D 2014

U.PORTO
FEUP FACULDADE DE ENGENHARIA
UNIVERSIDADE DO PORTO

INTEGRATION OF THE ASSESSMENTS OF ACCESS TO HEALTH CARE AND OF COMPETITION BETWEEN PROVIDERS INTO A NEW METHOD TO IDENTIFY TARGET GEOGRAPHIC MARKETS

PIERRE POLZIN
TESE DE DOUTORAMENTO APRESENTADA
À FACULDADE DE ENGENHARIA DA UNIVERSIDADE DO PORTO EM
ENGENHARIA INDUSTRIAL E GESTÃO

Integration of the assessments of access to health care and of competition between providers into a new method to identify target geographic markets

Pierre Polzin

A thesis submitted to Faculdade de Engenharia da Universidade do Porto for the Doctoral Degree in Industrial Engineering and Management

Supervisors:
José Luís Borges
António Fernando Coelho

Faculdade de Engenharia da Universidade do Porto
2014

ABSTRACT

This thesis proposes a new method to identify fine-resolution target geographic markets with low patients' access to health care and no potential competition problems for firms and public administrations to promote new supply. The proposed method corrects limitations of previous methods, enhancing the assessments of access and of competition. Moreover, it innovatively integrates the two kinds of assessment. This methodological integration makes it possible to process a comprehensive set of data, in order to, on one hand, identify geographic areas for a comparatively higher chance of firms' survival and expansion post-market entry. On the other hand, this method can be used to support public policies that aim at reducing access inequalities and increasing health care supply to disadvantaged populations.

The first stage of the method's development consists of the construction of indices with principal component analysis to capture disparities in access in different geographic areas. This stage also comprises an improved modeling of the potential travel patterns of patients to health care supply points. These enhancements are used in turn as a basis for the proposed method, namely the extended kernel density four-step floating catchment area (EKD4SFCA) method, which combines competition assessment into the analysis. The presentation of each development stage of the EKD4SFCA method is followed by case studies that show applications with actual data and how results are interpreted.

Finally, based on the scientific literature on validity and on the results of two validation exercises with different data sets, the method is validated. We conclude with remarks on the implications of the proposed method for firms' market entry strategies, public policies and the disadvantaged populations, and the identification of a number of paths for future research.

RESUMO

Esta tese propõe um novo método para a identificação de mercados geográficos alvo em alta resolução, onde o acesso de pacientes aos cuidados de saúde é baixo e onde não há problemas concorrenciais potenciais para empresas e administrações públicas promoverem oferta adicional. O método proposto ultrapassa limitações de métodos anteriores, aperfeiçoando as avaliações do acesso e da concorrência. Além disso, integra de forma inovadora os dois tipos de avaliação. Esta integração metodológica possibilita o processamento de um conjunto abrangente de dados para, por um lado, identificar áreas geográficas onde as empresas terão uma probabilidade relativamente maior de sobrevivência e expansão após a entrada no mercado. Por outro lado, este método pode ser adotado como suporte a políticas públicas que visam a redução de desigualdades no acesso e o aumento da oferta de cuidados de saúde para as populações desfavorecidas.

O primeiro estágio do desenvolvimento do método consiste na construção de índices através da análise de componentes principais, para a captura de variações no acesso em diferentes áreas geográficas. Envolve igualmente uma modelação aperfeiçoada dos padrões de viagem potencial dos pacientes até aos pontos de oferta de cuidados de saúde. Por sua vez, estas melhorias são utilizadas como base para o método proposto, concretamente o *Extended Kernel Density 4-Step Floating Catchment Area* (EKD4SFCA), que incorpora a avaliação concorrencial na análise. A apresentação de cada estágio de desenvolvimento do método EKD4SFCA é acompanhada por casos de estudo, que apresentam aplicações com dados concretos e mostram como os resultados são interpretados.

Por último, com base na literatura sobre validação e nos resultados de dois exercícios de validação com diferentes conjuntos de dados, o método é validado. Concluímos com observações sobre as implicações do método para estratégias empresariais, para políticas públicas e para as populações desfavorecidas, identificando oportunidades para investigação posterior.

ACKNOWLEDGEMENTS

I would like to thank my supervisors Professors José Luís Borges and António Coelho for their kindness, patience, rigor and motivation, the Portuguese Foundation for Science and Technology (Fundação para a Ciência e a Tecnologia) and the Portuguese Health Regulation Agency (Entidade Reguladora da Saúde), for their financial support, for the DEGI Department (Departamento de Engenharia e Gestão Industrial), for having accepted me to do this fulfilling work, and, foremost, all members of my family, in particular my mother, Norma Polzin, my father, Charles Frédéric Polzin, my wife, Alexandra Polzin, and my children, João Roberto Moreira Polzin and Pedro Filipe Moreira Polzin, for their love, patience and support.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

LIST OF ABBREVIATIONS

2SFCA	Two-Step Floating Catchment Area
ACCC	Australian Competition & Consumer Commission
CC	Competition Commission
CGAP	Consultative Group to Assist the Poor
DEGI	Departamento de Engenharia e Gestão Industrial
E2SFCA	Enhanced Two-Step Floating Catchment Area
EC	European Commission
EKD2SFCA	Extended Kernel Density Two-Step Floating Catchment Area
EKD4SFCA	Extended Kernel Density Four-Step Floating Catchment Area
ERS	Entidade Reguladora da Saúde
ESS	Espírito Santo Saúde
EU	European Union
GIS	Geographic Information System
HHI	Herfindahl-Hirschman Index
HPSA	Health Professional Shortage Areas
IOM	Institute of Medicine
JMS	José de Mello Saúde
KD2SFCA	Kernel Density Two-Step Floating Catchment Area
KMO	Kaiser-Meyer-Olkin
MAUP	Modifiable Areal Unit Problem
MSA	Measure of Sampling Adequacy
NUTS	Nomenclature of Territorial Units for Statistics
OECD	Organisation for Economic Co-operation and Development
OFT	Office of Fair Trading
PCA	Principal Component Analysis
RNCCI	Rede Nacional de Cuidados Continuados Integrados
SCP	Structure/Conduct/Performance
SPSS	Statistical Package for the Social Sciences
UKDH	United Kingdom Department of Health
UN	United Nations
US	United States
VBA	Visual Basic for Application
WHO	World Health Organization

1. Introduction

This thesis proposes a new method to identify fine-resolution target geographic markets with low patients' access to health care and no potential competition problems for firms and public administrations to promote new supply. The proposed method corrects limitations of previous methods, enhancing the assessments of access and of competition. Moreover, it innovatively integrates the two kinds of assessment.

The importance of the assessment of access to supply points is evidenced by the increasing number of publications found in the scientific literature and by its multidisciplinary character. In recent years, research on this subject has been proliferous and is being conducted in many different fields of research. These research fields are linked to several disciplines, such as, for example, public