UK model). Different from the first wave, establishing only an arrangement for the infrastructure management and ancillary services, leaving clinical management to the responsibility of the public sector. Initially, this wave included the hospitals of Faro, Évora, Guarda, Póvoa de Varzim/Vila do Conde, and Vila Nova de Gaia, all of them as replacement hospitals. However, after that, it was defined[15] the construction of the following units: hospital of Lisboa Oriental, Faro, Seixal, Évora, and hospital centre of Vila Nova de Gaia and Póvoa de Varzim/Vila do Conde.

None of these projects came to fruition, probably due to the financial crisis that affected Portugal in the early 2010s. However, it is predicted the construction of Hospital de Lisboa Oriental and, eventually, Hospital Central do Algarve. Therefore only four hospitals were created by the first wave of hospitals.

[15] Order No.12 891/2006, 21th June

2.6.1 Hospital de Loures

The creation of this PPP hospital involved two phases: the first one in December of 2003 to be cancelled in 2006, and the second one in early 2007 leading to the opening of a new hospital in February 2012.

The first proposal, for the construction of a new hospital with capacity for 650 beds, was cancelled because of serious flaws in the evaluation of the four candidates' bids. They were not in the public's best interest. Also, the integrity of the competitors' financial structure and the answers they gave were not satisfactory [24]. This tender attracted four competitors, consortia led by Espírito Santo Saúde (ESS), Hospitais Privados Portugueses (HPP), José de Mello Saúde (JMS) and Misericórdia do Porto.

A new tender was launched after the failure of the first one, also for the construction of a new hospital, now for 565 beds. Only two private entities applied: ESS and JMS, the latter won the tender [33]. HL – Sociedade Gestora do Edifício, is responsible for the management of the building, whereas the provision of health care services are provided by SGHL – Sociedade Gestora do Hospital de Loures [40].

2.6.2 Hospital de Cascais

For this tender, launched in May 2004, four consortia competed: ESS, HPP, JMS e Grupo Português de Saúde (GPS). HPP was the winner. Lusíadas Saúde - Parcerias Cascais, is responsible for the hospital establishment and providing health care, while TDHOSP – Gestão de Edifício Hospitalar, (Teixeira Duarte), is responsible for the building management [41]. The contract was only signed in 2008 [1] and consisted of the design, financing, construction, maintenance, and management of the Hospital de Cascais (replacement hospital) and for the management of the Cascais Hospital Center during the construction of the new hospital.

Hospital de Cascais was the first hospital in the NHS to work under a PPP regime. This hospital replaced the Cascais Hospital Center and covers the municipality of the same name, as well as eight parishes of the municipality of Sintra, serving a total of 285,000 citizens. The hospital started its operations in February 2010 with an installed capacity of 265 beds [42].

2.6.3 Hospital de Braga

A tender was launched in November 2004 for the construction of a new hospital with a capacity for 780 beds, and also for the management of the old hospital while the new one was under construction. This contract has the same goals and requirements as the one for Hospital de Cascais since it also involves a replacement hospital [33]. This new hospital would replace the former Hospital de São Marcos. Hospital de Braga covers the districts of Braga and Viana do Castelo.

With six competitors, among which four consortia that also competed for the Hospital de Cascais, the two best offers were made by ESS and the Escala Braga/JMS consortium. The latter was the winner of this tender. Escala de Braga – Sociedade Gestora do Edifício is responsible for the management of the building, whereas the provision of health care services is provided by Escala Braga – Sociedade Gestora do Estabelecimento (with a large share from JMS group) [43]. The management of the former hospital started in September 2009, whereas the new hospital only came into operation in May 2011.

2.6.4 Hospital de Vila Franca de Xira

Launched at the end of 2005, this tender intended to replace the old hospital, Hospital Reynaldo dos Santos. The project started with managing the former infrastructure while in parallel, the new hospital infrastructure was being built, for a capacity for 520 beds. Serving around 215,000 residents at launching [44].

The four consortia that also applied for the Hospital de Cascais presented a proposal, having the group JMS won the tender. The contract, signed in October of 2010, has Escala Vila Franca – Sociedade Gestora do Estabelecimento responsible for health care provision, and Escala Vila Franca, Sociedade Gestora do Edifício (with a large share from JMS group) responsible for the management of the building. Hospital Vila Franca de Xira went into operation in March 2013 [45].

2.7 Summary of chapter 2

This chapter introduced the reader to the Portuguese NHS. It was shown that the NHS provides health care services at different levels, belonging the hospital care to the secondary level. As of 2017, Portugal secondary care was comprised of 225 hospitals (107 EPE and 4 PPP hospitals). The definition of PPP is given as a contract where the state is the main supporter, but the management is controlled and made by a private entity through a contract established between both. This types of contracts can be applied to provide health care, for example, in the hospital sector. Three models can be used to implement a PPP hospital: UK model, Portuguese model, and Alzira model. Portugal's first wave of PPP hospitals resulted in 4 units: Hospital de Braga, Hospital de Cascais, Hospital de Loures, and Hospital de Vila Franca de Xira.

The next chapter is dedicated to studying aspects related to the literature review relevant to the topic under investigation.

3 | Literature review

This chapter addresses literature on the more technical aspects of the research. The first section does a review on international studies that evaluated PPP arrangements in the health sector, and the second section does similar research, although focusing on the Portuguese literature. The third section compiles works that do not consider PPP specifically but evaluate hospital performance of other types of hospitals using benchmarking tools, like DEA or similar. The last section is the result of an extensive literature review on the variables used in the studies that use DEA to compare performances between hospitals.

3.1 PPP hospitals in the world

Research and evaluation of performance in PPP arrangements have been a constant concern for health researchers, and several studies have been performed regarding this topic.

At the international level, there are some noteworthy works done, including one of Torchia *et al.* [46]. This global study addressed the issues of governance in health PPPs, conducting a systematic review of forty-six articles published in peer-reviewed journals from 1990 to 2011.

Drawing on the experience of countries such as Australia, Spain, and the United Kingdom, Mckee *et al.* [47] reviews the experience with variants of the PPP model in health care. The authors could not reach a consensus on whether the model was on itself flawed or whether the difficulties that appeared were the result of mistakes made during the execution of the model. Concluding, the authors refer that, this uncertainty surrounding the value of PPPs in health care needs urgent resolution

Barbetta *et al.* [48] used DEA to identify differences between public and private hospitals in Italy. The authors concluded that the differences between models resulted more from institutional settings in which they operate than an effect of the incentive structures embedded in the different proprietary forms. Barretta and Ruggiero [49] also study the PPP reality in the Italian health care.

Roerich *et al.* [7] analyzed over 1,400 publications about PPPs in health care considering a 20-year period. This study illustrates that further empirical research needs to explore pieces of evidence on the limitations of PPP arrangements in delivering public sector infrastructure and services.

Caballer-Tarazona *et al.* [38] evaluated the efficiency of the PPP experience in five Spanish hospitals to identify the influence of private management in the hospital outcomes. The authors concluded that the PPP group obtains good results, but not always better than those publicly managed. Acerete *et al.* [50] and Alonso *et al.* [51] also studied the Spanish health care PPPs.

Kruse *et al.* [52] compared the performance of public hospitals with the private counterparts regarding efficiency, accessibility, and quality of care in the European Union. The authors concluded that the private hospital sector seems to react more strongly to (financial) incentives than other provider types.

Some other works, that study PPPs in the health care sector in other countries, that are worth mentioning, such as, Sussex [53] and Waring *et al.* [54] in the UK, La Forgia and Harding [55] in Brazil, English [56] in Australia, Lim [57] in Singapore, Barlow *et al.* [58] includes several Europen countries in their work, and Hodge and Greve [59] and Buse and Harmer [8] which are global studies.

From the previous review, it is seen that there is an extensive literature regarding PPP hospitals. Nonetheless, only a few works assess and compare the performance of PPP hospitals against other secondary health providers. There is a trend in the conclusions from those works, either there are no relevant differences in performance between EPE and PPP managed hospitals, or PPPs show better results.

3.2 PPP hospitals in Portugal

Nonetheless, since the objective in this case study is focused on the Portuguese reality, a more thorough review is done on the literature that encompasses Portuguese hospitals. Table 3.1 summarizes the most important works studying the reality of hospital PPPs in Portugal.

From that table, some noteworthy works, such as, Ferreira [33] who seeks to understand if the existent hospital PPPs in Portugal have given rise to value gains for the financing entity (the State) or, in opposition, have led to inefficient management of the financed resources. As a result, he concludes that, in order to achieve a successful PPP, improvements are needed in decision making, budgeting, transparency, accountability and participation, and increased state bargaining capacity.

Entidade Reguladora da Saúde [60] a study requested by the Ministry of Health, evaluated four aspects to compare PPP hospitals with the EPE hospitals: technical efficiency, effectiveness, clinical quality and, regulatory costs. Regarding technical efficiency, measured with the DEA technology, it was found that PPP hospitals are globally efficient. However, no statistically significant differences were found. In terms of effectiveness, PPP hospitals presented a better relative performance in some indicators when compared to the non-PPP group average, but also showed the worst performance in others. Regarding quality, the study found that PPP hospitals have favourable results, on average, when compared to EPE hospitals. In the end, no overall conclusion is drawn on the advantage or disadvantage of PPPs.

Nunes and Matos [61] studied the performances of hospitals in a PPP regime, designing a benchmarking exercise with the DEA for the years 2013, 2014 and 2015. As results obtained that from the four hospitals analyzed, three of them were efficient in the three studied years, being only one considered inefficient. The aim of implementing PPPs in hospitals to improve hospital efficiency and productivity, according to the results of this study, was in part achieved. Most of the PPP hospitals were efficient when compared with themselves and with other hospitals of the NHS.

Ferreira and Marques [37] use Benefit of Doubt (BoD) alongside recent data about Portuguese hospitals (from FY2012-FY2017) to evaluate if PPP hospitals can deliver health care services with social performance levels at least as good as EPE hospitals. The authors conclude that PPPs are not expected to have lower performance levels when comparing to EPE hospitals, although there may be an interface problem leading to a potential conflict of interests between the two leading players.

From this last review, similar lessons to the ones obtained in the previous section can be retrieved. In Portugal, a few works are studying and comparing the performance of PPP hospitals against the EPE ones. In general, those works conclude that most PPP hospitals have good performance levels and do not tend to have lower performance levels when comparing to EPE hospitals. The question with these works is that they only consider one specific dimension in their analysis, either consider quality or efficiency dimensions, for example. Because a study considering a broader range of dimensions, to obtain more complete performance measurements of hospitals, is needed, this study is deemed to be compulsory.

3.3 Review on methodologies

Table 3.2 compiles some works that, although not using PPP hospitals into the analysis, evaluates and compares performances of different types of hospitals, in Portugal, using DEA or similar methodologies.

From this table, it is important to stress out some works such as Gonçalves [22], that measures the efficiency of SA and SPA hospitals using both parametric and non-parametric methodologies, Stochastic Frontier Analysis (SFA) and DEA, respectively. This study concluded that hospital corporations helped to improve the efficiency frontier of the overall hospitals. Nonetheless, EPE hospitals could show higher technical efficiency scores than those that were transformed into corporations.

Rego *et al.* [62] tried to understand to what extent private entities in public health care could improve hospital technical efficiency. By applying DEA to data from 2002 to 2004, concluded that introducing business management models to the public sector had a positive impact in Portuguese public hospitals. Noticing a particular impact in SA hospitals, although further studies are requested by the authors.

Ferreira and Marques [6] evaluated the efficiency of corporatized hospitals using DEA and Malmquist index for group comparison. The authors concluded that efficiency increased, albeit slightly when compared to the period before corporatization.

Nunes [63] evaluated hospital efficiency after the reforms made in 2002, using DEA and Malmquist index. In his conclusions, the author states that, in the short term, there is a slight improvement and that over the years, the positive effects are beginning to be seen with more than half of public hospitals with high efficiency scores.

Ferreira and Nunes [15] use DEA to measure and compare the use of inputs (resources) and the outputs (patient's discharge for example) produced by hospital unit in each RHA, in the year 2017. This study points out the existence of inequalities on a regional comparison, suggesting special attention from policymakers and hospitals managers to this diversity in interregional scores.

3.4 Inputs and outputs

This section provides a literature review on the central studies that are similar to this research, to find which variables are consensual in the literature. To the works from the previous tables, more works were considered to have a more significant sample size [2, 64–70], totalling for this review 28 different works. The following Figure 3.1 demonstrates the results obtained. Only the top 10 most used variables are shown, due to the high number of total variables (32 different inputs and 22 different outputs found). Sometimes variables had different names in various studies but regarded the same information, in those situations they were all clustered under the same variable.

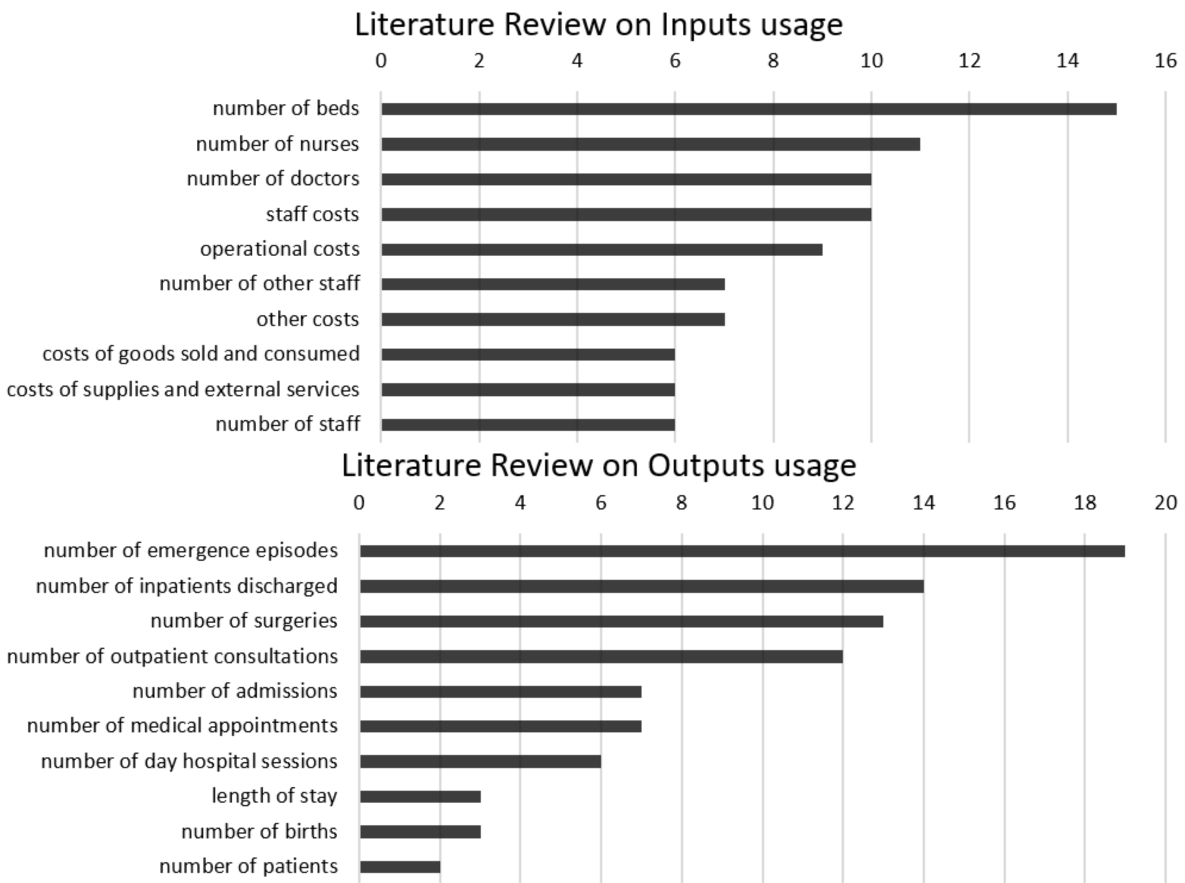

Figure 3.1: Inputs and outputs utilization in similar DEA studies.
Source: Author

Table 3.1: Literature review on PPPs in health care.

Author (Year)	Description	Methodology	Data Range	Main Conclusions
Tribunal de Contas (2009) [24]	Audit of the first wave of hospitals belonging to the PPP health program.	-	2003-2008	The audit recommends corrective measures to be instructed to the State, through the Ministry of Health.
Cruz and Marques (2013) [36]	Tries to overcome the gap in the literature regarding the review of PPP arrangements in a cross-country perspective.	-	-	It is the authors' opinion that vertical integration should be maintained as a PPP alternative.
Oliveira (2015) [44]	A statistical analysis is carried out, and a model is estimated to assess the impact of PPPs in the health sector.	Multiple linear regression	2009-2014	PPP evidence obtained is mixed: in some cases PPPs perform better than EPE hospitals but in other cases the reverse happens.
Ferreira (2016) [33]	Evaluation of positive and negative aspects of PPPs in the health sector.	-	2011-2015	Most significant flaw is the lack of mechanisms to specify boundaries of renegotiation capacity.
Entidade Reguladora da Saúde (2016) [60]	Assessment of the hospital management under PPPs, to identify "what were the positive things they brought to the NHS, while also identifying what comparative disadvantages they have.	DEA	2012-2015	No evidence found that PPP can lead to better or worse efficiency when comparing with other hospitals.
Nunes and Matos (2017) [61]	Evaluates the performance of PPP hospitals in the Portuguese NHS.	DEA	2013-2016	Three out of four PPP hospitals were efficient, so PPP hospitals are not always efficient, contrary to expectations.
Ferreira and Marques (2020) [37]	Demystifies the idea that PPP hospitals cannot deliver health care services with social performance levels at least as good as EPE hospitals.	BoD	2012-2017	No evidence supporting that EPE hospitals are better than the PPP ones.

Source: author

Table 3.2: Literature review on studies that use DEA or similar.

Author (Year)	Description	Methodology	Data Range	Main Conclusions
Tribunal de Contas (2006) [71]	Measures efficiency of hospitals that belonged to the SPA and were transformed in SA hospitals in 2002.	DEA	2001-2004	The adoption of the corporatized model did not result in efficiency losses, actually having, overall speaking, a relative increase in efficiency.
Barros *et al.* (2007) [72]	Analyzes the efficiency and productivity growth of a representative sample of Portuguese hospitals.	DEA Luenberger productivity indicator	1997-2004	Results show that, on average, Portuguese hospitals did not experience productivity growth during the period analyzed.
Giraldes (2007) [73]	Aims to evaluate the efficiency and quality of hospitals publicly owned with private management and hospitals of the public sector.	Aggregated efficiency and Quality ratings	2004	Concluded that corporate hospitals (EPE) were less effective in 2003, but in 2004 had the best efficiency scores.
Gonçalves (2008) [22]	Measurement of efficiency of SA and SPA hospitals using an efficient frontier approach.	DEA, SFA, Malmquist Index	2002-2004	Hospital corporations helped to improve the efficiency frontier of the entire hospital industry.
Moreira (2008) [74]	Comparison is made between the performances of the EPE hospitals and a control group composed of hospitals still within general government (SPA hospitals).	DEA	2001-2005	EPE hospitals, starting from a worse relative position, achieved relatively higher efficiency levels (significant statistically) when compared to SPA hospitals.
Simões and Marques (2009) [75]	Assesses the performance of the Portuguese hospitals and particularly the contribution of the congestion effect.	DEA	2005	DEA results seem to point out that SPA Hospitals had better performances than EPE hospitals.
Rego *et al.* (2010) [62]	The aim was to evaluate the impact of business management in Portuguese public hospitals with regards to efficiency.	DEA	2002-2004	Results suggest that the corporatization of hospitals have had a positive impact on public hospitals, particularly in SA hospitals.
Harfouche (2010) [76]	Two Health Policy Options were studied. An evaluation on the potential effects induced by both Options on the hospital efficiency was made.	DEA	2002-2008	Both Policy Options induced efficiency gains in corporate hospitals, either individually or by merger. Also noteworthy is the sharp decrease in hospital technical efficiency levels in 2004.

Source: author

Table 3.2: Literature review on studies that use DEA or similar (continued).

Author (Year)	Description	Methodology	Data Range	Main Conclusions
Ferreira and Marques (2015) [6]	Aims to investigate if the market structure reforms in the Portuguese health system have improved hospital performance and productivity.	DEA (order-m) Malmquist Index for group comparison	2002-2009	The more autonomy the hospital has from the Ministry of Health, the lower its productivity.
Nunes (2016) [63]	Analyzes the behavior and effects of corporatization in the efficiency of public hospitals, and in the efficiency of the hospitals after the intervention of the austerity program.	DEA Malmquist Index	2002-2013	Efficiency in business units only began to increase starting from the year 2005 (which corresponds to the transition to the EPE status), and the policy of merging health units did not have the expected outcome in terms of productivity.
Ferreira and Marques (2016) [20]	Discusses the importance of the Case-mix Index (CMI) to homogenize the sample of inpatient discharges.	Locally convex order-m method	2002-2009	The CMI shifts the efficiency frontier, but not the relative position of units against it.
Nunes (2017) [77]	Evaluates the efficiency variation of hospitals corporatization process in Portugal.	DEA	2002-2013	The adoption of business management improved the performance of hospital units over those that remained in the Public Administrative Sector.
Nunes (2017) [78]	Evaluation of the technical efficiency in nine Portuguese hospitals, which have undergone a change in their legal format.	DEA	2002-2013	The corporatization process of public hospitals generate efficiency gains in the medium/long term in eight of nine decision making units.
Nunes and Ferreira (2018) [17]	Compares efficiency scores in hospital units in the five RHAs.	DEA	2017	Portuguese public hospitals exhibit a considerable average performance.
Nunes *et al.* (2019) [79]	Econometric analysis to investigate the evolution of hospitals' performance.	DEA Hicks-Moorsteen productivity index	2013-2016	Since the financial crisis is still very recent, and no good system was monitoring the impact of austerity in the population, it was not possible to make concrete evaluations.
Ferreira and Nunes (2019) [15]	Reviews the process of decentralization of health in Portugal as well as the hospitals' corporatization and merging reforms.	DEA	2017	Diversity in interregional scores deserves special attention from the policymakers and hospital managers.

Source: author

3.5 Summary of chapter 3

This chapter did a profound literature review, and from the first two sections, it is learned that the concept pf PPPs in health care is an increasing subject of investigation around the world, with some studies already done in Portugal. Although several authors have studied Portuguese PPP hospitals (Figure 3.1), there are no efficiency studies that consider a comparative analysis between EPE and PPP hospitals. For this reason, the present work is innovative and will allow the development of future studies. Regarding the methodologies used to evaluate hospital efficiencies (Figure 3.2) it is seen that the majority of studies uses DEA or some particular case of it. The Malmquist index is also often used complementing the DEA methodology. This review is essential for the further research methodology made in Chapter 4. The last section results will help in the selection of variables needed to evaluate hospital performance, that is further addressed in Chapter 5.

The next chapter will study the methodologies that will better deal with the thesis problem.

4 | Research methodology

In this chapter, with the insights from the literature review made in Chapter 3, the methodology to address the problem of this thesis will be delineated. The most prominent benchmarking tools will be scrutinized to discover the most suitable tool to measure and compare performances between PPP and EPE hospitals.

4.1 Benchmarking: Overview

At its simplest, *Benchmarking* is a process of systematically comparing performance measures with a standard reference [80], being that reference the best practice (benchmark) or set of best practices. This chapter identifies and briefly describes the most important models available and used to estimate the performance of homogeneous services. Offers a critical description of the strengths and limitations of those models, when applied to the health sector.

4.2 Benchmarking models

Each hospital has its way of "producing" health care. In general, it consumes resources (inputs) such as financial resources (money), staff, equipment, and consumable items, turning them into health care outputs such as treated patients. The efficiency of this production process can be measured by examining the relationship between both the inputs and outputs [81].

A comparison between homogeneous entities exhibiting similar production technologies can be made, using a frontier where most efficient organizations within the sample of organizations under analysis are placed. This frontier can be called the *efficient frontier*. That is, the frontier is composed of organizations producing a certain level of output by using the least quantity of inputs, or achieve the maximum output by using a certain level of inputs. These benchmarks are dominant entities over the remaining [17]. The examples that do not lie on the efficient frontier are considered inefficient in comparison to the efficient ones. The larger the distance to the frontier, the higher the inefficiency level.

Models to estimate the efficiency frontier can be divided into two broad categories, parametric and non-parametric. Parametric models use econometric tools, and require specification of a particular functional form of the production function. Non-parametric models, place no conditions on the functional form, estimating the shape of the frontier using observed data, thus, requiring smaller assumptions [17].

To this day, Stochastic Frontier Analysis (SFA) is the predominant form of the parametric models used. In the non-parametric case, most methods take the form of DEA and its derived models [82].

Although, SFA and DEA are the most commonly used and known benchmarking models, there are several other models including: ratio analysis, Least Squares Regression (LSR), Total Factor Productivity (TFP), Ordinary Least Squares (OLS). A brief characterization of each model is going to be presented. After that, more attention is given to the SFA and DEA models.

4.2.1 Ratio analysis

Ratio analysis is the simplest method to estimate performance. It takes information from the relationship between inputs and outputs. It defines efficiency as the number of output units per unit of input. Its main weakness is the difficulty of identifying a reliable benchmark that includes all inputs and outputs of health care organizations [83].

$$\text{Efficiency (Productivity)} = \frac{\text{Output}}{\text{Input}} \tag{1}$$

4.2.2 Least squares regression

LSR is a parametric model, known for assuming that all health care organizations are efficient. It accommodates multiple inputs/outputs and also accounts for noise using an error term. Its weak points include: using common techniques, which are not effective relationships, lack of capacity to identify the individual inefficient units, and its parametric nature requiring a pre-specified production function [83].

4.2.3 Total factor productivity

TFP is a methodology that goes around the weaknesses of ratio analysis by incorporating multiple inputs and outputs into a single performance ratio. TFP uses index numbers to measure differences across health care organizations [83]. The main indices used include Laspeyres index, Pasche index, Fisher index, Tornqvist index, and Malmquist index.

4.2.4 Ordinary least squares

OLS uses regression models to do an estimation of the production function. Thus, using a sample of health care units, it can estimate their associated production function. Through that, it determines whether (and how much) specific providers are above or below average efficiency levels. Not truly identifying efficient behaviour, since efficiency estimates are based on average performance and not on the production frontier, is one of the major drawbacks of OLS [81].

4.2.5 Stochastic frontier analysis

SFA, just like LSR, is a parametric model. It improves LSR by assuming that all firms may not be efficient, at the same time that it accounts for noise [81]. SFA is used to estimate the production frontier for a given sample of health care providers using regression-based techniques. Does it by decomposing the error term into two parts: one is a one-sided error term that measures inefficiency, and another is a more usual error term that captures random influences affecting the organization.

SFA can be used to measure technical efficiency, scale economies, allocative efficiencies, and technical change. The need for specification of both functional form and specific distributional form for the inefficiency term is the major shortcoming of SFA [83]. The SFA model takes the form as follows [81]:

$$\mathrm{y_j} = \beta_j \mathrm{x_j} + u_j + v_j \tag{2}$$

where y_j is the vector of outputs, x_j is the vector of inputs, u_j is the one-sided inefficiency term ($u_j \geq 0$ for all j), v_j is the two-sided error term which is assumed to follow a Gaussian distribution N (0,1) and u_j and v_j have zero covariance.

The SFA cost frontier is usually estimated using a translog functional form that gives the possibility to test a broader range of assumptions about the nature of the cost function, not imposing restrictive *a priori* assumptions on its functional form. In hospital studies, where the sample size is usually limited, this may create measurement error and bias in inefficiency estimates through the inappropriate aggregation of inputs and outputs [81].

4.2.6 Data envelopment analysis

DEA is a non-parametric model that can take multiple inputs and outputs using them in a linear programming (optimization) model that gives a single score of technical efficiency per observation. DEA plots the production frontier for a sample of health care providers using linear programming techniques [81]. The providers that lie on the frontier are considered efficient, and the ones who lie inside are inefficient.

DEA also estimates the technical and allocative efficiency of all providers. It is non-stochastic, does not assume that the distance that a specific organization has from the efficient frontier may result from data noise. DEA can be used to calculate technical efficiency, scale efficiency, allocative efficiency, congestion efficiency, and technical change. In contrast with parametric methods, DEA (and its derivative models) is the only method available that can easily estimate multiple input–multiple output models [81].

4.2.7 What is the most frequently employed model to estimate efficiency?

As previously demonstrated, there are some models to evaluate performance in the health care sector, as well as in many other sectors of the economy. DEA has become the preferred model followed by SFA, as it can be seen in Figure 4.1.

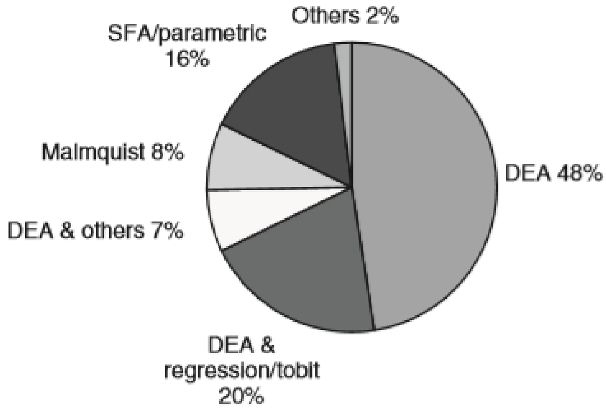

Figure 4.1: Methods used in reported studies
Source: Hollingsworth [81].[1]

To date, in the specific area of analysis of health care services and study of hospitals, the application of SFA has been limited. However, DEA has been primarily used, with hundreds of published applications, as it can be seen in Table 3.2 of the literature review made in Chapter 3.

4.3 Comparison of SFA and DEA

Being DEA and SFA the most used methodologies, what are the differences between them? What are the pros and cons of both? Which one is more adequate to evaluate hospital performance?

A simple way to describe what distinguishes both approaches lies in the fact that the econometric one seeks to determine the efficiency of the organization against a theoretical boundary. In contrast, the DEA approach seeks to determine the efficiency of the organization relative to other units within the same industry. There are three main differences between the two models, DEA is a non-stochastic approach, so it does not allow for statistical noise and random shocks, whereas SFA models random behaviour through the error term. SFA is parametric and DEA is non-parametric. The ability of DEA, contrary to SFA, to easily handle multiple outputs is also a noteworthy feature of the model, making it very appealing to the research [81].

One major problem of the SFA method, because of its parametric nature, is the need to provide a prior specification of functional form. This problem can be overcome by applying DEA. In DEA, the efficient

[1] Hollingsworth identified 289 studies, up to and including 2005, that measured efficiency in health services and provided statistics on the methods used.

frontier is created and shaped by data only, known beforehand, without any theoretical assumptions. As a consequence, DEA is more flexible than SFA, having the efficient frontier moulding itself to the data [82]. Hence, when measuring performance, when making a theoretical approach, maybe SFA is a better option. However, if an empirical standpoint is made, for example, when studying hospital performance, the flexibility offered by DEA can be seen as an attractive feature. Another drawback of SFA, is that it can introduce measurement error and bias in inefficiency estimates along with the aggregation of inputs and outputs, especially when the sample is small [81].

A DEA con is that it assumes that all data has no error (noise). SFA can model error, so if measurement error is present, and both methods have an identical frontier, SFA may give better efficiency estimates [82]. However, incorrect error specifications in SFA lead to biased results of efficiency.

Both models use two different principles when measuring the performance of hospitals. SFA does one overall evaluation with all the observations to achieve the estimates of inefficiency. Meanwhile, DEA does a separate optimization per hospital, giving a better fit for every single observation and a better basis for discovering sources of inefficiency per observation.

All these differences suggest that maybe there is not a “best method” for measuring performance. There is no need for consensus when the truth is that one might be preferred over the other to help solve different problems. For example, SFA can be useful when studying the future behaviour of an entire population of hospitals, whereas DEA might be the best option when studying individual hospitals, and how to remove inefficiency in those hospitals [84].

Considering the comparison made between both models, DEA will be the model used to evaluate the Portuguese PPP and EPE hospitals. As such, DEA will be detailed in the next section.

4.4 Data envelopment analysis

DEA was initially suggested by Charnes, Cooper, and Rhodes [85], who worked on measuring the efficiency of Decision Making Unit (DMU)s, based on the first studies of Farrel [86] on the measurement of efficiency of multiple input/output combinations. Since 1978, it has been significantly developed. In recent years, DEA has been applied to evaluate the performance levels of many different types of activities and contexts across the globe. Using DMUs of all different forms to evaluate the performance of entities, such as hospitals (see Sections 3.1 and 3.2), universities [87], business firms [88] and cities [89].

The first time that DEA was applied in health care was in 1983 by Nunamaker and Lewin [90]. The authors measured routine nursing service efficiency. To evaluate overall hospital efficiency, Sherman [91], applied DEA for the first time in 1984. Since then, DEA has been the leading model used to assess hospital efficiency across the world [83].

DEA is a non-parametric, non-stochastic, programming technique used to estimate production frontiers. Contrary to parametric techniques, DEA does not assume a particular functional form for the frontier. DEA empirical estimates an efficient frontier by applying a mathematical linear programming model to the observed data [81]. In some cases, non-linear programming has been associated with DEA, improving its features [92]. The shape of the efficiency frontier is determined by the data, simply considering that an organization that uses less input than another to produce the same amount of output can be considered more efficient. Observations with the highest ratios of output to input are considered efficient, and these are the observations that are going to create the efficiency frontier [82]. The frontier is made by a series of linear segments connecting one efficient observation to another. The ”best observed practice” is the foundation for the construction of the frontier and is, therefore, only an approximation to the true, unobserved efficiency frontier. Inefficient organizations are “enclosed” by the efficiency frontier in DEA [83]. The inefficiency of the organizations within the frontier boundary is calculated by the distance to the frontier [82]. DEA allows the evaluation of the performance of DMUs, peer entities (similar organizing units), which convert multiple inputs into multiple outputs, and also allows direct comparisons of efficiency between providers based on their observed production.

4.4.1 Model orientation

DEA models can follow two distinct orientations: Input and Output orientation. When efficiency is input oriented there is a focus on reducing inputs to increase efficiency, holding outputs. For example, in input orientation, it is assumed that health care managers can have more control over inputs rather than in the admission of patients. However, in output orientation, the goal is the output augmentation to increase efficiency, holding inputs. For example, health care managers can use marketing to attract more patients to their hospitals and increase the number of admissions [83]. In public services, as the public hospitals in Portugal, managers should focus on resources’ waste reduction, at the same time that the production increases, to mitigate waiting times and to improve access to the health care services. Therefore, often mix (non-oriented and directional) models can be preferred.

4.4.2 Basic frontier models

In DEA, there are two main approaches, Constant Returns to Scale (CRS) and Variable Returns to Scale (VRS), chosen according to the problem conditions. The initial basic frontier model was CRS, used when the scale of economies do not change as the size of the service increases. CRS reflects the fact that output will change by the same proportion as inputs are changed. VRS, contrary to CRS, does not assume that the scale of economies maintains constant as the size of the service facility increases [83]. In other words, VRS reflects the fact that production technology may exhibit increasing, constant and decreasing returns to scale.

Figure 4.2 illustrates DEA frontiers under both basic frontier models. Considering the point G, the distance from E to F measures the effects of economies of scale in production, whereas the distance from F to G measures inefficiency. It is visible that, under VRS more DMUs will be considered efficient since in CRS any economies of scale are soon incorrectly deemed as inefficient. [81]

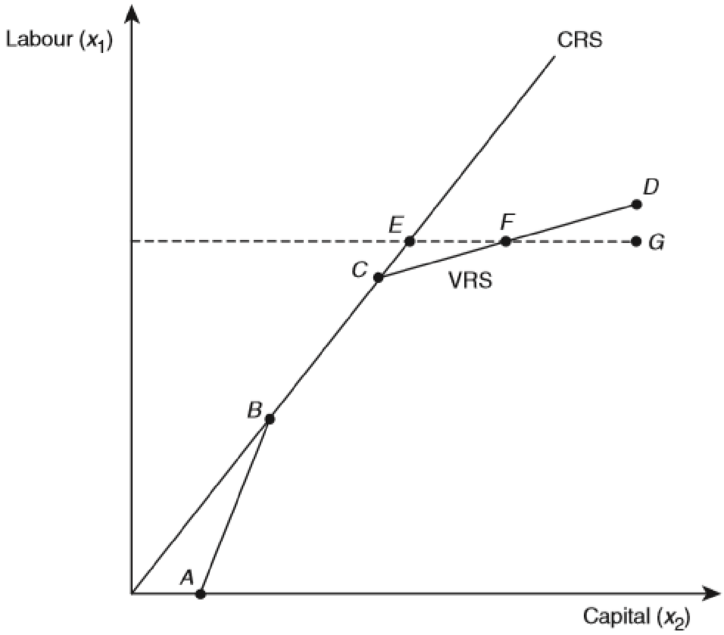

Figure 4.2: CRS and VRS under DEA
Source: Hollingsworth and Peacock [81]

4.4.3 Mathematical formulation

The general mathematical formulation of DEA corresponds to the resolution of a set of *n* linear programs, in order to estimate the efficiency scores of DMUs. Since the objective of this work is to measure performances of different hospitals to compare them, the DMUs in this example are hospitals. Let's consider a set of m inputs, x_i^j, i = 1, ...,m, and a set of outputs, y_r^j, r = 1, ...,s, evaluated for *n* hospitals (j = 1, ...,n). Hospital *k* can be represented by the vector (x_i^j, y_r^j) [17]. It is required the creation of a group of other hospitals to evaluate the efficiency of a given hospital, *k*, against which comparison will be undertaken.

Technical efficiency of hospital *k* can be calculated by minimizing or maximizing the objective function θ^k. Hospital *k* is technically efficient if $\theta^k = 1$, $\theta^k \in [0, 1]$, that is, input consumption is optimal [84]. Otherwise, to become efficient hospital *k* must reduce its resource consumption and waste. The input-oriented and output-oriented DEA models can be described as follows:

DEA input-oriented:

min θ^k
s.t.
$\sum_{j=1}^{n} \lambda^j x_i^j \leq \theta^k x_i^k, i = 1, \ldots, m$
$\sum_{j=1}^{n} \lambda^j y_r^j \geq y_r^k, r = 1, \ldots, s$ (3)
$\sum_{j=1}^{n} \lambda^j = 1 \quad (VRS)$
$\lambda^j \geq 0, j = 1, \ldots, n$

Optimal input/output combination for hospital *k* can be describes as:
$(x_i^*)^k = \sum_{j=1}^{n} \lambda^j x_i^j = \theta^k x_i^k$ (4)
$(y_r^*)^k = \sum_{j=1}^{n} \lambda^j y_r^j = y_r^k$

DEA output-oriented:

max Φ^k
s.t.
$\sum_{j=1}^{n} \lambda^j x_i^j \leq x_i^k, i = 1, \ldots, m$
$\sum_{j=1}^{n} \lambda^j y_r^j \geq \Phi^k y_r^k, r = 1, \ldots, s$ (5)
$\sum_{j=1}^{n} \lambda^j = 1 \quad (VRS)$
$\lambda^j \geq 0, j = 1, \ldots, n$

Optimal input/output combination for hospital *k* can be describes as:
$(y_r^*)^k = \sum_{j=1}^{n} \lambda^j y_r^j = \Phi^k y_r^k$ (6)
$(x_i^*)^k = \sum_{j=1}^{n} \lambda^j x_i^j = x_i^k$

where, n $\rightarrow$ number of hospitals; $m \rightarrow$ number of inputs; $s \rightarrow$ number of outputs;

$\lambda \rightarrow$ scale coefficient $\in [0, 1]$; $x \rightarrow$ inputs; $y \rightarrow$ outputs.

In order to calculate technical efficiency, some conditions must be met, as it is shown above. For example, the input-oriented model is subjected to four conditions [81]. The first condition to be met is that the weighted comparison group must use no more than a fraction (θ^k) of the m inputs which hospital k uses. The second condition is that the comparison group of hospitals must produce at least as much output, in all s dimensions, as the unit under study, hospital k. To allow for VRS technology, a third constraint is added. With the absence of this condition, the model would assume CRS, and hence θ^k would be constant across all inputs [81].

Due to the linear programming duality, the previous input-oriented model is equivalent to [93]:

$\max \sum_{r=1}^{s} x_r y_r^k + w$
s.t.
$\sum_{r=1}^{s} x_r y_r^j - \sum_{l=1}^{m} v_i x_i^j - w \leq 0, j = 1, \ldots, n$
$\sum_{i=1}^{m} v_i x_i^k = 1$ (7)
$x_r, v_i \geq 0$
$w >$-∞

4.5 Benefit of doubt

The label "BoD" derives from one conceptual point of DEA, which is, knowledge on the right weighting scheme for hospital performance benchmarking can be acquired from the hospital data themselves [94]. Behind this lies the idea that, when hospital k performs well in a particular indicator, that policy dimension is relatively essential to the hospital. Opposing, a hospital gives less importance to dimensions on which performs worst when comparing to the other hospitals in the group. Thus, giving this interpretation of higher importance to relative strengths, hospitals will set their own "optimal" weighting schemes [94].

The BoD implies that when having no information whatsoever about the correct weights, each hospital is considered the best possible comparing with the other hospitals in the group sample when its aggregate performance is determined [37].

The BoD approach can be considered equivalent to the original DEA model since it considers all indicators as outputs, but a "dummy input" equal to one for all the hospitals [95]. Considering inputs equal to one, for example, in the output-oriented model, it is intuitive that the first condition is no longer necessary, since x is always equal to one and the conditions one and three can be simplified, becoming:

DEA output-oriented:

$$
\begin{aligned}
&\text{max } \Phi^k \\
&\text{s.t.} \\
&\textstyle\sum_{j=1}^{n} \lambda^j y_r^j \geq \Phi^k y_r^k, r = 1, \ldots, s \\
&\textstyle\sum_{j=1}^{n} \lambda^j = 1 (VRS) \\
&\lambda^j \geq 0, j = 1, \ldots, n
\end{aligned}
\tag{8}
$$

The main advantage of applying the BoD is the flexibility that it provides in the weight choice. Any other weighting approach would worsen the position of the evaluated hospital relative to the other hospitals [95]. By using it, hospitals cannot argue that their poor relative performance in one indicator is due to a harmful or unfair weighting approach.

Nonetheless, there are also negative aspects to this flexibility. For example, it allows for a hospital to look like a good performer in a way that is not justifiable. Hospitals can ignore some indicators on which they do not perform well by giving (quasi) zero-weights [95]. Having trustworthy experts in the application of the BoD approach gives higher credibility and acceptance of the results obtained. In fact, it is possible to include some weight restrictions to the dual model of Equation (8) to avoid that problem.

4.6 Malmquist index

The Malmquist index was introduced by Caves *et al.* [96], being usually applied to measure productivity changes over time. Using this index, health care managers have the capability to compare the productivity of hospitals across two different periods. The Malmquist index can be obtained using frontier approaches such as SFA or DEA, but only the DEA one is going to be used here. Either the output-oriented or input-oriented approach can be used.

The DEA Malmquist index can multiplicatively decompose the overall efficiency measure into two components, efficiency change index (catching-up effect) and a technological/environmental change index (innovation). Efficiency change index measures change in performance that has effectively occurred due to hospital improvement, also measuring the difference in performance to the benchmark. The technological/environmental change index focuses on the way benchmarks evolved, measuring the changes in the best practices between the two analyzed periods [94]. It is important to stress that these two components can be mutually exclusive, for example, progress can be observed due to a robust and better performance of a specific hospital even if there is a less convenient environment than in the starting period. Alternatively, progress is observed due to better practices available, which the hospital still only takes advantage of partially.

The most widely applied Malmquist index approach uses distance functions, usually through DEA, using them to construct quantity indices as ratios [82]. Then, it decomposes this into both components mentioned above, efficiency and technological change.

The Malmquist index can be computed in either the input or the output orientations. The most common way of formulating the Malmquist index is [82, 97]:

$$\mathrm{M}_O^{t,t+1}\left(X_t, Y_t, X_{t+1}, Y_{t+1}\right) = \underbrace{\frac{D_O^{t+1}\left(X_{t+1}, Y_{t+1}\right)}{D_O^t\left(X_t, Y_t\right)}}_{=E(X_t,Y_t,X_{t+1},Y_r)} \underbrace{\left[\frac{D_O^t\left(X_{t+1}, Y_{t+1}\right) \cdot D_O^t\left(X_t, Y_t\right)}{D_O^{t+1}\left(X_{t+1}, Y_{t+1}\right) \cdot D_O^{t+1}\left(X_t, Y_t\right)}\right]^{1/2}}_{=T(X_t,Y_t,X_r,Y_{t+1})} \tag{9}$$

As mentioned, the Malmquist index is defined using distance functions, hence D_o represents those functions in periods t and $t+1$. If $M_o > 1$, performance has decreased between period t and period $t+1$. If $M_o = 1$, no change in performance from t to $t+1$ was detected. If $M_o < 1$, performance has increased between period t and period $t+1$ [83]. E represents the change in the technical efficiency levels between t and $t+1$, whereas T represents technological change, changes in productivity levels due to technical progress, in the same period [97].

A value of E less than one means that technical efficiency has improved in period $t+1$ compared to t in that the hospital has moved closer to the frontier. When E is greater than one, then the hospital is further away from the frontier. Regarding T, a value smaller than one means that the whole industry is producing more outputs in the period $t+1$ than in t, in other words, the sector experienced productivity improvements over time. When T is higher than one, it is a result of productivity loss in the industry between both time periods [82].

The technical efficiency change term E is used when calculating efficiency change under CRS. Although, when VRS exist, it is possible to decomposed E as follows [97]:

$$E(X_t, Y_t, X_{t+1}, Y_{t+1}) = \underbrace{\frac{D_{0,\mathrm{VRS}}^{t+1}\left(X_{t+1}, Y_{t+1}\right)}{D_{O,VRS}\left(X_t, Y_t\right)}}_{=P(X_t,Y_t,X_{t+1},Y_{t+1})} \cdot \underbrace{\frac{D_{0,\mathrm{CRS}}^{t+1}\left(X_{t+1}, Y_{t+1}\right) / D_{O,\mathrm{VRS}}^{t+1}\left(X_{t+1}, Y_{t+1}\right)}{D_{O,\mathrm{CRS}}^{t}\left(X_t, Y_t\right) / D_{O,\mathrm{VRS}}^{t}\left(X_t, Y_t\right)}}_{=S(X_t,Y_t,X_r,Y_r)} \tag{10}$$

Where, the efficiency change calculated under VRS is referred to as P (pure technical efficiency change), and the change in the deviation between both technologies, CRS and VRS, is referred as S (scale efficiency change).

A value of P less than one reflects efficiency gain which means the hospital is closer to the VRS frontier in period $t+1$ than it was in period t; the opposite holds for a value of P greater than one. If S is smaller than one, it means that there were improvements on scale-efficiency between the two periods [82].

4.7 Malmquist index for group comparison

The Malmquist index usually aggregates the values of all DMUs before the creation of the index in one "big" DMU. The main problem is that the "big" DMU is not clearly defined, which could lead to conflicting results. The Malmquist index for group comparison can be used to overcome this problem. That way, it is not required a subjective definition of the "big" DMU, as it is now possible to handle all the individual DMUs directly [98]. This method compares relative to group-specific frontiers only, does not create a common frontier, so it does not assume convex combinations of group-specific to be feasible.

Malmquist index for group comparison, already used by Ferreira and Marques [6] and Camanho and Dyson [98], as regular Malmquist index, can be multiplicatively decomposed into two indexes: one that reflects the efficiency of DMUs working in similar conditions, and another that reflects the difference in productivity between the best-practice frontiers of the different groups. A crucial characteristic of this index is to focus on group comparison for a given moment in time [98].

The general approach to use the Malmquist index for group comparison starts with applying DEA methodology to identify the group frontiers. Then, using Malmquist indexes, it is measured group performance. The Malmquist index adaptation used no longer measures productivity change amongst two time periods, allowing for a cross-sectional comparison of the performance of groups of DMUs operating under distinct conditions, at a given moment in time. With this, a new aggregated performance measure is obtained, which can be multiplicatively decomposed into the two indexes previously mentioned. The overall performance measure can be written as follows [98]:

$$\mathrm{I}^{AB} = \underbrace{\frac{\left[\prod_{j=1}^{\delta_A} D^A\left(X_j^A, Y_j^A\right)\right]^{1/\delta_A}}{\left[\prod_{j=1}^{\delta_B} D^B\left(X_j^B, Y_j^B\right)\right]^{1/\delta_B}}}_{=IE^{AB}} \cdot \underbrace{\left[\frac{\left(\prod_{j=1}^{\delta_A} D^B\left(X_j^A, Y_j^A\right)\right)^{1/\delta_A}}{\left(\prod_{j=1}^{\delta_A} D^A\left(X_j^A, Y_j^A\right)\right)^{1/\delta_A}} \cdot \frac{\left(\prod_{j=1}^{\delta_B} D^B\left(X_j^B, Y_j^B\right)\right)^{1/\delta_B}}{\left(\prod_{j=1}^{\delta_B} D^A\left(X_j^B, Y_j^B\right)\right)^{1/\delta_B}}\right]^{1/2}}_{=IF^{AB}} \tag{11}$$

IE^{AB} is a ratio that compares how efficiency spreads within groups, and IF^{AB} evaluates the difference in productivity between the frontiers of the two groups. This decomposition shows that improvements in performance can be caused by: low scatter in efficiency levels of the DMUs in one group when in comparison with another, and/or the best practice frontier is dominant. When IE^{AB} is less than one means that efficiency spread, *i.e.*, there is better consistency in efficiency levels in DMUs of group A than in those of group B. IF^{AB} is an index that measures the distance between the best-practice frontiers of groups A and B, when IF^{AB} is less than one means that productivity of the frontier of group A is greater than the one of group B [98]. The use of this Malmquist index version is likely to help to answer the research question: "Do PPP hospitals outperform the corporatized ones?". If one constructs two different clusters (one for PPPs and another for EPE hospitals), then they can be compared through the Malmquist index above. It is worth to mention that the applied model will couple the Malmquist index with the BoD with weight restrictions to include decision making preferences.

4.8 Summary of chapter 4

This chapter addresses the main benchmarking models to compare performances. In articulation with the literature previously reviewed, the two most used and adequate models are SFA, a parametric model, and DEA, a non-parametric model. By comparing both models, it is seen that both have pros and cons. Nonetheless, DEA has the advantage of providing a fit for every single observation, whereas SFA does one overall evaluation with all the observations to measure efficiency. For that reason, and because it is by far the most used methodology in the literature (4.1), the DEA will be the preferred methodology to address the thesis' objective. The BoD, a particular case of DEA, is also going to be used to take full advantage of the flexibility in the weight choice. By doing so, it is possible to say that poor relative performance in one indicator is not associated with unfair weighting. The Malmquist index for group comparison will be used to complement the performance analysis. This index uses DEA technology to identify group frontiers and then measures group performance.

In the next chapter, the data collection and the selection of variables will be made, as well as the definition of the models that will comprise this study.

5 | Case study

This chapter is divided into two parts. The first consists in gathering and analyzing all the data needed to process the following stages. The choice regarding the variables to be used in the model is going to be made, taking into consideration the literature review made in Chapter 3, data availability, and relevance to the topic. A statistical analysis will be made to the selected variables to make a first impression on the difference between both hospital models. The second part will explain the different models of this study and how those models were implemented.

5.1 Data and variables

In order to successfully evaluate the performance of hospitals, an appropriate, extensive, and reliable database is needed. As such, the selection of the variables, as well as the sources to be used, must be carefully made. In general, the choice of variables follows criteria like data availability and quality, as well as a comprehensive literature review.

5.1.1 Data collection

The data used in this study was essentially provided by the official database of the Portuguese Ministry of Health, the Central Administration of Health Systems (Administração Central do Sistema de Saúde (ACSS)), and also by the online platform *Transparência*. The ACSS publishes monthly, on its website,[1] benchmarking reports for EPE hospitals and PPP hospitals, of the NHS. The diversity and typology of indicators take into account the National Health Plan and is close to the internationally adopted methodologies aiming to compare the performance of Portuguese hospitals in the international context. It provides data on access, care performance, safety, volume and utilization, productivity, and economic and financial data that convey to some extent the efficiency of hospitals. *Transparência* is an Open Data initiative carried out by the Ministry of Health, with the logic of making available and making fully accessible the vast set of data underlying the operations and transactions that take place within the scope of the NHS activities. It centralizes, on its website,[2] in an accessible and intuitive online platform, the data produced by the systems inserted in the NHS. Both sources provide monthly data, for each variable, of over 40 hospitals. When encountered missing data from the sources mentioned, the research for data also included looking through official reports of hospitals, available on their websites.

[1] http://benchmarking.acss.min-saude.pt/
[2] https://www.sns.gov.pt/transparencia/

5.1.2 Variables

The selection of variables is a critical decision to assure that the real world is portrayed in the best possible way when performing efficiency analysis [83]. Biased results can result in the absence of a thoughtful selection of variables. When benchmarking hospitals, this selection is particularly tricky since each hospital provides a wide range of services. The variables used in this work were selected based on data availability, the characteristics of the Portuguese hospitals, and a literature review.

Using the literature review made in Subsection 3.4, a crossing between the most used inputs and outputs in the literature with the variables provided by the sources ACSS and *Transparência* was made. Most variables were available in the ACSS, being only two provided by the *Transparência* platform ("% of Hemorrhagic stroke mortality" and "% of Ischemic stroke mortality"). Some variables were not precisely the same, but proxies were made to overcome that obstacle.

Since this work uses BoD it will not consider inputs and outputs, rather using Key Performance Indicator (KPI)s. Nonetheless, the inputs and outputs of the literature review will be used to provide information on the best KPIs to choose. Looking to the top 10 inputs, the three most used could be found in the ACSS website giving origin to three related KPI's, "number of beds per inpatient", "Inpatient per FTE doctor", and "Inpatient per FTE nurse". ACSS provides information on doctors and nurses in Full Time Equivalent (FTE), a unit that indicates the workload of an employed person in a way that makes workloads comparable across various contexts. Most of the other inputs in the top 10 are related to hospital costs. Regarding those types of variables, information was more scarce due to the lack of information about PPP hospitals. Nonetheless, since it is essential to have a related cost variable, one KPI was chosen on operational costs, "operational costs per inpatient". However, it does not have information regarding two out of four PPP hospitals. Since the hospitals in the sample have different sizes, these KPI's provide information per inpatient to have a better point of comparison.

From the top 10 outputs, it is possible to cluster five of those outputs in access variables (number of emergency episodes, number of outpatient consultations, number of admissions, number of medical appointments, and number of day hospital sessions). The KPI's that were chosen, regarding these access variables and what the ACSS offers, were "% of first non-urgent medical appointments within the maximum (legislated) guaranteed time", "% List of Surgery Subscribers (LSS) with waiting time less than or equal to Maximum Response Time Guaranteed (MRTG) (P1-270 days)", "average delay before surgery (in days)", and "% of hip surgeries in the first 48 hours after fracture". It was selected as proxies for the number of surgeries, "ambulatory surgery (Homogeneous Diagnostic Groups (HDG)) for outpatient procedures" and "surgical outpatient (surgical interventions" (number of ambulatory surgeries over the total number of surgeries) were chosen from the ACSS website. The number of births was not considered in this study, although it was selected the variable "% of caesarean deliveries". For the number of inpatients discharged and length of stay, four KPI's were picked, "readmission within 5 days after discharge (different civil years)", "readmission within 30 days after discharge (different civil years)", "readmission within 31-180 days after discharge (different civil years)", and "% of admissions with delay over 30 days".[3]

The previous KPI's were selected regarding the information of Figure 3.1. However, since it was deemed necessary to have safety related KPI's, it was added to the work, "Pressure ulcer rate", "Catheter related bloodstream infections rate", "Postoperative pulmonary embolism/deep vein thrombosis per 100 inpatients", "Postoperative septicaemia cases per 100 inpatients", "% of Instrumented vaginal births with 3rd and 4th degree lacerations", "% of Hemorrhagic stroke mortality", and "% of Ischemic stroke mortality" [37, 92]. It was also added to the KPIs "Hospital occupancy rate" since the occupancy of a hospital provides essential information on their productivity. Initially, three more variables were considered, two economic inputs, EBITDA and Operational results, and % Urgency episodes served within the expected time. However, when the database started to be created, the first two were excluded since they did not have any data concerning PPP hospitals, and the other was removed because it only had data starting in mid-2016.

As it is possible to see from Figure 5.1, three different types of KPI's were created, efficiency, access, and quality. Efficiency is subdivided in services availability, which regards the existence of resources to be used when needed, productivity, providing ratios between outputs and inputs of the hospital, and financial KPI's, measurements that describe economic units. Access, in this particular case, access to healthcare services, is defined by the ability of one citizen to use a specific service whenever necessary at her/his will. Quality KPI's refer to patients' care appropriateness and clinical safety, where the first regards the ability to provide patient-centred care services backed by evidence-based guidelines. The second is the ability to prevent complications or even preventable deaths from happening during care.

[3] Some variables selected are not the same as the ones in the top 10, either due to its inexistence in the ACSS or *Transparência* platform, or because it was considered more relevant to the study to use a proxy.

- KPIs
  - Efficiency
    - Services Availability
      - k_1^+M, Inpatient per hospital beds,
    - Productivity
      - k_2^+M, Inpatient per FTE doctor,
      - k_3^+M, *Inpatient per FTE nurse*,
      - k_4^+, Hospital occupancy rate,
    - Financial
      - k_5^-, Operational costs per inpatient,
  - Access
    - k_1^-m, Inpatient per hospital beds,
    - k_2^-m, Inpatient per FTE doctor,
    - k_3^-m, *Inpatient per FTE nurse*,
    - k_6^+, % of first non-urgent medical appointments within the maximum (legislated) guaranteed time,
    - k_7^+, % LSS subscribers with waiting time less than or equal to MRTG (P1 - 270 days),
    - k_8^-, Average delay before surgery (in days),
    - k_9^+, % of hip surgeries in the first 48 hours after fracture,
  - Quality
    - Care Appropriatness
      - k_{10}^+, % Ambulatory surgery (HDG) for outpatient procedures,
      - k_{11}^-, % Readmission within 5 days after discharge (different civil years),
      - k_{12}^-, *% Readmission within 30 days after discharge (different civil years)*,
      - k_{13}^-, % Readmission within 31-180 days after discharge (different civil years),
      - k_{14}^-, % of admissions with delay over 30 days,
      - k_{15}^-, % of caesarean deliveries,
      - k_{16}^+, % Surgical outpatient (surgical interventions),
    - Clinical Safety
      - k_{17}^-, Pressure ulcer rate,
      - k_{18}^-, % of Hemorrhagic stroke mortality,
      - k_{19}^-, % of Ischemic stroke mortality,
      - k_{20}^-, Catheter related bloodstream infections rate,
      - k_{21}^-, Postoperative pulmonary embolism/deep vein thrombosis per 100 inpatients,
      - k_{22}^-, *Postoperative septicaemia cases per 100 inpatients*,
      - k_{23}^-, *% of Instrumented vaginal births with 3rd and 4th degree lacerations*,

Figure 5.1: KPI's selected. Note: k^+ identifies the desirable variables while k^- identifies the undesirable ones. Difference between km and kM is explained in Subsection 5.2. Variables in *italic* mean that they will be subsequently removed from the dataset.

Source: Author

5.1.3 Statistical analysis

A statistical analysis of the sample was made for each KPI selected. The analysis distinguishes both the legal regimes as well as the years of the data. Data from each year (from 2015 to 2019) and also from 2015-2019 altogether is analyzed in four dimensions: overall hospitals (comprising all 28 hospitals), EPE hospitals (including 24 EPE hospitals), PPP hospitals (comprising four PPP hospitals), and an individual analysis of each of the four PPP hospitals. The statistical analysis measures the maximum value (max value), 75th percentile (Q3), mean, median, 25th percentile (Q1), minimum value (min value), standard deviation (σ), and Coefficient of Variation (CV) for each KPI of the study.

To complement the statistical analysis, two tests were implemented, the two-sample t-test and Kruskal-Wallis test. The t-test refers to any statistical hypothesis test in which the test statistic follows a Student's t-distribution under the null hypothesis. The Kruskal-Wallis test is a non-parametric test that does not require normality of distribution. Both tests are used to determine if there are statistically significant differences between two or more groups of an independent variable on a continuous or ordinal dependent variable [99, 100]. In this case, it was used to compare the two groups of hospitals (EPE and PPP hospitals). These tests reject the null hypothesis at the 5% significance level.

Table 5.1 provides the results of both tests for each KPI, presenting p-values, confidence intervals on the difference of means (Δ), T statistics, and the best performer in the t-test, and only the p-value and best performer for the Kruskal-Wallis test. Since the significance level is placed at 5%, whenever the p-value $>$ 0.05, the null hypothesis cannot be rejected. The rejection of the null hypothesis in the t-test means that there is no indication that the two groups under analysis have different means. In the Kruskal-Wallis test, the rejection of the null hypothesis implies that there is no indication that the two groups under study have different medians. From this analysis, in only three KPIs (k_1, k_{13}^-, and k_{20}^-) it is rejected the null hypothesis, that is, both EPE and PPP hospitals show similar results. k_{21}^- was not considered before because, although concluding that both hospitals have similar results, the same does not happen in the Kruskal-Wallis test, being the EPE hospitals the best performer (A).

KPIs k_2 and k_3 have both groups as the best performers because it depends on the interpretation made. These KPIs can be analyzed from an access or efficiency/productivity point of view.[4] In terms of access, it is better to have a low number of patients per doctor and nurse, being, in this case, the best performer the EPE hospitals (A). In terms of efficiency/productivity, it is wanted a higher number of patients treated per doctor and nurse. In this case, the best performer is the PPP hospitals (B). k_4^+ is a unique case since the best performer is not related with the highest mean. The optimal value for hospital occupancy occurs when the values lay in the range between 80 to 90%. Since the average of the EPE hospitals is 85,86%, and the PPP hospitals is 91,73%, the best performer in k_4^+ is the EPE hospitals (A).

Excluding some exceptions, explained before, both tests (t-test and Kruskal-Wallis) provide similar

[4] This KPI characteristic will be further developed in Subsection 5.2

Table 5.1: Statistical analysis: Student's t-test for means and Kruskal-Wallis non-parametric test for distributions for the variables of this study.

	Student's t-test					Kruskal-Wallis test	
	p-value	**Δ, lower bound**	**Δ, upper bound**	**T statistic**	**Best performer**	**p-value**	**Best performer**
k_1	0.7335	-0.3501	0.2466	-0.3405	similar	0.6165	similar
k_2	0.0022	-1.1537	-0.2539	-3.0688	A/B	0.0000	A/B
k_3	0.0000	-1.8972	-1.5049	-17.0138	A/B	0.0000	A/B
k_5^-	0.0000	564.0153	846.6634	9.7901	B	0.0000	B
k_6^+	0.0000	0.0615	0.0973	8.6926	A	0.0000	A
k_7^+	0.0000	-0.0414	-0.0164	-4.5456	B	0.0000	B
k_4^+	0.0000	-0.0699	-0.0475	-10.2568	A*	0.0000	A*
k_8^-	0.0000	0.3892	0.4990	15.8682	B	0.0000	B
k_{10}^+	0.0000	-0.0640	-0.0360	-7.0077	B	0.0000	B
k_{11}^-	0.0000	0.0017	0.0036	5.5560	B	0.0000	B
k_{12}^-	0.0001	0.0029	0.0087	3.9274	B	0.0000	B
k_{13}^-	0.1697	-0.0010	0.0055	1.3737	similar	0.0512	similar
k_{14}^-	0.0000	0.0062	0.0097	8.9040	B	0.0000	B
k_{15}^-	0.0000	0.0174	0.0326	6.4381	B	0.0000	B
k_9^+	0.0000	-0.1197	-0.0505	-4.8254	B	0.0000	B
k_{16}^+	0.0000	-0.0926	-0.0700	-14.1236	B	0.0000	B
k_{17}^-	0.0000	-0.0025	-0.0017	-10.7611	A	0.0000	A
k_{20}^-	0.5386	-0.0003	0.0006	0.6150	similar	0.6987	similar
k_{21}^-	0.1274	-0.1368	0.0171	-1.5254	similar	0.0001	A
k_{22}^-	0.0000	-0.5073	-0.2697	-6.4147	A	0.0000	A
k_{23}^-	0.0004	-0.0115	-0.0033	-3.5602	A	0.0000	A
k_{18}^-	0.0000	0.0408	0.0722	7.0692	B	0.0000	B
k_{19}^-	0.0000	0.0216	0.0361	7.8138	B	0.0000	B

A - when EPE hospitals are the best performers; B - when PPP hospitals are the best performers; similar - when both types of hospitals have equal performances; A/B - when both groups can be considered as best performers, depending on the point of view (access or productivity); A* - when the best performer is not related with the highest or lowest mean but to a range of values.
Δ represents the confidence intervals on the difference of means.
Source: author.

conclusions on the best performers, which may result from the considerable size of the sample. Based on the results obtained, some conclusions can be extracted regarding the types of KPIs:

- PPPs outperform EPE hospitals regarding the health care appropriateness;
- EPE hospitals outperform PPPs concerning clinical safety, although not in all KPIs;
- In terms of access KPIs there is a superior performance of PPP hospitals, with exception to one KPI;
- There is no clear evidence supporting the hypothesis that one group outperforms the other concerning efficiency.

These findings do not provide any consistent conclusion on the best type of hospital management, then, surfaces the need to use a benchmarking tool to optimise the weights given to each dimension. BoD and Malmquist index for group comparison are the tools that will provide a composite index allowing to discover which group (EPE or PPP) has a better overall performance.

5.1.3.A Pearson's correlation coefficients

It was calculated the Pearson's correlation coefficient for the variables used in this work to understand if there are variables providing the same information, i.e., highly correlated. According to Table A.4, Pearson's correlation coefficients between $k_l^{\pm}$ and $k_{l'}^{\pm}$, for $l \neq l' = 1, ..., 23$, are small or non-statistically significant at the level 5% of significance except for five coefficients. Those five coefficients show a considerable correlation among them. However, only two have $r \geq 70$, which means that those KPIs have a high association degree between them, possibly bringing redundant information into the model. The variables in question are "Inpatient per hospital beds" and "inpatient per FTE nurse" ($r = 0.84$), and "% readmission within 5 days after discharge(different civil years)" and "% readmission within 30 days after discharge(different civil years)" ($r = 0.81$). In order to reduce redundant information, one variable of each pair will be removed from the database. There are no specific criteria to choose between the variables so in the first case "inpatient per FTE nurse" will be removed, leaving "Inpatient per hospital beds", and in the second pair, stays "% readmission within 5 days after discharge(different civil years)", removing "% readmission within 30 days after discharge(different civil years)". The dataset will from now on comprise 21 variables instead of 23.

All Tables, containing the statistical analysis, are available online,[5] it was also added to the Appendix Tables A.1, A.2, and A.3 corresponding to the overall data (2015-2019), not all Tables were put for the sake of brevity, but as said they can all be seen online.

[5] https://drive.google.com/file/d/19dt7V4JoFxtkC6VwSnZob2xUFL7iOXdb/view?usp=sharing

5.1.4 Data pre-processing

In this study, it is expected that higher levels of quality contribute to better hospital overall performance. When in the presence of two quality observations, the largest one provides higher utility to the hospital. This characteristic holds for all desirable variables. However, quality can be measured through undesirable dimensions as well. In this case, an appropriate rescaling must be done.

To better understand the meaning of a desirable variable, an example is provided using k^+_{10}. Minor surgeries should always occur in ambulatory services. Otherwise, the inpatients will undergo more complex procedures, increasing the total costs of the surgery, and also the risks of a severe infection. Performing minor surgeries in ambulatory services, besides providing better care appropriateness, also enhances the efficiency level of the hospital. Hence, with higher values of k^+_{10} comes better levels of quality. All desirable variables are identified in Figure 5.1 with k^+. The undesirable variables, such as k^-_8, function oppositely. Higher levels of average delay or waiting time for a service, the longer the inpatient remains in the hospital, means less further admissions. Hence, with higher values of k^-_8 comes lower levels of quality. All undesirable variables are identified in Figure 5.1 with k^-.

As noted before, this study assumes that the utility function associated to each quality level should be increasing, which means, $\forall k^{q(\ell)}_{r(\ell)} \geqslant k^{p(\ell)}_{r(\ell)} \Longrightarrow U_{r(\ell)}(q) \geqslant U_{r(\ell)}(p), q, p \in \Omega$ [92]. Which means that data needs to be transformed using the following three regimes [101]:

- Larger-the-better (k^+):

$$\tilde{k}^{q(\ell)}_{r(t)} = 100 \frac{k^{q(\ell)}_{r(t)} - \min K_{r^{(\theta)}}}{\max K_{r^{(i)}} - \min K_{r^{(\ell)}}} \tag{12}$$

- Smaller-the-better (k^-):

$$\tilde{k}^{q(\ell)}_{r(t)} = 100 \frac{\max K_{r^{(l)}} - k^{q(\ell)}_{r^{(l)}}}{\max K_{r^{(l)}} - \min K_{r^{(l)}}} \tag{13}$$

- Nominal-the better (k_4):

$$\tilde{k}^{q(\ell)}_{r(t)} = 100 - 100 \frac{|k^{q(\ell)}_{r^{(l)}} - OB|}{\max\{\max K_{r^{(l)}} - OB; OB - \min K_{r^{(l)}}\}} \tag{14}$$

By applying the previous Equations, all KPI's were rescaled to the interval [0,100]%, being hospital p $\in \Omega$ with $\tilde{k}^{p(\ell)}_{r(\ell)}$ = 100 the best performer in the KPI $r(\ell) \in \Gamma(\ell)$. Whereas, the worst performer is given by $\tilde{k}^{p'(\ell)}_{r(\ell)}$ = 0, $\in \Omega$, in the same KPI. Equation 14 is used when the quality level of a variable is at its best when the data lays inside a reference band. Hospital occupancy desirable outcome, according to the ACSS website, stands between 80 and 90%. Therefore, since the Equation 14 receives only a single number (OB - object value) and not an interval, it was assumed that the targeted value for OB is 85%, so the best performer, in this case, is $\tilde{k}^{p(\ell)}_{r(\ell)}$ = 85, $\in \Omega$.

5.1.5 Sample

Sources provide information for more than 40 Portuguese hospitals, however, not all the hospitals have complete data, because the variable in question does not apply to that facility, or because it is unknown/not provided, or it was not measured in that period. Therefore, hospitals that had a considerable amount of missing information were excluded from the sample, remaining the ones that provided reliable information. Highly differentiated hospitals (oncology centres) were also removed from the sample in order to ensure homogeneity, as well as local health units since they arise from the vertical integration of a general hospital and primary health care centres. Given that, the sample comprises 28 hospitals, listed in Figure 5.3, where 20 are hospital centres, and 8 are singular hospitals (of which 4 are the PPP hospitals). Figure 5.2 gives a better understanding of the geographic representation of the selected hospitals.

In terms of the period of the sample, it was taken into consideration that the last PPP to start its operations was Hospital Vila Franca de Xira, in March 2013, so it was considered to begin the time of the sample by January 2014, ending in the last month available which is September 2019. Since 2014 has a considerable amount of missing data, in the variables selected, the beginning of the time was moved to January 2015. It is important to mention, as said in Subsection 2.6, that Hospital de Braga lost his PPP status in September 2019, becoming an EPE hospital. Having monthly data of 28 hospitals regarding their activity between January 2015 and September 2019, the database comprises of 1596 observations.

5.2 Models

This study considers two different models (Models I and II) to analyze hospital efficiency. These models intended to improve the robustness of findings and conclusions of this research. The creation of these two models is a result of having key variables with data missing where their exclusion was not an option. Productivity (Inpatient per FTE doctor) and economic (Operational costs per inpatient) KPIs do not provide any information on two PPP hospitals (Cascais and Loures). To consider all inputs, Model I was created, which contemplates all the hospitals (having the four PPPs) excluding these two variables in question, and Model II that excludes Cascais and Loures hospitals from the sample.

- *Model I* - variables with missing data are removed from the database (this model uses 15 out of 17 KPIs). Resulting in a sample of 1596 DMUs (relative to all 28 hospitals), from which 228 DMUs regard PPP hospitals.
- *Model II* - hospitals with missing data are removed from the database (this model uses all KPIs). Resulting in a sample of 1482 DMUs (relative to 24 out of 28 hospitals), from which 114 DMUs regard PPP hospitals.

Figure 5.2: Map of the hospitals.
Source: Author

NORTE
- Hospital de Braga, PPP
- Centro Hospitalar Póvoa de Varzim/Vila do Conde, EPE
- Centro Hospitalar do Médio Ave, EPE
- Centro Hospitalar Trás-os-Montes e Alto Douro, EPE
- Centro Hospitalar Tâmega e Sousa, EPE
- Centro Hospitalar Universitário de São João, EPE
- Centro Hospitalar Universitário do Porto, EPE
- Centro Hospitalar Vila Nova de Gaia/Espinho, EPE
- Centro Hospitalar Entre Douro e Vouga, EPE

CENTRO
- Centro Hospitalar do Baixo Vouga, EPE
- Centro Hospitalar Tondela-Viseu, EPE
- Centro Hospitalar Universitário Cova da Beira, EPE
- Centro Hospitalar e Universitário de Coimbra, EPE
- Centro Hospitalar de Leiria, EPE

LISBOA E VALE DO TEJO
- Centro Hospitalar Médio Tejo, EPE
- Hospital Distrital de Santarém, EPE
- Hospital de Vila Franca de Xira, PPP
- Hospital de Loures, PPP
- Hospital Fernando Fonseca, EPE
- Centro Hospitalar Universitário Lisboa Norte, EPE
- Hospital de Cascais, PPP
- Centro Hospitalar de Lisboa Ocidental, EPE
- Centro Hospitalar Universitário de Lisboa Central, EPE
- Hospital Garcia de Orta, EPE
- Centro Hospitalar Barreiro/Montijo, EPE
- Centro Hospitalar de Setúbal, EPE

ALENTEJO
- Hospital Espírito Santo de Évora, EPE

ALGARVE
- Centro Hospitalar Universitário do Algarve, EPE

Figure 5.3: Regions.
Source: Author

Each of these two models will unfold into two versions, *m* and *M*. The difference between them lays in the utilization of different equations, on KPIs k_1 and k_2, upon data pre-processing. These two variables are unique because they can be classified as desirable or undesirable variables depending on the way they are interpreted. If considered as an access KPI, it is better to have a low number of patients per doctor, and patients per bed, being, in this case, an undesirable variable, so Equation 13 is used in data pre-processing. From an efficiency/productivity approach, it is wanted a higher number of patients treated per doctor, and a higher number of patients per bed, being a desirable variable, consequently Equation 12 is used.

Models with *m* version will consider KPI k_1 and k_2 as undesirable ($k_1^- m$ and $k_2^- m$), whereas models with *M* version considers them as desirable ($k_1^+ M$ and $k_2^+ M$). All the other KPIs remain equal in both versions.

In conclusion, this study has two different models with two variants each: *Model I m*, *Model I M*, *Model II m*, and *Model II M*.

5.2.1 Model implementation

All the computations of the DEA model and Malmquist indexes were performed recurring to the Matlab® R2018a[6] due to its high-performance properties. The computational framework was all created by the author with help provided by ”Data Envelopment Analysis Toolbox”, a package for Matlab, developed by Álvarez, Barbero, and Zofío (2016) [102], that comprises functions to calculate the main DEA models. This package includes code that allows the application of CRS and VRS, including both desirable and undesirable outputs when measuring efficiency and productivity, i.e., Malmquist and Malmquist-Luenberger indices. Data Envelopment Analysis Toolbox is a free software, available for download at `http://www.deatoolbox.com`, and also on the open-source repository GitHub.[7] The main functions of this study, BoD and Malmquist index for group comparison, were not comprised in the toolbox mentioned above. However, the toolbox helped at many other aspects.

Regarding the DEA methodology, as said in Section 4.4, the model can follow two distinct orientations (input or output orientation) and two main approaches (CRS or VRS). In this study, it will only be used the CRS orientation of the DEA since both CRS and VRS orientations provide the same results in the BoD model. Concerning the orientation, it was decided to use an output orientation in the analysis. The reason for this preference is that the goal of healthcare facilities should not mainly focus on cutting costs and reducing inputs but focus on increasing their outputs. Improving the performance of treatments, hospitalization, and patients’ release, leading to a better quality of life and health, less mortality, and inadequate treatment.

[6] `https://www.mathworks.com/products/matlab.html`
[7] `https://github.com/javierbarbero/DEAMATLAB`

5.2.2 Model running

After running the model for the first time, using the BoD, poor results were obtained, being near all the efficiency scores equal to one. When searching for the cause of the problem, it was found that some variables were disturbing the results. Those variables were "catheter related bloodstream infections rate", "Postoperative pulmonary embolism or deep vein thrombosis per 100 inpatients", "Postoperative septicaemia cases per 100 inpatients", and "% of Instrumented vaginal births with 3^{rd} and 4^{th} degree lacerations". The origin of this issue may be related to the dimensionality of these variables in question since their rescaled values are pretty much all close to 100. A test run was made discarding these four variables and the results obtained were satisfying. As a result, they were removed from the dataset, reducing the number of KPIs from 21 to 17.

An improvement was achieved by removing those variables, but DEA (and all non-parametric methods) are particularly sensitive when using a high number of variables. Therefore, since it is considered mandatory for good results, it was employed the Principal Component Analysis (PCA) technique to narrow down the number of variables. PCA is a well-known technique to aggregate data with no significant loss of information, it seeks the linear combination of the original variables such that the resulting variables capture maximal variance [103].

PCA was applied to four different sets of KPIs, one for each model. This operation resulted in having only one KPI (k_{PCA}) for each model. In all the PCAs the first component can explain more than 97% of the original KPIs, as it can be seen in Table A.5, meaning that it is a worthy representation of the variance in the corresponding KPIs.

5.3 Summary of chapter 5

The data used in this study is provided by two websites, ACSS and *Transparência*. After processing, the final database comprises 17 variables and 28 hospitals. The statistical analysis does not provide clear evidence supporting the idea that one hospital model outperforms the other in terms of efficiency, cementing the need to use the benchmarking tools. Two different models were created, having two versions each. Model I removes variables with missing data from the database, having a sample with 1596 DMUs, Model II removes hospitals with missing data, having a sample with 1482 DMUs. Each model unfolds in two versions, m and M. The difference between them is solely in variables k_1, and k_2, whereas in m they are processed as undesirable, in M they are treated as desirable.

The chapter that follows presents and discusses all the results and findings of this research, obtained with the different models.

6 | Results and discussion

After implementing both models, BoD and Malmquist index for groups comparison, and using the database collected and treated in Chapter 5, a comprehensive range of results will be drawn in this chapter. This interpretation demands some enlightening regarding the efficiency measures in this study. The first aspect is that only technical efficiency will be calculated. Particularly, since an output orientation BoD model will be used, it is measured how much outputs (KPIs) can increase proportionally. The second aspect refers to the relative nature of results since the efficiency will be calculated by comparing one individual hospital against an efficiency frontier composed of all the hospitals (these efficiencies are not absolute calculations). At last, it is crucial to have in consideration that the homogeneity of hospitals, the quality of data, and the systematic measurement errors can profoundly impact on the consistency of the results and their estimation.

This section will also discuss these results, starting with the performances obtained with the BoD model, providing analysis in three different stages. The Malmquist group comparison will be assessed and commented, also for the four models of this study (Model I m, Model I M, Model II m, and Model II M).

Since this chapter presents the discussion and conclusion of the work developed, this chapter dispenses a summary, being this fulfilled by the general discussion.

6.1 BoD results

The BoD model estimates efficiency scores per DMU, resulting in 1596 efficiency scores per Model I, and 1482 efficiency scores per Model II. These results were exported from Matlab to Excel tables in order to further analyse them (those tables are provided online).[1] A statistical analysis, similar to the one made in Chapter 5 to the KPI's, was conducted for these results. The analysis was made per year, providing monthly, yearly, and global values. The same dimensions of the previous analysis were used: the overall set of hospitals, EPE hospitals, PPP hospitals, as well as an individual analysis per PPP. The statistical analysis evaluated the maximum and minimum value, the upper and lower bounds of the 90% confidence intervals, the quartiles, mean, median, the frequency of good social performers (F(τ=0.90)), the standard deviation (σ), and the Coefficient of Variation (CV). The frequency of good social performers provides the probability of getting hospitals whose technical efficiency is, at least, equal to τ, here defined as τ = 90%. The two-sample t and Kruskal-Wallis test were used.

Not all the Tables, providing the results of analysis previous described, will be portrayed in this

[1] https://drive.google.com/file/d/1bjnB9ZdYTuUybBSN4nenAPfSY3aJ1V9I/view?usp=sharing

chapter, although they can all be seen online.[2]

For each model, in a first stage, a table comprising the statistical analysis of the performances was presented, followed by plots of the monthly and yearly efficiencies' evolution over the years. After, the results of the statistical tests were provided. The four different models of the study were discussed. However, due to the similarity in some results, only the first model, Model I m, was discussed in depth.

6.1.1 Model I m

Model I m uses all data from hospitals but does not include KPIs k_2 nor k_5^-. Nonetheless it considers KPI k_1 as undesirable ($k_1^- m$). Concerning this model, see Table 6.1 which gives the basic statistics for all the efficiency scores between 2015 and 2019.

First impressions provided by Table 6.1:

- Only EPE hospitals have efficient hospitals (performance = 1.00), the best PPP performer has a performance level of 0.977;
- The worst performer belongs to the EPE hospital group as well, with a performance score of 0.749. The lowest performance in a PPP hospital was 0.798;
- PPP hospitals have a higher average performance consistency (θ^{PPP}=0.891, [0.895; 0.888]$_{90\%}$) than EPE hospitals (θ^{PH}=0.855, [0.859; 0.856]$_{90\%}$);
- The standard deviation, in both groups, is considerably low, performances only vary four centesimal points. The same stands for the CV, both groups have a coefficient of 0.04, meaning that the performance results have low dispersion (or heterogeneity);
- The probability of finding good social performers ($\theta >$0.9) is way higher for the PPP hospitals (39%) than the EPE hospitals (12%).

Although the best hospital performance belongs to the EPE hospital group, it is possible to say that PPPs are not expected to be the worst performers than EPE hospitals. PPPs have a considerably higher average score, and a much higher percentage of good performers, being these two quite essential characteristics when comparing both types of hospitals.

Results suggest that to become efficient, both EPE and PPP hospitals need to amplify their performances by 14% and 11%, respectively. It is important to note, that there is only one efficient hospital (θ) = 1.00) which is Centro Hospitalar Póvoa de Varzim/Vila do Conde, a EPE hospital.

[2] https://drive.google.com/file/d/15IdfX1zXjcnWsrLicjE4GkM3Em4POeNf/view?usp=sharing

Table 6.1: Global performances basic statistics (Model I m).

2015/2019	Hospital Overall	EPE Hospitals	PPP Hospitals
Entries	1596	1368	228
Maximum value	1.000	1.000	0.977
Maximum {CI}	0.864	0.859	0.895
Percentile 75%	0.885	0.879	0.916
Mean	0.863	0.858	0.891
Median	0.860	0.855	0.887
Percentile 25%	0.834	0.831	0.868
Minimum {CI}	0.861	0.856	0.888
Minimum value	0.749	0.749	0.798
F(τ =0.90)	15.60	11.70	39.04
Standard deviation (σ)	0.039	0.037	0.035
Coeficient ofvariation (CV)	0.05	0.04	0.04

CV measures heterogeneity among observations, it is calculated by dividing σ by the mean.
CI - confidence interval
Source: Author.

Referring to the other three models, the same conclusions can be drawn since the global results are very similar. These tables, Table A.6, Table A.7, Table A.8, were added to the Appendix A.

Table 6.1 provided a big picture of the performances obtained with BoD. A more detailed analysis will be carried out by exhibiting the evolution of both hospital models performance throughout the years. Figure 6.1 presents the monthly evolution for 2015.

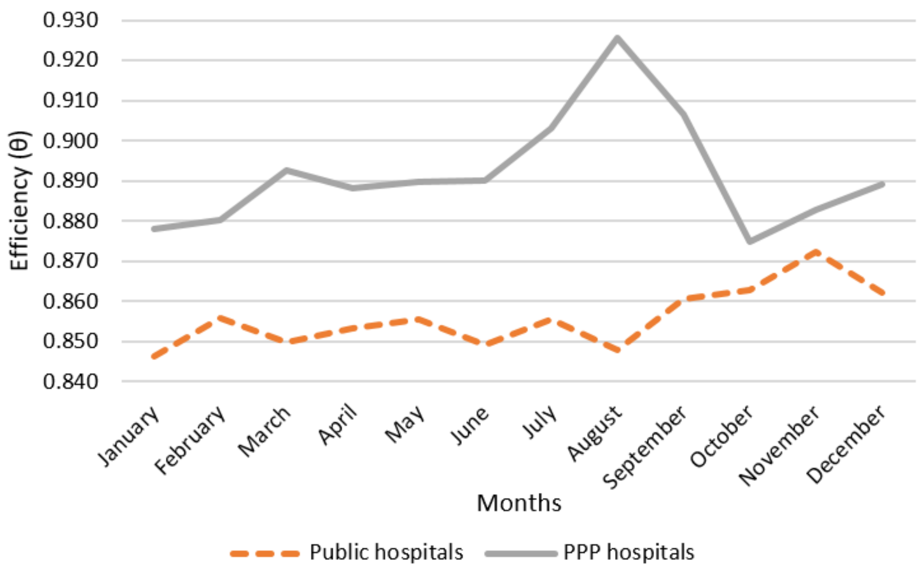

Figure 6.1: Hospitals monthly average performance variation in 2015 (Model I m).
Source: Author

From Figure 6.1, regarding each individual regime evolution, EPE hospitals' performance is somewhat stable until August. From August until November there is an increase of almost 2.5% in performance. In the PPP perspective, there is a slight growth from January to the middle of the year (except for April), experiencing a peak in the performance of more than 3.5% in August. Right after, in the following two months, there is a sudden decrease of 5%, and from October until the end of the year, a constant increase.

In Figure 6.1, it is visible the superiority of the PPP hospitals in terms of performance. PPPs outperform the EPE hospitals in every month of the year, on average terms. The difference in performances, between both models, in the first half of the year, is pretty much constant surrounding 3.5%. August is the month with the most significant discrepancy with a difference of 8%. In October and November, both performances approach the minimum difference of the year of only 1%. The year ends with a difference close to 3% favouring the PPP regime. Albeit the variations, both models, EPE and PPP improved their performances in 2015, with an average rise of 1.6% and 1.1%, respectively.

The evolution of hospital efficiency, in 2016, is shown in Figure 6.2. In this year, the performance of PPP hospitals also dominated the EPE ones throughout the year, except for one month, August, where a sudden drop in performance puts their value below the EPE ones.

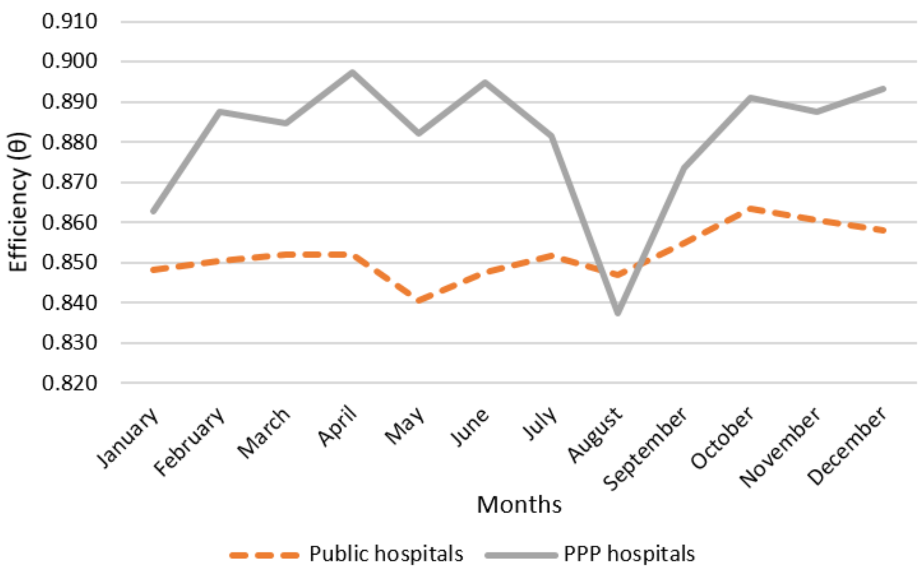

Figure 6.2: Hospitals monthly average performance variation in 2016 (Model I m).
Source: Author

By analyzing the individual performance of each model, it is possible to see some differences. The EPE hospitals do not show considerable variations, having a net improvement of 1% this year. PPP performance has a more lumpy evolution, with a particular focus in August where it experienced a decrease of more than 4%. By looking to the individual performance of the PPP hospitals, it is seen that Hospital de Vila Franca de Xira is the main responsible for this August drop, experiencing a decrease of 10% in efficiency that month. This drop may be related with a maximum occupancy peak verified in that month, where a rate of almost 180% of occupancy was reached [104].[3] Nonetheless, that drop was immediately recovered in the adjacent month, ending the year with an overall gain of 3%.

When comparing both models, in the first half of the year, the PPPs always have a better performance, of at least 1.5%, sometimes reaching 5%. In August, PPPs have a considerable drop in performance, being surpassed by the EPE ones. However, it is the only month where that occurs since in September PPPs increase their performance. Then until the end of the year, they are always above the EPE hospitals. The year ends with a difference of 3.5% in favour of the PPP regime.

Figure 6.3 shows the average efficiency variation in 2017. Like the previous years, PPPs show higher levels of efficiency during the whole year.

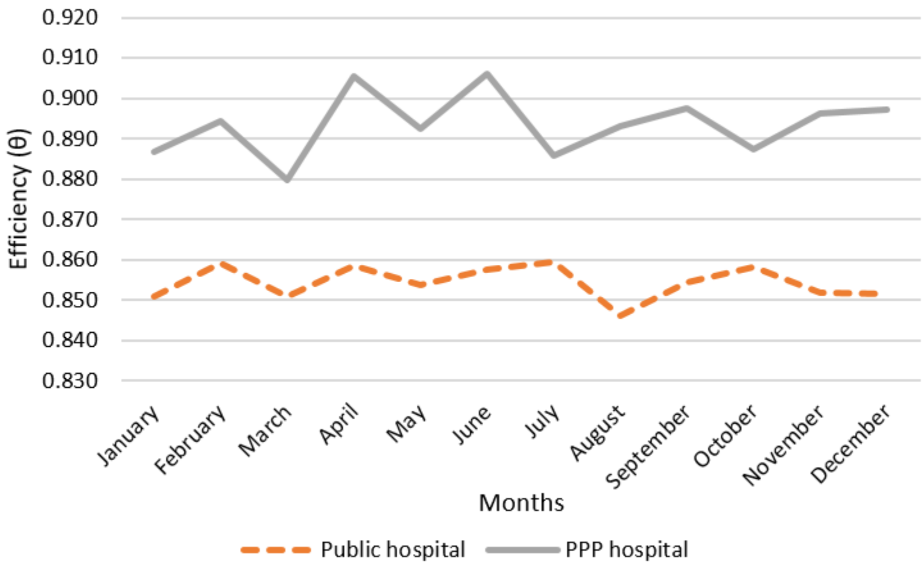

Figure 6.3: Hospitals monthly average performance variation in 2017 (Model I m).
Source: Author

In 2017, both models show limited individual evolution. EPE hospitals' efficiency barely leaves the range of efficiency of 0.850-0.860, ending the year with almost the same performance as when it started it. PPP hospitals have more peaks and troughs always inside a 3% fluctuation, ending the year with an overall improvement of 1%. It is possible to see that, when contrasting both models, throughout the year, the PPP performance is always 3 to 4% greater than the EPE hospitals. It is also worth mentioning that the fluctuations in both models are similar, with the exception for the last two months of the year. In November and December, the efficiency levels of EPE hospitals decrease, whereas in the PPPs it increases. This contrary expansion leads to one of the more substantial differences between the two models, having PPPs at the end of the year a 4.5% higher efficiency level.

[3] Hospital de Vila Franca de Xira internou doentes em refeitórios e casas de banho, Ana Maia, Jornal Público, 2019-05-29, https://www.publico.pt/2019/05/28/sociedade/noticia/hospital-vila-franca-xira-internou-doentes-refeitorios-casas-banho-1874499, Accessed: 2020-05-28

Figure 6.4 represents the evolution of hospitals monthly efficiency in 2018. This year is very similar to 2017 since there are no significant variations in performance, and PPPs always outperform EPE hospitals.

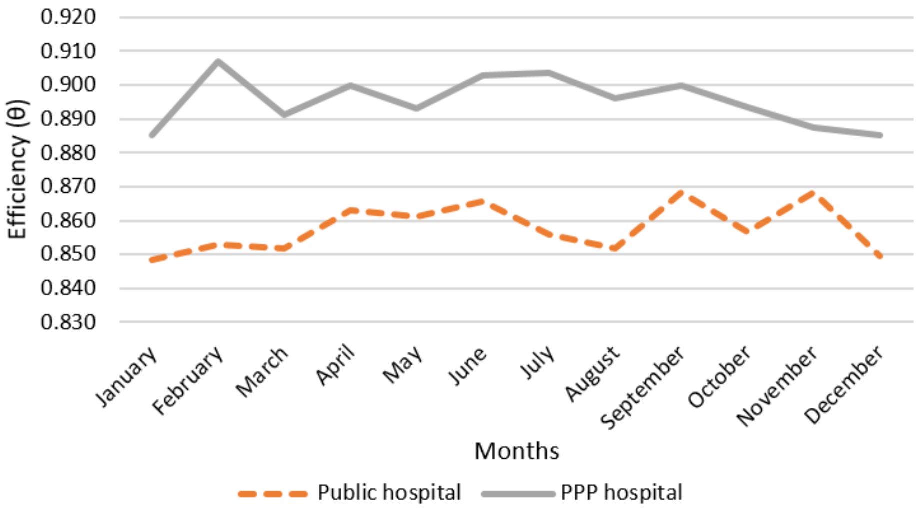

Figure 6.4: Hospitals monthly average performance variation in 2018 (Model I m).
Source: Author

In terms of individual evolution, both models do not present any significant fluctuations. 2018 first half for EPE hospitals is positive, having an almost 2% increase. The second half is more tumultuous, varying their performance up and down 1% nearly every month. They end the year with basically the same efficiency as when they started it. PPP hospitals start the year with a 2% increase, right after, their efficiency, stays around 0.890-0.900, having minor peaks and troughs until September. From September onwards, PPP's performance starts decreasing, closing the year with the same efficiency level of January. When comparing both models, PPPs are always above EPE hospitals, having a higher difference, in February, of more than 5%, and July and August, close to 5%. During the rest of the year the distinction between models is essentially 3.5%, except for November, the month were both models have the closest performance with a difference of only 2%. The year ends with PPPs outperforming by 3.5% EPE hospitals.

Figure 6.5 portrays the evolution of performance for the last year under analysis, 2019. This year only shows the development until September, the last month with available data. Just like all the previous years, PPPs always have better average monthly performances, when comparing to the EPE hospitals.

Taking into consideration their individual progression along the year, EPE hospitals start with a peak in February, after an increase of almost 2%. From March until July there is a steady growth in performance of more than 1%, ending with a drop in August. The year ends with an increase of 1.6%. PPPs' first quarter is precisely the opposite of the EPE hospitals' quarter, starting with a decrease, and right

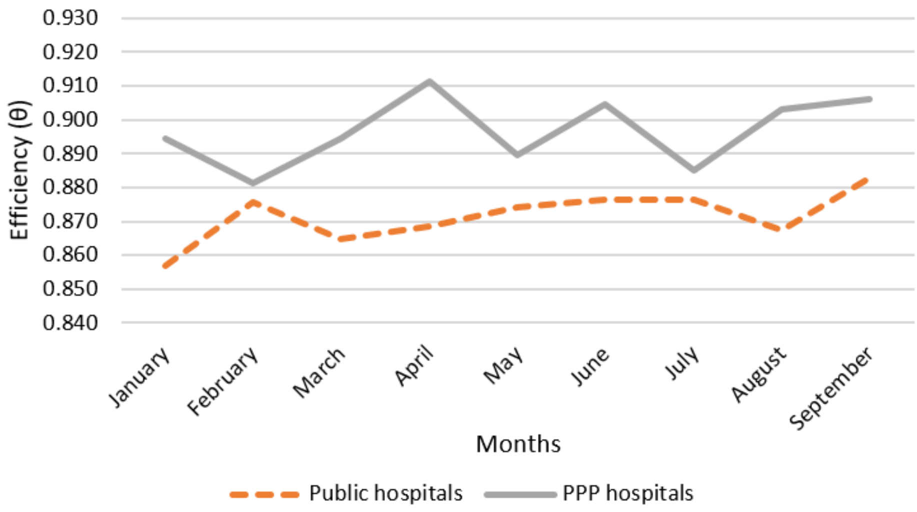

Figure 6.5: Hospitals monthly average performance variation in 2019 (Model I m).
Source: Author

after, from February until April an increase of 3%. From this point there are several peaks and troughs, ending the year with a two-month rise of 2%. In the overall EPE hospitals increased 2.6%, and PPPs 1.2%. It is possible to see that, when comparing both models of hospital management, in February, they converge to their lowest difference, of only 0.5%. After that, there are considerable variations, peaks in April, June, and August achieving differences of 4%, and drops in May and July, with differences of only 1%, always having the PPP regime on top. The last month of the analysis ends with a PPP efficiency level 2% higher than their EPE equivalent.

Figure 6.6, just like the previous figures, represents the monthly average efficiency variation, but in this case, for all the years under analysis (from January 2015 to September 2019). This figure allows to understand if there is seasonality in hospital performance. Seasonality refers to predictable changes that occur over a year.

In the individual evolution of EPE hospitals, it is possible to observe that they have more unsatisfactory performances in January, March, August and December. August performance can be related to being the vacation month, which usually is related to massive decreases in patients. From March to July it is witnessed a more stable performance. The last quarter is marked by having an increase in performances, except for the last month. PPP hospitals, just like the EPE ones, are at their lowest in January. They have three positive peaks in April, June and September, being consistent in the other months. Unlike EPE hospitals, they do not have a drop in efficiency in December. Comparing both management models, it is visible that, on average, PPPs outperform the EPE ones in every month. Since both models are roughly consistent, the difference between them does not leave the range of 2.5-4%.

Analog figures for the other three models of this study, of Figure 6.1 to 6.6, were also made. Due to

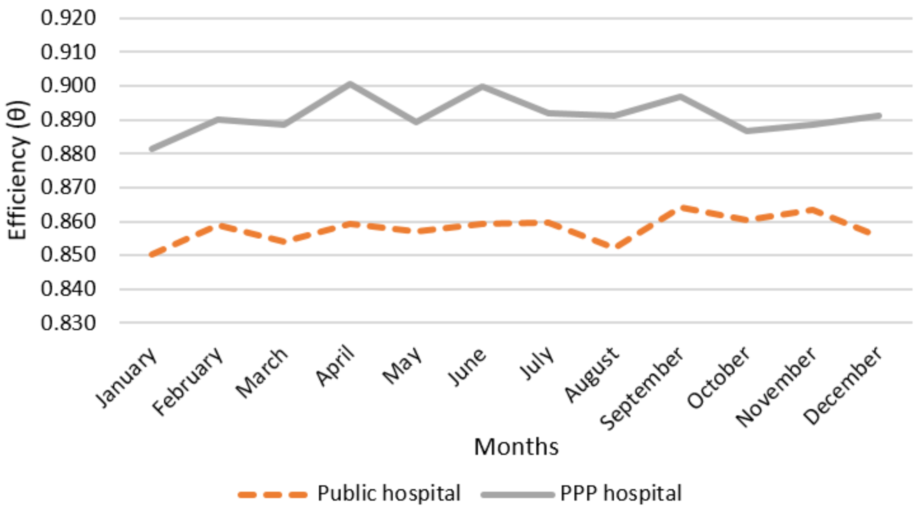

Figure 6.6: Hospitals monthly average performance variation between 2015 and 2019 (Model I m).
Source: Author

the similarity of the results obtained, Figure A.1 to Figure A.18 were added to the Appendix A.

Figure 6.7 gives the average efficiency of each regime for every year of the study (2015 to 2019). By studying this figure, it is possible to understand how hospital performance developed throughout the years. In addition to both models, it was also added in the figure the evolution of the hospital overall efficiency. Hospital overall provides the development of all the Portuguese hospitals' performance (all 28 hospitals taken into consideration into this study).

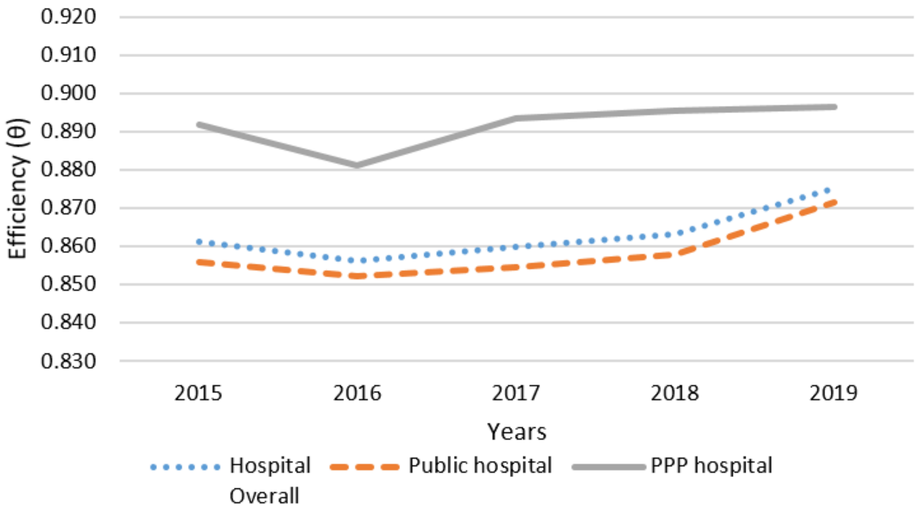

Figure 6.7: Hospitals yearly average performance variation between 2015 and 2019 (Model I m).
Source: Author

Despite, in 2016, having a small drop of less than 0,5%, EPE hospitals increase their performance in every year of the analysis, having an overall increase of 1,5%. PPPs, just like EPE hospitals, increase their efficiencies each passing year, except for 2016. In that year, they experience a drop of more than 1% in efficiency. The net improvement of the PPP's performance, along the studied period, is 0,5%. It is also clearly visible that, PPPs have a better performance in every year when compared with their EPE counterparts. The difference between both models is around 3 to 4%, but in the past three years, they appear to be converging. In 2019, the difference between the two models reached their lowest, 2,6%. The hospital overall evolution is very similar to the EPE ones since they are the majority (24 in 28 hospitals are EPE). Despite being only four, PPPs increase significantly the overall hospital performance. In hospital overall, it is possible to see a tendency of improvement each year, excluding 2016. This tendency can be associated with the better environment that the Portuguese economy has been experiencing in these recent years.

To assess the statistical significance, of the differences in performance, between both models for the several years under analysis, two statistical techniques consistent with the BoD methodology were used. These were the student's t-test for means and Kruskal-Wallis non-parametric test for distributions, the same tests used in the statistical analysis for the variables. These tests determine if there are statistically significant differences between the two groups of hospitals. Both tests will use a 5% significance level, which means that there is only a 5% probability of error in rejecting the null hypothesis that both populations are identical. Table 6.2 provides results for both tests for every year of the analysis (2015 to 2019), and all the years together (2015-2019). The table presents p-values, confidence intervals on the difference of means (Δ), T statistics, and the best performer in the t-test, and only the p-value and best performer for the Kruskal-Wallis test.

Table 6.2: Statistical analysis: Student's t-test for means and Kruskal-Wallis non-parametric test for distributions for the performance values of Model I m.

	2015	**2016**	**2017**	**2018**	**2019**	**2015-2019**
			Student's t-test for means			
p-value	0.0000	0.0000	0.0000	0.0000	0.0017	0.0000
Δ. lower bound	-0.0464	-0.0394	-0.0501	-0.0486	-0.0408	-0.0390
Δ. upper bound	-0.0252	-0.0185	-0.0282	-0.0264	-0.0096	-0.0285
T statistic	-6.6238	-5.4395	-7.0296	-6.6467	-3.1799	-12.7032
Best performer	PPP	PPP	PPP	PPP	PPP	PPP
			Kruskal-Wallis non-parametric test for distributions			
p-value	0.0000	0.0000	0.0000	0.0000	0.0001	0.0000
Best performer	PPP	PPP	PPP	PPP	PPP	PPP

Source: Author.

Similar outputs were obtained for both tests, having all the years presented p-values well below the significance level of 5%. All years have p-values of zero, in both tests, except for the t-test for 2019. Both tests reject the null hypothesis of similar performances of the hospital groups, which means that there is one group that outperforms the other. The best performer is determined by looking to the confidence intervals on the difference of means (Δ), in the student's t-test, and by analyzing the medians, in the Kruskal-Wallis test. Concluding, for this model (Model I m), both tests declare PPP hospitals as the best group for every year under analysis.

6.1.2 Model I M

Model I M is the model that uses all the data from the hospitals but does not include KPIs k_2 and k_5^-, and considers KPI k_1 as desirable ($k_1^+ M$). Figure 6.8 shows the average efficiency of the different types of hospital management for every year of the study (2015 to 2019).

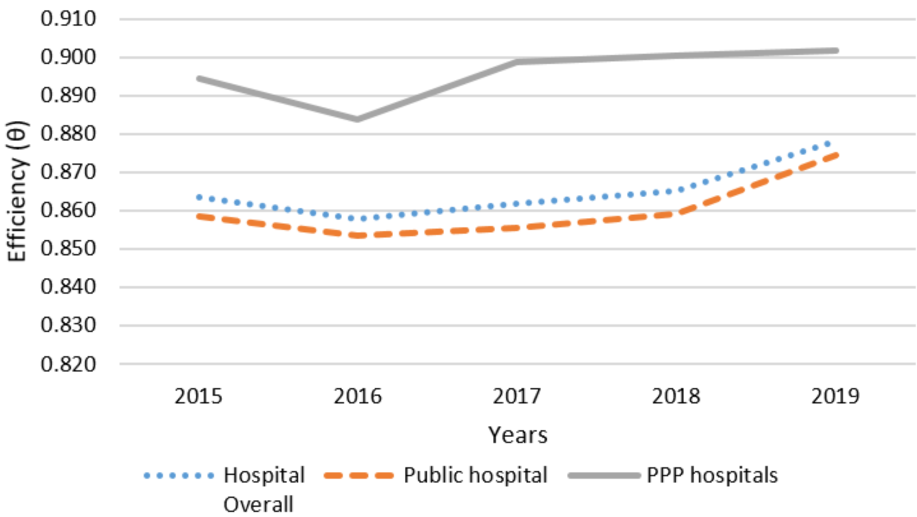

Figure 6.8: Hospitals yearly average performance variation between 2015 and 2019 (Model I M).
Source: Author

This figure is very similar to the one in Figure 6.7 since the development of both hospitals, as well the hospital overall, is identical in both models (Model I m and Model I M). Meaning that the interpretations made for Figure 6.7 also fit Figure 6.8 and the Model I M. In summary, PPPs have a superior performance throughout the whole analysis. However, in the last three years, that superiority has been reducing. There was an increase in performance in every year, except for 2016, where there is a small drop in efficiency in both models.

Table 6.3 presents the results of the statistical tests performed to the efficiencies obtained for every year of the analysis (2015 to 2019, and 2015-2019). Since the efficiency results obtained in this model are similar to the ones obtained in Model I m, Table 6.3 is also identical to the one obtained for that model (6.2).

Bottom line, both tests reject the null hypothesis for all the years. Just as in Model I m, PPPs are the best performer group, for both tests, and all the years under consideration.

Table 6.3: Statistical analysis: Student's t-test for means and Kruskal-Wallis non-parametric test for distributions for the performance values of Model I M.

	2015	2016	2017	2018	2019	2015-2019
Student's t-test for means						
p-value	0.0000	0.0000	0.0000	0.0000	0.0009	0.0000
Δ. lower bound	-0.0474	-0.0415	-0.0546	-0.0528	-0.0433	-0.0415
Δ. upper bound	-0.0243	-0.0191	-0.0318	-0.0297	-0.0112	-0.0305
T statistic	-6.1015	-5.3266	-7.4600	-7.0261	-3.3471	-12.8726
Best performer	PPP	PPP	PPP	PPP	PPP	PPP
Kruskal-Wallis non-parametric test for distributions						
p-value	0.0000	0.0000	0.0000	0.0000	0.0001	0.0000
Best performer	PPP	PPP	PPP	PPP	PPP	PPP

Source: Author.

Ultimately, it is possible to conclude that the distinction in Model I of the KPI k_1 as undesirable or desirable (m or M), do not affect the performances of the hospital units since the results and conclusions obtained are very similar in both models.

6.1.3 Model II m

Model II m is the model that uses all the KPIs, removing the hospitals that have missing data from the database (Hospital de Cascais and Hospital de Loures). It considers KPIs k_1 and k_2 as undesirable ($k_1^- m$ and $k_2^- m$). Starting by analyzing the yearly average efficiencies of the hospital models, it is possible to see, on Figure 6.9, that the PPPs are always above EPE hospitals in terms of performance.

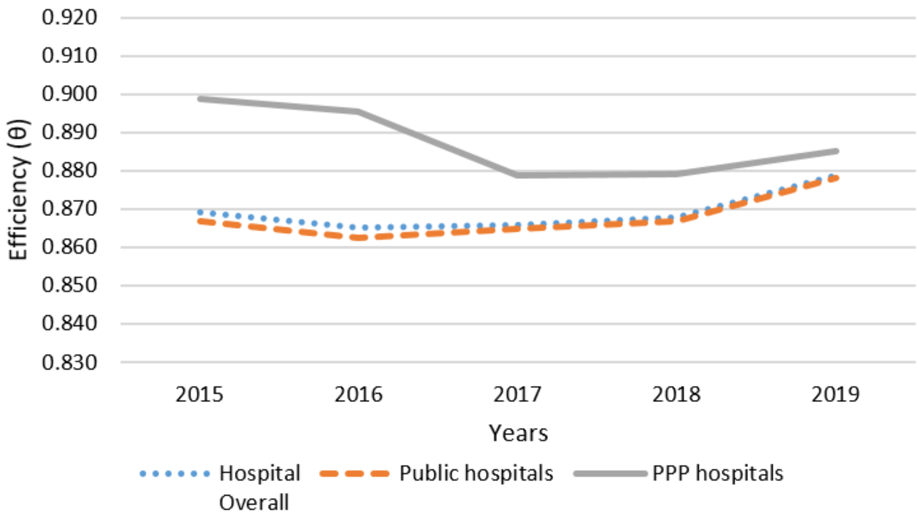

Figure 6.9: Hospitals yearly average performance variation between 2015 and 2019 (Model II m).
Source: Author

By analyzing the individual evolution of EPE hospitals, it is possible to see that they increase their efficiency every year, except for 2016, where they experience a minimal drop of less than 0.5%. EPE hospitals have an overall increase of 1.1%. PPP's performance starts by having their maximum value in 2015, and from that year until 2017, has a drop-in efficiency of 2%. Their performance stays the same in 2018, increasing in 2019 0.6%. PPPs throughout this study had a decrease in efficiency with an overall drop of 1.4% from 2015 to 2019. Comparing both models, it is clear that PPPs outperform EPE hospitals every year, but that superiority has been decreasing. Starting with a favourable difference in performance for the PPPs of more than 3%, in 2015, that difference converges through the years to their minimum of only 0.7%, in 2019. If in Model I, the hospital overall evolution was similar to the EPE ones, because they comprised almost all of the sample (24 in 28 hospitals were EPE), in this model they are practically identical (24 in 26 hospitals are EPE).

Results from the statistical tests made on the average efficiencies, for every year of the analysis can be found in Table 6.4.

Table 6.4: Statistical analysis: Student's t-test for means and Kruskal-Wallis non-parametric test for distributions for the performance values of Model II m.

	2015	2016	2017	2018	2019	2015-2019
			Student's t-test for means			
p-value	0.0000	0.0000	0.0333	0.0695	0.4797	0.0000
Δ. lower bound	-0.0440	-0.0455	-0.0265	-0.0255	-0.0255	-0.0263
Δ. upper bound	-0.0195	-0.0198	-0.0011	0.0010	0.0120	-0.1390
T statistic	-5.1083	-4.9952	-2.1386	-1.8215	-0.7080	-6.3819
Best performer	PPP	PPP	PPP	similar	similar	PPP
			Kruskal-Wallis non-parametric test for distributions			
p-value	0.0000	0.0000	0.0063	0.0106	0.1271	0.0000
Best performer	PPP	PPP	PPP	PPP	similar	PPP

Source: Author.

The outputs obtained for both tests are very similar, being only 2018 the exception. The student's t-test rejects the null hypothesis in 2015, 2016, 2017 and 2015-2019 since their p-values are below the significance level defined as 5%. In the years that the t-test rejects the null hypothesis, it also says that the best performer is the PPP hospitals' group. In 2018 and 2019, the null hypothesis can not be rejected, meaning that both types of hospital management have similar performance results. The Kruskal-Wallis test only differs from the t-test in 2018, rejecting the null hypothesis in that year, being PPP hospitals the best performers.

6.1.4 Model II M

Model II m is the model that uses all the KPIs, removing the hospitals that have missing data from the database (Hospital de Cascais and Hospital de Loures). It considers KPIs k_1 and k_2 as desirable (k_1^+M and k_2^+M). Figure 6.10 shows the yearly efficiency variation between 2015 and 2019. It is possible to see the resemblance between this figure and Figure 6.9 in the evolution of both hospitals' performance.

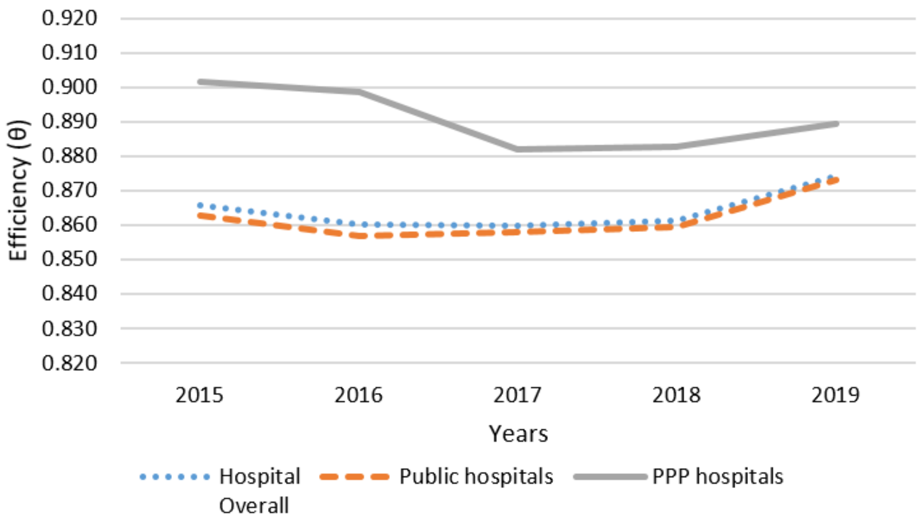

Figure 6.10: Hospitals yearly average performance variation between 2015 and 2019 (Model II M).
Source: Author

EPE hospitals increase their average performance every year of the analysis, except for 2016, where they sustain a minimal drop of 0,4% in efficiency. Their net improvement is 1% until 2019. PPPs decrease their average efficiency from 2015 until 2017, losing two percentage points. Until 2019, they rise their performances, however, not enough to have an overall positive gain in efficiency, dropping 1,3% from 2015 to 2019. Hospital overall evolution is similar to EPE hospitals. The difference between PPPs and EPE hospitals along the years has been declining. In the first couple of years, the difference was about 4%, but then started decreasing, reaching their minimum in 2019 of only 1,6%.

Table 6.5 provides the results of the two statistical tests made for the efficiency values of this model.

Only in 2019, the student's t-test does not reject the null hypothesis, which means that only in that year the there is no significant differences between the efficiencies of the EPE and PPP hospitals. In all the other years, as well as in the overall period (2015 until 2019), the p-values are below the significance level. In all those years, the test also acknowledges, PPPs as the best performers. The Kruskal-Wallis test rejects the null hypothesis for all the years, always giving as the best performer the PPPs. A similar behaviour, previously spotted in Model II m, Figure 6.9 and Table 6.4, can also be seen in this model.

In this model, the distinction between undesirable and desirable KPIs (m or M), have a considerable impact on the results and conclusions obtained. The main difference resides in the lower efficiency

Table 6.5: Statistical analysis: Student's t-test for means and Kruskal-Wallis non-parametric test for distributions for the performance values of Model II M.

	2015	2016	2017	2018	2019	2015-2019
			Student's t-test for means			
p-value	0.0000	0.0000	0.0012	0.0026	0.1336	0.0000
Δ. lower bound	-0.0540	-0.0562	-0.0386	-0.0385	-0.0374	-0.0366
Δ. upper bound	-0.0240	-0.0272	-0.0095	-0.0082	0.0050	-0.0224
T statistic	-5.1147	-5.6689	-3.2587	-3.0305	-1.5053	-8.1450
Best performer	PPP	PPP	PPP	PPP	similar	PPP
			Kruskal-Wallis non-parametric test for distributions			
p-value	0.0000	0.0000	0.0001	0.0005	0.0368	0.0000
Best performer	PPP	PPP	PPP	PPP	PPP	PPP

Source: Author.

values of PPPs in Model m, when comparing with Model M, leading to different results in the statistical tests. Model m always considers as the best performer PPPs, except for 2018 and 2019 in the t-test, and 2019 in the Kruskal-Wallis test. Model M always considers the PPPs as the best performer in the Kruskal-Wallis test, and in the t-test only does not consider PPPs as the best performer in 2019. In Figure 6.11 it can be found a graph that compares all four models of this study, providing the global average efficiencies of each type of hospital (EPE and PPP), and for the overall hospitals. It is seen that the PPP, on average, have a superior performance, between 2015 and 2019, in all four models.

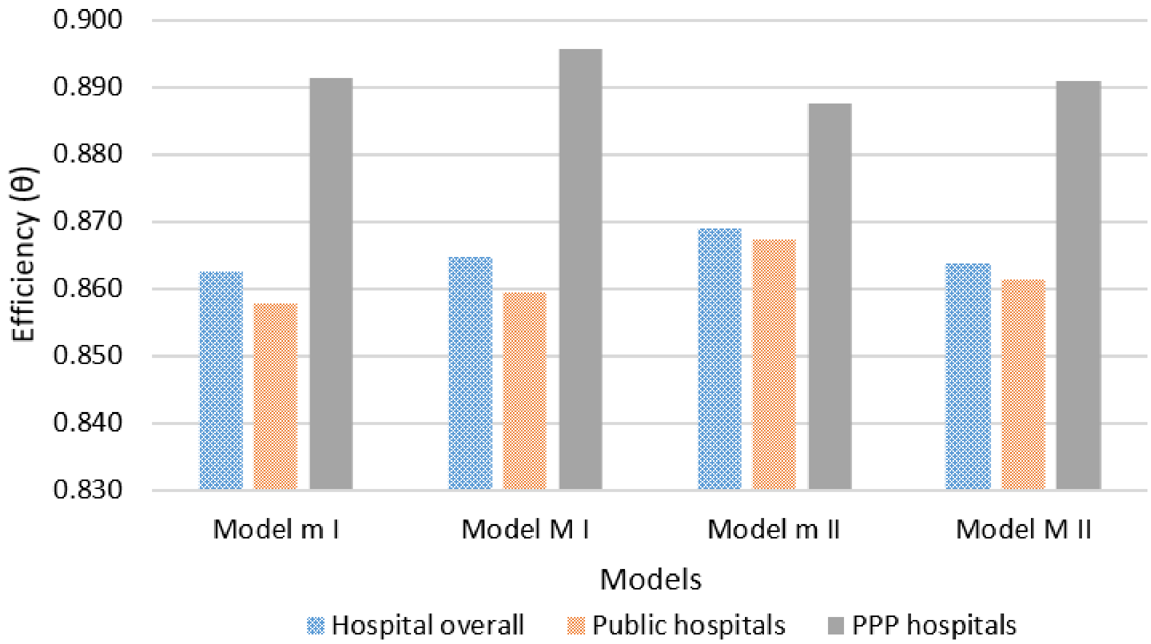

Figure 6.11: Hospitals' global performance across all four models.
Source: Author

In Model II, it is visible the effect of the distinction of KPIs k_1 and k_2 as desirable or undesirable (m or M) on the yearly average performance. Model II m has a considerably smaller gap between both hospitals' performance. Nonetheless, the results obtained with the statistical tests have similar conclusions.

The impact of removing two PPP hospitals from the sample primarily affected the evolution of the performance of the PPP hospitals throughout the years. Whereas in Model I, they tend to grow annually (except for 2016), Model II drops the PPP performance almost every year (except for 2018). The evolution of EPE hospitals does not suffer considerable changes. However, due to the differences in the PPP performances, the gap between both models is considerably smaller in Model II, having repercussions on the statistical analysis. Model I statistical tests reject the null hypothesis for all the years, showing PPPs as best performers. Model II rejects the null hypothesis all the years (with PPP also has the best performer) except for 2019, meaning both hospital models have similar performances.

6.1.5 Global vs partial

The previous sections provided the main findings of the current research using the metatechnology, created by the merger of both samples. Which means that the efficient frontier ($metafrontier = \Psi_{AUB}$)[4] is created using both hospital models. Consequently, each DMU will be projected against that metafrontier. Nonetheless, social performance can be measured relative to the EPE and PPPs own technologies, in that case, each DMU is projected against their own single frontier, Ψ_A and Ψ_B.

Table 6.6: Social performance of EPE hospitals and PPPs, regarding own-group technology and the metatechnology, basic statistics for all four models.

		Performance regarding own-group technology				Performance regarding the metatechnology			
Model	**Group**	**Max {CI}**	**Mean**	**Min {CI}**	**F(τ)**	**Max {CI}**	**Mean**	**Min {CI}**	**F(τ)**
Model I m	EPE	0.859	0.858	0.856	11.70	0.859	0.858	0.856	11.70
	PPP	0.916	0.913	0.909	63.16	0.895	0.891	0.888	39.04
Model I M	EPE	0.861	0.860	0.858	13.30	0.861	0.860	0.858	13.30
	PPP	0.925	0.921	0.917	74.56	0.899	0.896	0.892	46.93
Model II m	EPE	0.869	0.867	0.866	14.55	0.869	0.867	0.866	14.55
	PPP	0.930	0.927	0.923	88.60	0.891	0.887	0.884	21.05
Model II M	EPE	0.863	0.861	0.860	13.30	0.863	0.861	0.860	13.30
	PPP	0.942	0.939	0.935	97.37	0.894	0.891	0.888	38.60

F(τ) is the probability of getting hospitals whose social performance is, at least, equal to τ, here defined as τ = 0.9000.

Source: Author.

Results for the different models are presented in Table 6.6, and can also be found online.[5] This table

[4] A - EPE hospitals; B - PPP hospitals

[5] https://drive.google.com/file/d/1XoKWfN4LJZ8vXarxgyAsI6wxMr06bgq1/view?usp=sharing

contains data concerning the average of both performance analysis, their 90% confidence intervals, $CI_{90\%} = [min\{CI\}, max\{CI\}]$, and the frequency of good social performers, F($\tau = 0.90$). Looking at Table 6.6, PPPs are slightly better performers than EPE hospitals in all the models of the study. The differences between them range from 5 to 8% considering own-group technology, and 2 to 4% considering the metatechnology. In a statistical sense, by looking to the 90% confidence intervals of both hospitals, it is possible to conclude that those differences are considerable because there is no overlapping in the efficiencies. In all models the minimum {CI} of the PPPs is always more significant than the maximum {CI} of the EPE hospitals. The frequency of best performers is also considerably higher for the PPPs, particularly when efficiency is measured against the Ψ_A and Ψ_B frontiers.

It is possible to compare every single frontier with their metafrontier. To do so, the student's t-test and the Kruskal-Wallis test have been, once again, applied.

Table 6.7: Statistical analysis: Student's t-test for means and Kruskal-Wallis non-parametric test for distributions forthe performance values of Ψ_A and Ψ_B, and Ψ_{AUB} frontiers.

	Model I m	**Model I M**	**Model II m**	**Model II M**
		Student's t-test for means		
p-value	0.0357	0.0167	0.0175	0.0134
Δ, lower bound	-0.0058	-0.0066	-0.0055	-0.0066
Δ, upper bound	-0.0002	-0.0007	-0.0005	-0.0008
T statistic	-2.1012	-2.3940	-2.3776	-2.4734
		Kruskal-Wallis non-parametric test for distributions		
p-value	0.1232	0.0738	0.1063	0.1245

In Table 6.7, it appears that both tests have contrary results. Student's t-test, at a significance level of 5%, rejects the null hypothesis that single frontiers overlap their corresponding metafrontier, for all the models. However, in the Kruskal-Wallis test, there is no significant evidence to reject the null hypothesis, meaning there is not an important frontier shift between 2015 and 2019. It is important to note that the values obtained for the EPE hospital's efficiency are the same, whether they are measured through a projection to the metafrontier or their own frontier. This characteristic is a result of having EPE hospitals comprising most of the hospital's sample, creating a small bias in the frontier. That is why the PPPs have a better performance when analyzed against their own single frontier. The existence of this frontier shift requires the use of the Malmquist index for group comparisons to better compare the performance of the hospitals operating under the EPE and PPP regime.

6.2 Malmquist index for group comparison results

This section provides the results of the Malmquist index for group comparison methodology. Results can also be found online.[6] Since this index measures performances using partial frontiers, firstly, it is used the DEA methodology to identify the group frontiers. For this Malmquist index, four sets of efficiency results will be created by the DEA since each hospital group efficiency will be measured relative to their own frontier, Ψ_A and Ψ_B, and the frontier of their counterpart, Ψ_{A_B} and Ψ_{B_A} (A - EPE hospital, B - PPP hospital). With these results, the Malmquist four group comparison can be applied.

As explained in Subsection 4.7, Malmquist index for group comparison is represented by I^{AB}, which when less than one indicates that group A has a better performance than group B, showing the opposite when superior than one. I^{AB} can be decomposed in two sub-components, IE^{AB} that reflects the efficiency of DMUs working in similar conditions, and IF^{AB} that manifests the difference in productivity between the best-practice frontiers of both hospital groups. To recap, when IE^{AB} is less than one, group A has more consistent efficiency values (spread efficiency) than group B. When IF^{AB} is less than one, the productivity of the frontier of group A is higher than the one of group B. Table 6.8 presents the results for I^{AB} and its sub-components.

Table 6.8: Results for the Malmquist index for group comparison analysis.

	Model I m	Model I M	Model II m	Model II M
IE^{AB}	1.064	1.072	1.069	1.090
IF^{AB}	0.977	0.972	0.958	0.949
I^{AB}	1.040	1.042	1.024	1.035

Source: Author.

The results of the hospital comparison should not only be based on the information of the overall index (I^{AB}), but complemented with its sub-indexes (IE^{AB} and IF^{AB}).

In Model I m, the EPE hospital's efficiency spread is 6.4% higher than the PPP hospitals, i.e., their DMUs were located further from their own frontier. EPE hospitals present the best productivity among all samples and have benchmarks 2.3% more efficient than PPPs. In the overall index, it is seen that PPPs outperform EPE hospitals by 4%.

Regarding Model I M, the efficiency spread of EPE hospitals is 7.2% higher than the one from PPPs, and the best practices of EPE hospitals are 2.8% more efficient than the PPPs benchmarks. PPPs are the hospital group with the best overall performance with a difference of 4.2%, the highest in all four models.

Model II m also has EPE hospitals with the least consistency in efficiency, having their efficiency spread 6.9% higher than the PPPs. EPE hospitals have higher productivity levels, outperforming the

[6] https://drive.google.com/file/d/10Rt8KHpaIAST725ij55BAx7N1ZpeDlJp/view?usp=sharing

PPPs by 4.2%. The difference in the overall performance is only 2,4%, the lowest in all four models, favouring the PPPs.

Finally, in Model II M, the hospital's efficiency spread is higher in the EPE hospital's group, being 9% less consistent. The EPE hospital group presents the best productivity in this model, having benchmarks 5.1% more efficient than the PPPs. PPPs are the group with the best overall performance, outperforming the EPE group by 3.5%.

Concluding, despite EPE hospitals showing better productivity levels in all four models, the PPPs also have a considerable lower efficiency spread, therefore, being in the overall performance superior to the EPE hospitals.

7 | Concluding remarks

The beginning of the millennium, the Portuguese health sector, was marked by several reforms affected by the introduction of NPM. This trend had as consequences the separation of functions between financing (by the State), and the provision of hospital services and infrastructure management by a private entity (on a contract basis). As a result, the first wave of PPP hospitals was launched, which consists of four PPP hospitals: Hospital de Loures, Hospital de Cascais, Hospital de Braga, and Hospital de Vila Franca de Xira. The latter was the last hospital to start its operations in 2013. The goal of this thesis is to find out whether those four hospitals outperform traditional public hospitals, to understand if it is advantageous the utilization of PPPs in the hospital sector. Several important conclusions could be drawn from the developed empirical work.

From the global results of the metafrontier, it is seen a considerable difference between social performance levels of both groups, being the performance of PPPs 2 to almost 4% higher (Figure 6.11), on average, than the EPE hospitals. Note also that both groups have room for improvement to achieve efficiency. However, being the PPPs closer to the efficient frontier they still can improve their performance levels, this holds in all adopted models. These results do not translate the conclusions made in Subsection 5.1.3 about Table 5.1, which are most likely flawed since they are based on statistical tests that ignore the heterogeneous environment under which hospitals deliver health care to citizens. Most studies and reports still use these weak tests to compare groups of entities that mostly provide unreliable biased results. This study uses a considerable amount of robust methodologies and tools to reach its conclusions to reduce bias and flaws given by the previous statistical tests.

The yearly monthly analysis, in Model II, has PPPs with always superior performance levels than EPE hospitals, except for one month in 2016. In Model II, there is also the superiority of the PPP hospitals but, especially in 2019, the difference between both models is almost absent, even having EPE hospitals higher levels of performance in some months. In both models, seasonality appears to have no impact on performance levels. However, it is possible to see a trend of converging performances of both models, starting in 2018 in Model I, and in 2017 in Model II. Note that this converging effect results of improvements in the EPE hospital's average performance since PPPs, either maintained their levels or experienced slow increases in comparison.

Statistical tests were made to understand if the differences between performance levels of both groups are considerable. In Model I, the difference was always significant, clearly showing that PPPs consistently outperform the EPE hospitals. In Model II, the impact of the trend detected before, is already enough in 2019 to make the differences between models no longer considerable. Thus, in Model II PPPs outperform EPE hospitals every year, except for 2019 where they have similar performances.

Due to significant differences in the single frontiers, caused by group effects since they are heterogeneous, the Malmquist index for group comparison was used to reach a more reliable conclusion about the comparison of both models. With this index, it was observed that although EPE hospitals have presented the best productivity levels when compared with hospitals under PPP management, they are also the group with the highest efficiency spread. Nonetheless, the overall index indicates that the best hospital performance between both groups is from the PPP hospitals. These results support the previous conclusions where the PPP hospitals also outperformed the EPE ones.

The results obtained in this research give support to the empirical evidence provided from the lessons learned from past researches, in subsections 3.1 and 3.2. The short literature available comparing both models of hospitals provided evidence that PPP hospitals have proper levels of performance and did not show a tendency to have lower performance levels. This study complements those results by giving evidence that PPP hospitals show better performance levels when compared with EPE hospitals.

Regarding the methodology used in this study, the reader must be aware that the BoD results cannot be seen as factors compelling the same reduction/expansion of undesirable/desirable variables, in a radial sense, since the model is non-radial. The BoD model maximizes slacks with the optimization of dual variables (virtual weights), being the performance of the hospital under analysis the best possible (DEA is a benevolent methodology) [37], implying that each hospital has their origins of inefficiency.

In sequence, the Malmquist index calculations are based on distance functions, which are dependent on the results obtained through normal DEA models [98]. Therefore, usually, Malmquist does not incorporate slacks in its analysis. This research constructs the Malmquist index by replacing the conventional DEA, that measures efficiency in a radial sense, with the BoD, that incorporates non-radial forms of inefficiency in the investigation.

Based on Ferreira and Marques [37], the results of this research can affect decisions of policymakers, hospital managers and clinical staff, operational researchers and other academics, and most importantly has implications with the health of citizens.

Policymakers and Regulators. It comes down to the policymakers to decide whether or not a new PPP deal is happening, being revoked, or renovated. For example, when a PPP agreement is coming close to its deadline, the Portuguese government faces a decision between three possibilities: renews the contract with the same private entity, opens a new tender for a new contract, or changes the PPP management to a EPE one. This research's results may help the decisions of policymakers since the PPP hospital that is in activity exhibits better performance levels than the EPE ones. Meaning that there is no reason to change the management nature of the hospital to a EPE one. Furthermore, one could say that an opposite change may be of the state's interest since PPPs outperform the EPE ones. Nevertheless, since the switching costs of changing are considerably high, and the difference between the hospital's performances are not that extensive, a more thorough analysis should be made. Despite, PPPs showing better performance levels, both groups have room for improvement so that they can be good performers. One idea that could help that improvement could be the creation of an annual report by the regulator(s). This report would analyze both types of hospitals, checking for inconsistencies and missing data, and perform a similar benchmarking study to the one made in this thesis.

Hospital Managers and Clinical Staff. The ability to provide quality, in time, appropriated, and safe health care services to all citizens, regardless of their willingness to pay, are the main objectives of all health care providers, whether they are publicly managed or are under a PPP agreement. The clinical staff of both types of hospitals should copy the better practices observed, i.e., the benchmarks belonging to the metafrontier.

Operational Researchers and Other Academics. The tools implemented in this thesis can also be applied to do benchmarks in other fields of study. Measuring the performance of businesses, comparing them against the best performer, or making comparisons between two groups are standard practices to improve performance. The tools used in this study are the most adequate, modern and solid to apply to these types of performance analysis.

Citizens. There is a popular, perhaps uninformed, opinion that PPP hospitals can not deliver health care services with performance levels at least as good as public hospitals. The main idea behind this is that private partners only mission is to maximize profit, not caring if the social component is harmed. This idea is not necessarily true since this study concludes that the PPP hospitals outperform the public ones. Meaning that, the private entities demand for profit does not affect their performance levels.

This research analyzes the performance of four hospitals, that resulted from the first wave of PPP hospitals, comparing them with the public ones. As already mentioned before, this first wave employs a PPP contract know as the Portuguese model. This contract includes the construction and management of the infrastructure, ancillary services, and clinical management. The characteristics of this PPP contract received a great deal of attention because there is much criticism regarding the inclusion of clinical management. According to Silva [105], this inclusion is not advisable given the difficult measurability of clinical activity and the difficulty of identifying and managing the risks present in public services. In fact, because of the critics, the second wave of PPP hospitals no longer intends to include clinical management in the contracts. However, according to the results obtained, there is no empirical evidence supporting those critics. The PPP hospitals, created under the Portuguese model, show better performance levels than the public ones. Therefore, the results of this study provide evidence that there is no empirical support to follow the decision of removing clinical management from the PPP contracts.

7.1 System Limitations and Future Work

This research focuses on comparing the performances of PPP hospitals against the corporatized ones. The main conclusions of this study could be more robust if the sample included a more significant number of years, of variables, and less missing data. These three aspects are correlated to the sample size, which has a considerable impact on non-parametric methods. This study has a substantial number of observations in both models, resulting in a high number of frontier combinations, otherwise, in a sample with a relatively low number of observations, the model would classify a high number of hospitals as efficient. Regarding the variables, this study considers access, clinical safety, care appropriateness, and efficiency variables. Despite having a vast number of variables available, the number of dimensions is more limited. Adding more variables to a study does not necessarily mean better results due to redundancy problems. However, if there is a broad range of dimensions available and reliable, the addition of variables could complement the research. For example, structural quality (related to the hospital facility) and outcomes (overall effects of care) indicators are not considered in this study. It is required to have variables relating to all the dimensions to have a better social performance evaluation, although this study examines a broad range, some more could be used. The lack of availability was not the only reason to not consider more dimensions, some variables had a lot of missing data, in the years of this study, and there were many gaps in the PPP hospital's data. For example, productivity and financial indicators had no information on PPP hospitals, hence the necessity of Model I and Model II. Another aspect that influenced the database was the ranges assumed to replace some gaps of missing data.

All these shortcomings could be addressed by improving the diversity and quality of the benchmarking indicators provided by the ACSS website. A new website could be created to overcome those shortcomings, increasing the diversity in dimensions included and having more complete information. Considering that the results obtained in this thesis have substantial implications from a societal point of view, it is advisable to redo the social performance analysis, performed by the current study, complemented by a broader set of indicators, when available. A posteriori, it is also recommended to compare this research with those future researches.

Finally, note that the reproducibility of these results should be restricted to PPP hospitals that follow the Portuguese model or the Alzira model, where the contract also has included the clinical services. The reason to advise against the use of PPP under the UK model is that their clinical services are provided by a public entity, being only the infrastructure and ancillary services managed by a private entity. In these types of arrangements, it is not expected to have different performances from public hospitals. However, it is possible to have a conflict of interests between both. Nevertheless, as future work, if the second wave of PPP hospitals comes to fruition, despite being planned that those contracts will no longer include clinical services, a similar study should be made considering these new hospitals. Such a study allows for a comparison between both PPP models and could give good empirical evidence on which one is better.

Bibliography

[1] P. P. Barros, S. R. Machado, and A. Simões, *Portugal Health System Review*. European Observatory on Health Systems and Policies, 2011, vol. 13, no. 4.

[2] D. C. Ferreira and R. C. Marques, "Identifying congestion levels, sources and determinants on intensive care units: the Portuguese case," *Health Care Management Science*, vol. 21, no. 3, pp. 348–375, 2018. [Online]. Available: http://dx.doi.org/10.1007/s10729-016-9387-x

[3] PORDATA, Instituto Nacional de Estastística, and Direção Geral de Saúde/Ministério da Saúde, "PorData - Ambiente de Consulta." [Online]. Available: https://www.pordata.pt/DB/Ambiente+de+Consulta/Nova+Consulta

[4] P. P. Barros, *Economia da Saúde*. Almedina, 2019.

[5] M. M. F. Ferreira, "Empenhamento organizacional de profissionais de saúde em hospitais com diferentes modelos de gestão," Ph.D. dissertation, Universidade do Minho, 2005. [Online]. Available: http://hdl.handle.net/1822/4383

[6] D. C. Ferreira and R. C. Marques, "Did the corporatization of Portuguese hospitals significantly change their productivity?" *European Journal of Health Economics*, vol. 16, no. 3, pp. 289–303, 2015.

[7] J. K. Roehrich, M. A. Lewis, and G. George, "Are Public-Private Partnerships a healthy option? A systematic literature review," *Social Science and Medicine*, vol. 113, pp. 110–119, 2014.

[8] K. Buse and A. Harmer, "Power to the partners?: The politics of public-private health partnerships," *Development*, vol. 47, no. 2, pp. 49–56, 2004.

[9] P. P. Barros and X. Martinez-Giralt, "Contractual design and PPPs for hospitals: Lessons for the Portuguese model," *The European Journal of Health Economics*, vol. 10, no. 4, pp. 437–453, 2009.

[10] J. Simões, P. P. Barros, and M. Temido, "Public-Private Partnerships in the Portuguese health sector," *World hospitals and health services: the official journal of the International Hospital Federation*, vol. 46, no. 1, pp. 6–9, 2010.

[11] A. C. d. Campos, "Hospital-empresa: Crónica de um nascimento retardado: Políticas de saúde: Em foco," *Revista portuguesa de saúde pública*, vol. 21, no. 1, pp. 23–33, 2003.

[12] D. C. Ferreira, A. M. Nunes, and R. C. Marques, "Doctors, nurses, and the optimal scale size in the Portuguese public hospitals," *Health Policy*, vol. 122, no. 10, pp. 1093–1100, 2018.

[13] Serviço Nacional de Saúde (SNS), "Doutor António Duarte Arnaut." [Online]. Available: https://www.sns.gov.pt/home/doutor-antonio-duarte-arnaut/

[14] A. M. Nunes and D. C. Ferreira, "The health care reform in Portugal: Outcomes from both the New Public Management and the economic crisis," *International Journal of Health Planning and Management*, vol. 34, no. 1, pp. 196–215, 2019.

[15] D. C. Ferreira and A. M. Nunes, "Technical efficiency of Portuguese public hospitals: A comparative analysis across the five regions of Portugal," *International Journal of Health Planning and Management*, vol. 34, no. 1, pp. e411–e422, 2019.

[16] "Serviço Nacional de Saúde – SNS." [Online]. Available: https://www.sns.gov.pt/sns/servico-nacional-de-saude/

[17] A. M. Nunes and D. C. Ferreira, "Eficiência na assistência hospitalar em Portugal: um estudo comparativo," *Jornal Brasileiro de Economia da Saúde*, vol. 10, no. 3, pp. 278–284, 2018.

[18] A. M. Nunes and D. C. Ferreira, "Reforms in the Portuguese health care sector: Challenges and proposals," *International Journal of Health Planning and Management*, vol. 34, no. 1, pp. e21–e33, 2019.

[19] Instituto Nacional de Estastistíca (INE), "Estatísticas da Saúde 2017," 2019.

[20] D. C. Ferreira and R. C. Marques, "Should inpatients be adjusted by their complexity and severity for efficiency assessment? Evidence from Portugal," *Health Care Management Science*, vol. 19, no. 1, pp. 43–57, 2016.

[21] A. P. C. Monte and L. V. Franca, "Comparação entre sistemas de gestão hospitalar: SPA, SA e EPE, na perspectiva do planeamento e controlo orçamental – Um estudo de caso," p. 29, XIV Congreso Internacional de la Academia de Ciencias Administrativas A. C. (ACACIA), 2010.

[22] L. F. M. Gonçalves, "Análise da eficiência dos hospitais SA e SPA segundo uma abordagem de fronteira de eficiência," Ph.D. dissertation, ISCTE Business School, 2008.

[23] Relatório de Primavera, "Observatório português dos sistemas de saúde," 2006.

[24] Tribunal de Contas, "Auditoria ao programa de Parcerias Público Privadas da saúde - Primeira vaga de hospitais," 2009.

[25] M. R. Reich, "Public-Private Partnerships for public health," *Nature medicine*, vol. 6, no. 6, p. 617, 2000.

[26] World Bank Institute, "Public-Private Partnerships: Reference guide version 1.0," *World Bank*, 2012.

[27] Commission of the European Communities, "Green Paper on Public-Private Partnerships and community law on public contracts and concessions," *Vasa*, 2004. [Online]. Available: http://medcontent.metapress.com/index/A65RM03P4874243N.pdf

[28] J. M. Sarmento, *Parcerias Público-Privadas*. Fundação Francisco Manuel dos Santos, 2016.

[29] C. O. Cruz and R. C. Marques, *O Estado e as Parcerias Público-Privadas*. Sílabo, 2012.

[30] R. C. Marques and D. Silva, "As Parcerias Público-Privadas em Portugal. Lições e recomendações," *Tékhne - Revista de Estudos Politécnicos*, no. 10, pp. 33–50, 2008.

[31] C. P. Rebelo, "As Parcerias Público-Privadas na saúde em Portugal : Análise crítica ." Master Thesis, Instituto Superior de Ciências Sociais e Políticas - Universidade de Lisboa, 2014.

[32] J. P. L. Mendes, "Repensar as Parcerias Público-Privadas no setor da saúde?" Master Thesis, Faculdade de Medicina - Universidade do Porto, 2017.

[33] D. C. Ferreira, "Parcerias Público-Privadas no sector da saúde," 2016.

[34] J. Simões, "As Parcerias Público-Privadas no sector da saúde em Portugal," *Revista Portuguesa de Saúde Pública*, vol. 4, pp. 79–90, 2004.

[35] C. O. Cruz and R. C. Marques, "Contribution to the study of PPP arrangements in airport development, management and operation," *Transport Policy*, vol. 18, no. 2, pp. 392–400, 2011.

[36] C. O. Cruz and R. C. Marques, "Integrating infrastructure and clinical management in PPPs for health care," *Journal of Management in Engineering*, vol. 29, no. 4, pp. 471–481, 2013.

[37] D. C. Ferreira and R. C. Marques, "Public-Private Partnerships in health care services : Do they socially outperform public hospitals regarding quality and access? Evidence from Portugal," *Socio-Economic Planning Sciences*, 2020.

[38] M. Caballer-Tarazona, A. Clemente-Collado, and D. Vivas-Consuelo, “A cost and performance comparison of Public Private Partnership and public hospitals in Spain,” *Health Economics Review*, vol. 6, no. 1, pp. 1–7, 2016.

[39] Direção Geral do Tesouro, “Parcerias público-privadas e concessões–relatório de 2012,” 2012.

[40] Administração Central do Sistema de Saúde (ACSS), “Hospital de Loures.” [Online]. Available: http://www2.acss.min-saude.pt/areaseunidades/parceriassaude/apresentacao/hospitaldeloures/tabid/511/language/pt-pt/default.aspx

[41] Administração Central do Sistema de Saúde (ACSS) , “Hospital de Cascais.” [Online]. Available: http://www2.acss.min-saude.pt/areaseunidades/parceriassaude/apresentacao/hospitalcascais/tabid/509/language/pt-pt/default.aspx

[42] Sistema Nacional Saúde (SNS), “Dados e factos - Hospital de Cascais.” [Online]. Available: https://www.hospitaldecascais.pt/pt/hospital/Paginas/dados.aspx

[43] Administração Central do Sistema de Saúde (ACSS), “Hospital de Braga.” [Online]. Available: http://www2.acss.min-saude.pt/areaseunidades/parceriassaude/apresentacao/hospitaldebraga/tabid/510/language/pt-pt/default.aspx

[44] A. M. A. Oliveira, “Public-Private Partnerships and public hospitals in Portugal: Evidence from delivery data,” Master Thesis, Instituto Superior Técnico, 2015.

[45] Administração Central do Sistema de Saúde (ACSS), “Hospital de Vila Franca de Xira.” [Online]. Available: http://www2.acss.min-saude.pt/areaseunidades/parceriassaude/apresentacao/hospitaldevilafrancadexira/tabid/512/language/pt-pt/default.aspx

[46] M. Torchia, A. Calabrò, and M. Morner, “Public-Private Partnerships in the health care sector: A systematic review of the literature,” *Public Management Review*, vol. 17, no. 2, pp. 236–261, 2015.

[47] M. McKee, N. Edwards, and R. Atun, “Public-Private Partnerships for hospitals,” *Bulletin of the World Health Organization*, vol. 84, pp. 890–896, 2006.

[48] G. P. Barbetta, G. Turati, and A. M. Zago, “Behavioral differences between public and private not-for-profit hospitals in the Italian national health service,” *Health economics*, vol. 1213, no. December 2007, pp. 1207–1213, 2008.

[49] A. Barretta and P. Ruggiero, “Ex-ante evaluation of PFIs within the Italian health-care sector: What is the basis for this PPP?” *Health Policy*, vol. 88, no. 1, pp. 15–24, 2008.

[50] B. Acerete, A. Stafford, and P. Stapleton, “Spanish healthcare Public Private Partnerships: The ‘Alzira model’,” *Critical Perspectives on Accounting*, vol. 22, no. 6, pp. 533–549, 2011.

[51] J. M. Alonso, J. Clifton, and D. Díaz-Fuentes, "Public Private Partnerships for hospitals: Does privatization affect employment?" *Journal of Strategic Contracting and Negotiation*, vol. 2, no. 4, pp. 313–325, 2016.

[52] F. M. Kruse, N. W. Stadhouders, E. M. Adang, S. Groenewoud, and P. P. Jeurissen, "Do private hospitals outperform public hospitals regarding efficiency, accessibility, and quality of care in the European union? A literature review," pp. e434–e453, 2018.

[53] J. Sussex, "Public-Private Partnerships in hospital development: Lessons from the UK's private finance initiative'," *Research in Healthcare Financial Management*, vol. 8, no. 1, pp. 59–77, 2003.

[54] J. Waring, G. Currie, and S. Bishop, "A contingent approach to the organization and management of Public-Private Partnerships: An empirical study of English health care," *Public Administration Review*, vol. 73, no. 2, pp. 313–326, 2013.

[55] G. M. La Forgia and A. Harding, "Public-Private Partnerships and public hospital performance in São Paulo, Brazil," *Health Affairs*, vol. 28, no. 4, pp. 1114–1126, 2009.

[56] L. M. English, "Using Public-Private Partnerships to achieve value for money in the delivery of healthcare in Australia," *International Journal of Public Policy*, vol. 1, no. 1-2, pp. 91–121, 2005.

[57] M. Lim, "Shifting the burden of health care finance: A case study of Public-Private Partnership in Singapore," *Health Policy*, vol. 69, no. 1, pp. 83–92, 2004.

[58] J. Barlow, J. Roehrich, and S. Wright, "Europe sees mixed results from Public-Private Partnerships for building and managing health care facilities and services," *Health Affairs*, vol. 32, no. 1, pp. 146–154, 2013.

[59] G. A. Hodge and C. Greve, "Public-Private Partnerships: An international performance review," *Public Administration Review*, vol. 67, no. 3, pp. 545–558, 2007.

[60] Entidade Reguladora da Saúde, "Estudo de Avaliaçao das Parcerias Público-Privadas na Saúde," 2016. [Online]. Available: https://www.ers.pt/

[61] A. M. Nunes and A. A. Matos, "Análise do desempenho dos hospitais em parceria público privada no serviço nacional de saúde português," *Jornal Brasileiro de Economia da Saúde*, vol. 9, no. 3, pp. 304–309, 2017.

[62] G. Rego, R. Nunes, and J. Costa, "The challenge of corporatisation: The experience of Portuguese public hospitals," *European Journal of Health Economics*, vol. 11, no. 4, pp. 367–381, 2010.

[63] A. M. Nunes, “Reformas na gestão hospitalar: Análise dos efeitos da empresarialização,” Ph.D. dissertation, Instituto Superior de Ciências Sociais e Políticas, 2016.

[64] M. G. Guedes, “A análise da eficiência hospitalar Hospitais EPE,” Ph.D. dissertation, 2014.

[65] P. F. d. L. P. e. C. Figueiredo, “Avaliação de Desempenho dos Hospitais Portugueses : Aplicação de Métodos Não-Paramétricos,” Ph.D. dissertation, 2010.

[66] D. C. Ferreira, R. C. Marques, and A. M. Nunes, “Economies of scope in the health sector: The case of Portuguese hospitals,” *European Journal of Operational Research*, vol. 266, no. 2, pp. 716–735, 2018. [Online]. Available: https://doi.org/10.1016/j.ejor.2017.09.044

[67] D. C. Ferreira and R. C. Marques, “Do quality and access to hospital services impact on their technical efficiency?” *Omega (United Kingdom)*, vol. 86, pp. 218–236, 2019. [Online]. Available: https://doi.org/10.1016/j.omega.2018.07.010

[68] S. Grosskopf and V. Valdmanis, “Measuring hospital performance: A non-parametric approach,” *Journal of Health Economics*, vol. 6, no. 2, pp. 89–107, 1987.

[69] A. Xavier, “Measuring the efficiency of portuguese hospitals with dea: an approach using the general algebraic modeling system,” in *5th WSEAS Congress on Applied Computing conference*, pp. 62-67, 2012.

[70] P. C. Souza, J. H. Scatena, R. T. Kehrig, and B. S. Souza, “Seleção de variáveis inputs e outputs na análise envoltória de dados aplicada a hospitais,” *Revista de Administração em Saúde*, vol. 17, no. 69, 2017.

[71] Tribunal de Contas, “Relatório Global de Avaliação do Modelo de Gestão dos Hospitais do SEE,” vol. I, 2006.

[72] C. P. Barros, A. G. De Menezes, N. Peypoch, B. Solonandrasana, and J. C. Vieira, “An analysis of hospital efficiency and productivity growth using the Luenberger indicator,” *Health care management science*, vol. 11, no. 4, p. 373, 2008.

[73] M. D. R. Giraldes, “Avaliação da eficiência e da qualidade em hospitais: Entidades públicas empresariais e sector público administrativo,” *Acta Medica Portuguesa*, vol. 20, no. 5, pp. 471–490, 2007.

[74] S. Moreira, “Efficiency analysis of public hospitals transformed into public corporations : An application of Data Envelopment Analysis ,” *Economic Bulletin, Spring*, no. 1, 2008.

[75] P. Simões and R. C. Marques, “Performance and congestion analysis of the Portuguese hospital services,” *Central European Journal of Operations Research*, vol. 19, no. 1, pp. 39–63, 2009.

[76] A. P. Harfouche, “Opções políticas em saúde: Efeitos sobre a eficiência hospitalar,” Ph.D. dissertation, Instituto Superior de Ciências Sociais e Políticas, 2010.

[77] A. M. Nunes, “Análise dos resultados de eficiência da política de empresarização de unidades hospitalares em Portugal,” *Jornal Brasileiro de Economia da Saúde*, vol. 9, no. 3, pp. 310–315, 2017.

[78] A. M. Nunes, “Avaliação da eficiência técnica em hospitais portugueses: Efeitos Da empresarização,” *Revista de Administração Hospitalar e Inovação em Saúde (Rahis)*, vol. 14, no. 2, 2017.

[79] A. M. Nunes, D. C. Ferreira, and A. C. Fernandes, “Financial crisis in Portugal: Effects in the health care sector,” *International Journal of Health Services*, vol. 49, no. 2, pp. 237–259, 2019.

[80] R. C. Marques and K. De Witte, “Towards a benchmarking paradigm in European water utilities,” *Public Money and Management*, vol. 30, no. 1, pp. 42–48, 2010.

[81] B. Hollingsworth and S. J. Peacock, *Efficiency Measurement in Health and Health Care*. Routledge International Studies in Health Economics, 2008.

[82] R. Jacobs, P. C. Smith, and A. Street, *Measuring Efficiency in Health Care: Analytic Techniques and Health Policy*. Cambridge University Press, 2006.

[83] Y. A. Ozcan, *Health Care Benchmarking and Performance Evaluation*. Springer, 2008.

[84] W. W. Cooper, L. M. Seiford, and J. Zhu, *Handbook on Data Envelopment Analysis*. Springer Science & Business Media, 2011, vol. 164.

[85] A. Charnes, W. W. Cooper, and E. Rhodes, “Measuring the efficiency of decision making units,” *European Journal of Operational Research*, vol. 2, no. 6, pp. 429–444, 1978.

[86] M. Farrel, “The measurement of productive efficiency,” *Journal of the Royal Statistical Society, Series A, Part 3*, 1957.

[87] C. S. Sarrico, S. M. Hogan, R. G. Dyson, and A. D. Athanassopoulos, “Data Envelopment Analysis and university selection,” *Journal of the Operational Research Society*, vol. 48, no. 12, pp. 1163–1177, 1997.

[88] R. Malhotra, D. Malhotra, and C. A. Lafond, “Analysing financial services industry using Data Envelopment Analysis,” *International Journal of Applied Management Science*, vol. 1, no. 3, pp. 217–246, 2009.

[89] T. Sueyoshi, "Measuring the industrial performance of Chinese cities by Data Envelopment Analysis," *Socio-Economic Planning Sciences*, vol. 26, no. 2, pp. 75–88, 1992.

[90] T. R. Nunamaker and A. Y. Lewin, "Measuring routine nursing service efficiency: A comparison of cost per patient day and Data Envelopment Analysis models/comment," *Health Services Research*, vol. 18, no. 2 Pt 1, p. 183, 1983.

[91] H. D. Sherman, "Hospital efficiency measurement and evaluation: Empirical test of a new technique," *Medical Care*, pp. 922–938, 1984.

[92] D. C. Ferreira, R. C. Marques, and A. M. Nunes, "Pay for performance in health care: A new best practice tariff-based tool using a log-linear piecewise frontier function and a dual–primal approach for unique solutions," *European Journal of Operational Research*, 2019.

[93] D. C. Ferreira and R. C. Marques, "A step forward on order-α robust nonparametric method: Inclusion of weight restrictions, convexity and non-variable returns to scale," *European Journal of Operational Research*, pp. 1–36, 2017.

[94] L. Cherchye, W. Moesen, N. Rogge, and T. V. Puyenbroeck, "An introduction to 'benefit of the doubt' composite indicators," *Social Indicators Research*, vol. 82, no. 1, pp. 111–145, 2007.

[95] L. Cherchye, W. Moesen, N. Rogge, and T. Van Puyenbroeck, "Constructing composite indicators with imprecise data: A proposal," *Expert Systems with Applications*, vol. 38, no. 9, pp. 10 940–10 949, 2011.

[96] D. W. Caves, L. R. Christensen, and W. E. Diewert, "The economic theory of index numbers and the measurement of input, output, and productivity," *Econometrica: Journal of the Econometric Society*, pp. 1393–1414, 1982.

[97] D. C. Ferreira and R. C. Marques, "Malmquist and Hicks-Moorsteen productivity indexes for clusters performance evaluation," *International Journal of Information Technology and Decision Making*, vol. 15, no. 5, pp. 1015–1053, 2016.

[98] A. S. Camanho and R. G. Dyson, "Data Envelopment Analysis and Malmquist indices for measuring group performance," *Journal of Productivity Analysis*, vol. 26, no. 1, pp. 35–49, 2006.

[99] K. K. Yuen, "The two-sample trimmed t for unequal population variances," *Biometrika*, vol. 61, no. 1, pp. 165–170, 1974.

[100] W. H. Kruskal and W. A. Wallis, "Use of ranks in one-criterion variance analysis," *Journal of the American statistical Association*, vol. 47, no. 260, pp. 583–621, 1952.

[101] C. Y. Kung and K. L. Wen, “Applying grey relational analysis and grey decision-making to evaluate the relationship between company attributes and its financial performance - A case study of venture capital enterprises in Taiwan,” *Decision Support Systems*, vol. 43, no. 3, pp. 842–852, 2007.

[102] I. Álvarez, J. Barbero, J. L. Zofío *et al.*, “A Data Envelopment Analysis toolbox for matlab,” Universidad Autónoma de Madrid (Spain), Department of Economic Analysis, Tech. Rep., 2016.

[103] H. Zou, T. Hastie, and R. Tibshirani, “Sparse principal component analysis,” *Journal of computational and graphical statistics*, vol. 15, no. 2, pp. 265–286, 2006.

[104] Entidade Reguladora da Saúde, “Deliberação do conselho de administração da Entidade Reguladora da Saúde.” [Online]. Available: https://www.ers.pt/uploads/document/file/16532/Publica__o_-_vers_o_n_o_confidencial_-_ERS_057_2016.pdf

[105] P. Silva, “Fundamentos e modelos nas parcerias público-privadas na saúde: O estudo dos serviços clínicos,” *Coimbra: Edições Almedina*, 2009.

A | Appendix A

Table A.1: Statistical analysis: hospital overall 2015-2019.

	Entries	Max value	Q3	Mean	Median	Q1	Min value	σ	CV
k_1	1596	21.82	8.32	7.03	6.56	5.43	1.15	2.13	0.30
k_2	1482	51.04	7.18	6.56	6.4	5.63	2.04	2.36	0.36
k_3	1482	13.28	4.81	4.15	4.1	3.26	1.05	1.12	0.27
k_5^-	1482	11 149.00	3 526.00	3 154.95	3 088.0	2 727.00	873.00	762.11	0.24
k_6^+	1596	99.18%	77.58%	69.92%	69.59%	62.25%	0.00%	13.06%	0.19
k_7^+	1596	100.00%	92.68%	85.79%	86.41%	80.53%	3.27%	8.94%	0.10
k_4^+	1596	177.35%	91.39%	86.70%	86.41%	81.84%	62.59%	8.26%	0.10
k_8^-	1596	3.90	1.08	0.86	0.79	0.57	0.00	0.42	0.49
k_{10}^+	1596	100.00%	86.08%	80.41%	81.99%	76.51%	0.00%	10.13%	0.13
k_{11}^-	1596	4.89%	2.40%	2.01%	1.98%	1.60%	0.00%	0.67%	0.33
k_{12}^-	1596	17.93%	8.96%	7.75%	7.70%	6.66%	0.00%	2.07%	0.27
k_{13}^-	1596	40.54%	10.89%	9.67%	9.72%	8.53%	0.00%	2.32%	0.24
k_{14}^-	1596	13.89%	4.26%	3.50%	3.41%	2.66%	0.00%	1.27%	0.36
k_{15}^-	1596	60.07%	31.37%	28.33%	28.00%	24.63%	12.77%	5.49%	0.19
k_9^+	1596	100.00%	65.00%	46.22%	44.44%	27.03%	0.00%	24.83%	0.54
k_{16}^+	1596	97.30%	69.58%	63.63%	64.18%	58.47%	17.14%	8.53%	0.13
k_{17}^-	1596	0.0286	0.0017	0.0013	0.0000	0.0000	0.0000	0.0028	2.12
k_{20}^-	1596	77.78%	33.33%	27.38%	25.93%	20.55%	0.00%	11.34%	0.41
k_{21}^-	1596	40.91%	14.71%	12.52%	11.58%	9.35%	0.00%	5.25%	0.42
k_{22}^-	1596	0.1229	0.0000	0.0002	0.0000	0.0000	0.0000	0.0033	13.81
k_{23}^-	1596	16.67	0.33	0.21	0.00	0.00	0.00	0.55	2.66
k_{18}^-	1596	8.08	1.20	0.75	0.55	0.00	0.00	0.86	1.14
k_{19}^-	1596	33.33%	2.14%	1.58%	0.00%	0.00%	0.00%	2.91%	1.84

Q3 - stands for the 75th percentile; Q1 - stands for the 25th percentile; σ - is the standard deviation; CV - which measures heterogeneity among observations, is calculated by σ dividing by the mean.
Source: author.

Table A.2: Statistical analysis: EPE hospitals 2015-2019.

	Entries	Max value	Q3	Mean	Median	Q1	Min value	σ	CV
k_1	1368	21.82	8.24	7.02	6.54	5.44	1.69	2.11	0.30
k_2	1368	51.04	7.12	6.51	6.35	5.58	2.04	2.42	0.37
k_3	1368	13.28	4.71	4.02	3.95	3.22	1.05	1.01	0.25
k_5^-	1368	11 149.00	3 571.00	3 209.21	3 165.50	2 782.25	873.00	758.52	0.24
k_6^+	1368	99.18%	79.12%	71.05%	70.00%	62.75%	34.51%	12.22%	0.17
k_7^+	1368	100.00%	91.27%	85.38%	85.92%	80.05%	58.58%	8.52%	0.10
k_4^+	1368	144.25%	90.46%	85.86%	85.76%	80.90%	62.59%	7.85%	0.09
k_8^-	1368	3.90	1.12	0.92	0.86	0.65	0.00	0.41	0.44
k_{10}^+	1368	100.00%	85.77%	79.70%	81.28%	75.64%	0.00%	10.68%	0.13
k_{11}^-	1368	4.89%	2.49%	2.05%	2.02%	1.61%	0.00%	0.69%	0.34
k_{12}^-	1368	17.93%	9.11%	7.83%	7.79%	6.70%	0.00%	2.16%	0.28
k_{13}^-	1368	40.54%	11.03%	9.70%	9.75%	8.50%	0.00%	2.43%	0.25
k_{14}^-	1368	13.89%	4.46%	3.61%	3.57%	2.75%	0.00%	1.31%	0.36
k_{15}^-	1368	60.07%	31.64%	28.69%	28.27%	25.00%	14.69%	5.54%	0.19
k_9^+	1368	100.00%	62.96%	45.01%	43.48%	26.32%	0.00%	24.10%	0.54
k_{16}^+	1368	97.30%	68.21%	62.47%	62.86%	57.45%	17.14%	8.43%	0.13
k_{17}^-	1368	0.0286	0.0016	0.0010	0.0000	0.0000	0.0000	0.0021	2.02
k_{20}^-	1368	77.78%	34.06%	28.19%	26.87%	21.13%	0.00%	11.56%	0.41
k_{21}^-	1368	40.91%	15.15%	12.94%	12.04%	9.56%	0.00%	5.44%	0.42
k_{22}^-	1368	0.1229	0.0000	0.0003	0.0000	0.0000	0.0000	0.0036	13.71
k_{23}^-	1368	16.67	0.31	0.20	0.00	0.00	0.00	0.58	2.94
k_{18}^-	1368	4.54	1.15	0.70	0.51	0.00	0.00	0.78	1.11
k_{19}^-	1368	33.33%	1.92%	1.48%	0.00%	0.00%	0.00%	2.93%	1.98

Q3 - stands for the 75th percentile; Q1 - stands for the 25th percentile; σ - is the standard deviation; CV - which measures heterogeneity among observations, is calculated by σ dividing by the mean.
Source: author.

Table A.3: Statistical analysis: PPP hospitals 2015-2019.

	Entries	Max value	Q3	Mean	Median	Q1	Min value	σ	CV
k_1	228	12.98	9.13	7.08	6.76	5.27	1.15	2.20	0.31
k_2	114	10.62	8.13	7.21	7.31	6.41	3.15	1.28	0.18
k_3	114	7.85	6.84	5.72	6.18	4.58	3.12	1.21	0.21
k_5^-	114	4 563.18	2 802.55	2 503.87	2 392.00	2 178.00	1 136.00	432.60	0.17
k_6^+	228	84.50%	71.68%	63.11%	66.57%	59.70%	0.00%	15.61%	0.25
k_7^+	228	98.71%	94.77%	88.27%	90.37%	84.45%	3.27%	10.82%	0.12
k_4^+	228	177.35%	95.70%	91.73%	91.16%	86.28%	70.99%	8.82%	0.10
k_8^-	228	1.69	0.54	0.48	0.43	0.30	0.00	0.29	0.61
k_{10}^+	228	100.00%	86.54%	84.70%	84.78%	82.53%	74.59%	3.53%	0.04
k_{11}^-	228	2.82%	2.11%	1.79%	1.87%	1.51%	0.23%	0.47%	0.26
k_{12}^-	228	10.64%	8.09%	7.25%	7.23%	6.42%	0.73%	1.29%	0.18
k_{13}^-	228	15.34%	10.33%	9.47%	9.61%	8.62%	1.02%	1.42%	0.15
k_{14}^-	228	4.77%	3.31%	2.82%	2.83%	2.33%	1.07%	0.70%	0.25
k_{15}^-	228	45.34%	29.16%	26.19%	25.97%	22.69%	12.77%	4.63%	0.18
k_9^+	228	100.00%	77.38%	53.52%	52.79%	34.61%	0.00%	27.70%	0.52
k_{16}^+	228	81.85%	74.48%	70.60%	70.92%	66.92%	51.90%	5.17%	0.07
k_{17}^-	228	0.0266	0.0039	0.0031	0.0000	0.0000	0.0000	0.0051	1.64
k_{20}^-	228	50.00%	27.55%	22.54%	22.43%	17.67%	0.00%	8.46%	0.38
k_{21}^-	228	24.00%	11.53%	10.05%	9.74%	8.40%	3.25%	2.92%	0.29
k_{22}^-	228	0.0022	0.0000	0.0001	0.0000	0.0000	0.0000	0.0004	3.16
k_{23}^-	228	1.49	0.43	0.26	0.15	0.00	0.00	0.30	1.15
k_{18}^-	228	8.08	1.61	1.08	0.84	0.00	0.00	1.19	1.09
k_{19}^-	228	17.07%	2.94%	2.22%	1.50%	0.00%	0.00%	2.75%	1.24

Q3 - stands for the 75th percentile; Q1 - stands for the 25th percentile; σ - is the standard deviation; CV - which measures heterogeneity among observations, is calculated by σ dividing by the mean.
Source: author.

Table A.4: Correlation coefficients (r) for all KPIs.

	k_1	k_2	k_3	k_5^-	k_6^+	k_7^+	k_4^+	k_8^-	k_{10}^+	k_{11}^-	k_{12}^-	k_{13}^-	k_{14}^-	k_{15}^-	k_9^+	k_{16}^+	k_{17}^-	k_{18}^-	k_{19}^-
k_1	1	0.29	0.84	-0.42	-0.12	-0.12	0.19	0.03	0.07	-0.17	-0.14	-0.08	0.15	-0.10	-0.01	-0.06	0.04	-0.22	-0.21
k_2	0.29	1	0.37	-0.31	0.02	-0.01	0.04	-0.01	-0.03	-0.01	0.09	0.15	-0.03	0.03	-0.02	0.00	0.11	0.06	-0.03
k_3	0.84	0.37	1	-0.55	-0.16	-0.12	0.20	-0.06	0.09	-0.17	-0.12	-0.03	0.12	-0.11	-0.10	0.03	0.08	-0.27	-0.27
k_5^-	-0.42	-0.31	-0.55	1	0.07	-0.08	-0.10	0.04	-0.09	-0.01	0.00	-0.03	0.05	0.08	-0.04	-0.01	-0.11	0.16	0.17
k_6^+	-0.12	0.02	-0.16	0.07	1	-0.01	-0.16	0.08	-0.15	-0.04	-0.05	-0.09	0.04	0.10	0.08	-0.28	0.01	-0.06	-0.07
k_7^+	-0.12	-0.01	-0.12	-0.08	-0.01	1	-0.06	-0.27	-0.23	0.10	0.02	0.04	-0.36	0.07	0.30	-0.15	-0.07	-0.06	-0.17
k_4^+	0.19	0.04	0.20	-0.10	-0.16	-0.06	1	-0.07	0.18	0.01	-0.01	0.02	0.14	-0.09	-0.07	0.20	0.03	0.06	0.11
k_8^-	0.03	-0.01	-0.06	0.04	0.08	-0.27	-0.07	1	0.17	0.04	0.15	0.07	0.37	-0.10	-0.15	-0.10	-0.01	0.01	0.04
k_{10}^+	0.07	-0.03	0.09	-0.09	-0.15	-0.23	0.18	0.17	1	0.01	0.04	0.09	0.12	-0.11	-0.09	0.56	0.01	0.07	0.13
k_{11}^-	-0.17	-0.01	-0.17	-0.01	-0.04	0.10	0.01	0.04	0.01	1	0.81	0.42	-0.04	0.02	0.03	-0.03	-0.06	0.18	0.18
k_{12}^-	-0.14	0.09	-0.12	0.00	-0.05	0.02	-0.01	0.15	0.04	0.81	1	0.65	0.06	0.01	-0.02	-0.01	-0.05	0.24	0.19
k_{13}^-	-0.08	0.15	-0.03	-0.03	-0.09	0.04	0.02	0.07	0.09	0.42	0.65	1	0.12	0.04	-0.06	0.01	0.00	0.15	0.14
k_{14}^-	0.15	-0.03	0.12	0.05	0.04	-0.36	0.14	0.37	0.12	-0.04	0.06	0.12	1	-0.06	-0.28	-0.07	0.11	0.08	0.11
k_{15}^-	-0.10	0.03	-0.11	0.08	0.10	0.07	-0.09	-0.10	-0.11	0.02	0.01	0.04	-0.06	1	0.11	0.00	0.00	0.08	-0.04
k_9^+	-0.01	-0.02	-0.10	-0.04	0.08	0.30	-0.07	-0.15	-0.09	0.03	-0.02	-0.06	-0.28	0.11	1	-0.07	0.11	-0.09	-0.18
k_{16}^+	-0.06	0.00	0.03	-0.01	-0.28	-0.15	0.20	-0.10	0.56	-0.03	-0.01	0.01	-0.07	0.00	-0.07	1	-0.04	0.23	0.15
k_{17}^-	0.04	0.11	0.08	-0.11	0.01	-0.07	0.03	-0.01	0.01	-0.06	-0.05	0.00	0.11	0.00	0.11	-0.04	1	-0.07	-0.03
k_{18}^-	-0.22	0.06	-0.27	0.16	-0.06	-0.06	0.06	0.01	0.07	0.18	0.24	0.15	0.08	0.08	-0.09	0.23	-0.07	1	0.30
k_{19}^-	-0.21	-0.03	-0.27	0.17	-0.07	-0.17	0.11	0.04	0.13	0.18	0.19	0.14	0.11	-0.04	-0.18	0.15	-0.03	0.30	1

Green cells identify $r \geq 70$, and yellow cells identify $50 \leq r \leq 70$.
Source: author.

Table A.5: PCA results.

Model I				Model II			
m		M		m		M	
Explained	**Coefficient**	**Explained**	**Coefficient**	**Explained**	**Coefficient**	**Explained**	**Coefficient**
97.50	0.25	97.37	0.10	97.85	0.23	97.57	0.10
0.73	0.33	0.77	0.34	0.60	0.29	0.68	0.03
0.33	0.25	0.33	0.25	0.30	0.25	0.32	0.27
0.27	0.30	0.28	0.31	0.20	0.23	0.23	0.25
0.21	0.27	0.22	0.28	0.18	0.28	0.20	0.29
0.18	0.16	0.20	0.17	0.16	0.30	0.18	0.32
0.15	0.28	0.16	0.29	0.13	0.25	0.15	0.27
0.14	0.21	0.15	0.21	0.13	0.26	0.14	0.28
0.12	0.27	0.13	0.27	0.10	0.19	0.12	0.20
0.10	0.26	0.11	0.27	0.08	0.25	0.09	0.26
0.07	0.23	0.08	0.24	0.06	0.24	0.06	0.26
0.06	0.20	0.06	0.21	0.06	0.22	0.06	0.23
0.05	0.33	0.06	0.34	0.05	0.15	0.05	0.16
0.04	0.23	0.04	0.23	0.04	0.18	0.04	0.20
0.03	0.24	0.03	0.25	0.03	0.31	0.04	0.33
				0.03	0.21	0.03	0.22
				0.02	0.22	0.02	0.24

Source: author.

Table A.6: Global performance basic statistics (model I M).

2015/2019	Hospital Overall	EPE Hospitals	PPP Hospitals
Entries	1596	1368	228
Maximum value	1.000	1.000	0.972
Maximum {CI}	0.866	0.861	0.899
Percentile 75%	0.890	0.884	0.918
Mean	0.865	0.860	0.896
Median	0.865	0.860	0.896
Percentile 25%	0.836	0.832	0.875
Minimum {CI}	0.863	0.858	0.892
Minimum value	0.734	0.734	0.788
F(τ =0.90)	18.11	13.30	46.93
Standard deviation (σ)	0.041	0.040	0.033
Coeficient of variation (CV)	0.05	0.05	0.04

CV

measures heterogeneity among observations. is calculated by σ dividing by the mean.
CI - confidence interval
Source: Author.

Table A.7: Global performance basic statistics (model II m).

2015/2019	Hospital Overall	EPE Hospitals	PPP Hospitals
Entries	1482	1368	114
Maximum value	1.000	1.000	0.958
Maximum {CI}	0.870	0.869	0.891
Percentile 75%	0.889	0.887	0.898
Mean	0.869	0.867	0.887
Median	0.867	0.865	0.889
Percentile 25%	0.845	0.844	0.877
Minimum {CI}	0.868	0.866	0.884
Minimum value	0.746	0.746	0.820
F(τ =0.90)	15.05	14.55	21.05
Standard deviation (σ)	0.033	0.033	0.021
Coeficient of variation (CV)	0.04	0.04	0.02

CV

measures heterogeneity among observations, is calculated by σ dividing by the mean.

CI - confidence interval

Source: Author.

Table A.8: Global performance basic statistics (model II M).

2015/2019	Hospital Overall	EPE Hospitals	PPP Hospitals
Entries	1482	1368	114
Maximum value	1.000	1.000	0.949
Maximum {CI}	0.865	0.863	0.894
Percentile 75%	0.889	0.885	0.907
Mean	0.864	0.861	0.891
Median	0.865	0.862	0.891
Percentile 25%	0.838	0.835	0.877
Minimum {CI}	0.862	0.860	0.888
Minimum value	0.737	0.737	0.796
F(τ =0.90)	15.25	13.30	38.60
Standard deviation (σ)	0.038	0.038	0.022
Coeficient of variation (CV)	0.04	0.04	0.03

CV

measures heterogeneity among observations, is calculated by σ dividing by the mean.

CI - confidence interval

Source: Author.

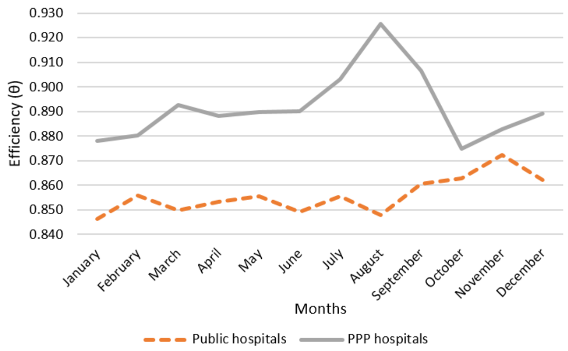

Figure A.1: Hospitals monthly average performance variation in 2015 (Model I M).
Source: Author

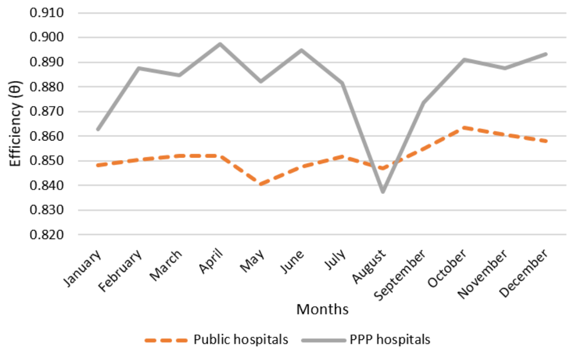

Figure A.2: Hospitals monthly average performance variation in 2016 (Model I M).
Source: Author

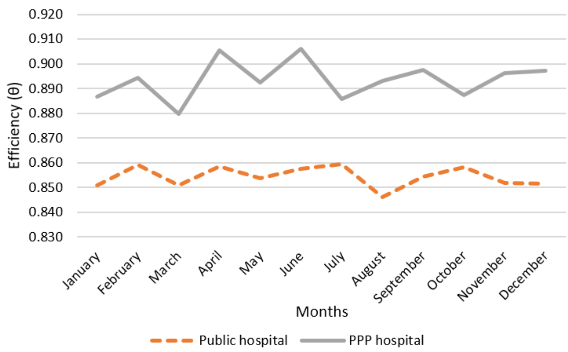

Figure A.3: Hospitals monthly average performance variation in 2017 (Model I M).
Source: Author

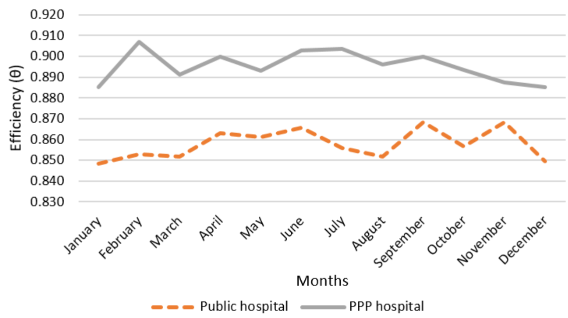

Figure A.4: Hospitals monthly average performance variation in 2018 (Model I M).
Source: Author

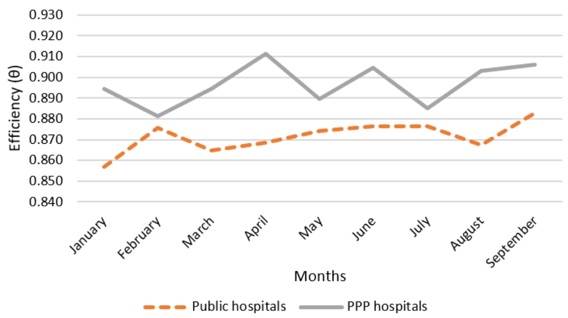

Figure A.5: Hospitals monthly average performance variation in 2019 (Model I M).
Source: Author

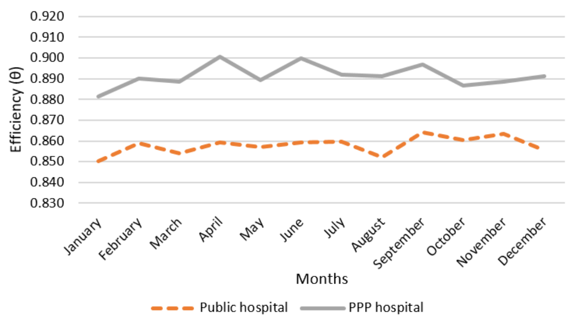

Figure A.6: Hospitals monthly average performance variation between 2015 and 2019 (Model I M).
Source: Author

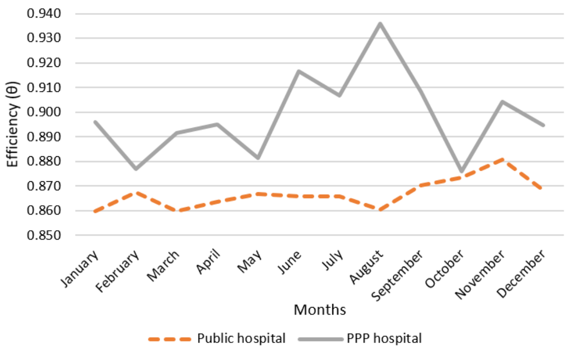

Figure A.7: Hospitals monthly average performance variation in 2015 (Model II m).
Source: Author

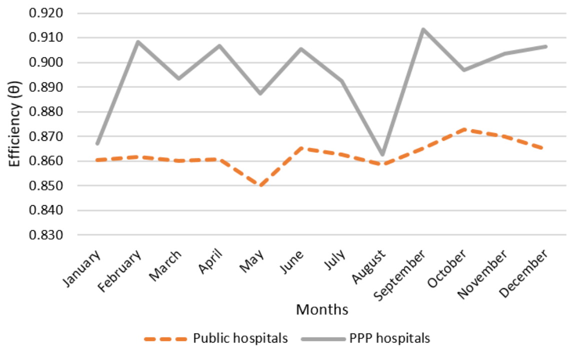

Figure A.8: Hospitals monthly average performance variation in 2016 (Model II m).
Source: Author

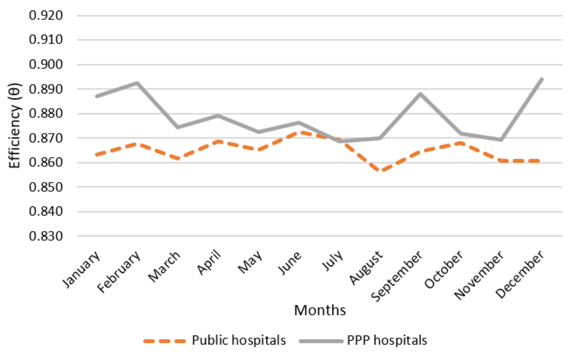

Figure A.9: Hospitals monthly average performance variation in 2017 (Model II m).
Source: Author

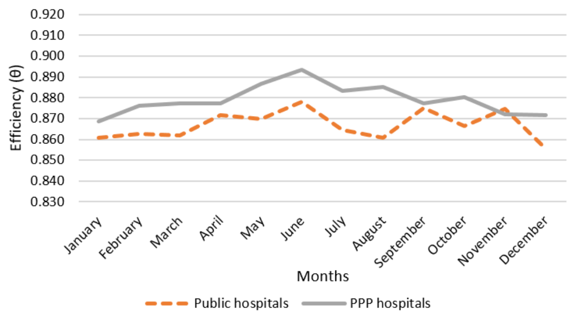

Figure A.10: Hospitals monthly average performance variation in 2018 (Model II m).
Source: Author

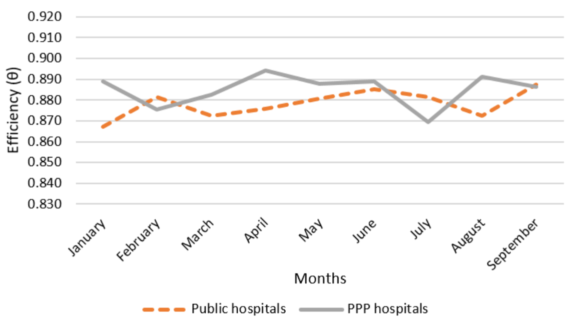

Figure A.11: Hospitals monthly average performance variation in 2019 (Model II m).
Source: Author

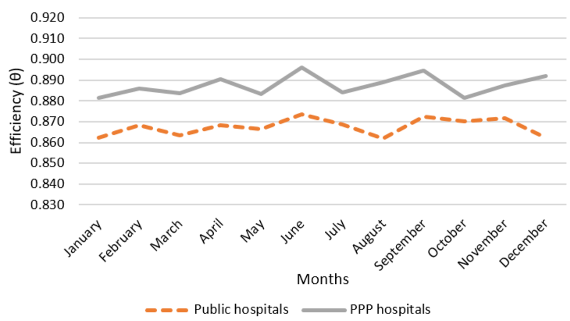

Figure A.12: Hospitals monthly average performance variation between 2015 and 2019 (Model II m).
Source: Author

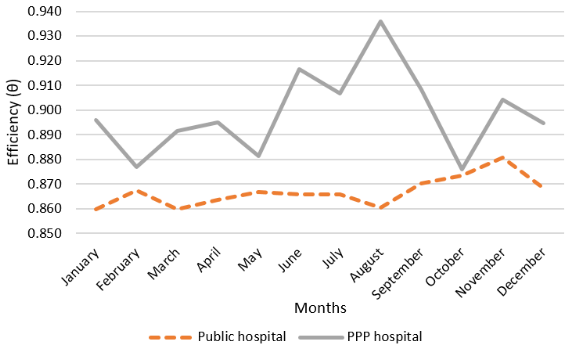

Figure A.13: Hospitals monthly average performance variation in 2015 (Model II M).
Source: Author

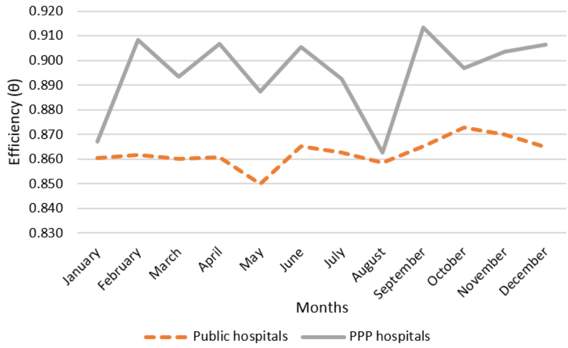

Figure A.14: Hospitals monthly average performance variation in 2016 (Model II M).
Source: Author

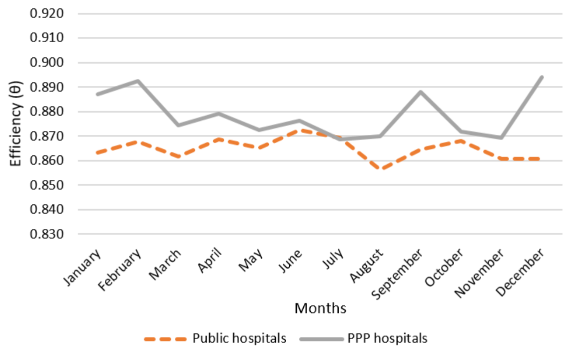

Figure A.15: Hospitals monthly average performance variation in 2017 (Model II M).
Source: Author

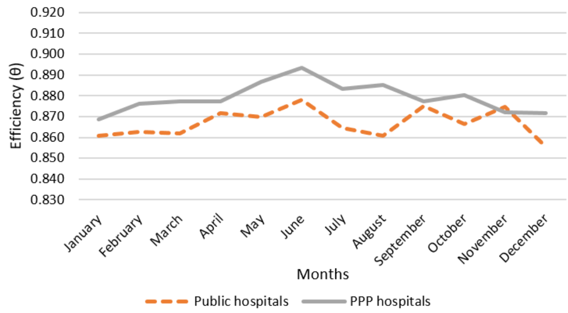

Figure A.16: Hospitals monthly average performance variation in 2018 (Model II M).
Source: Author

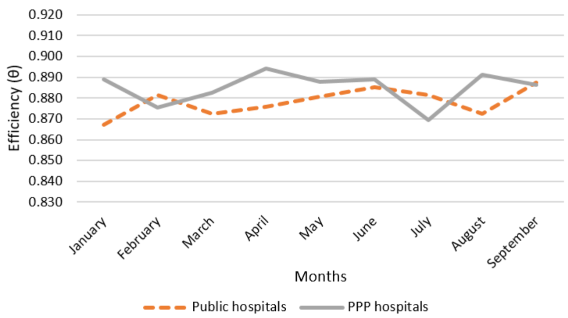

Figure A.17: Hospitals monthly average performance variation in 2019 (Model II M).
Source: Author

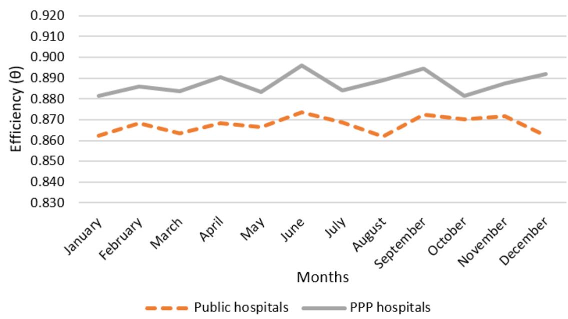

Figure A.18: Hospitals monthly average performance variation between 2015 and 2019 (Model II M).
Source: Author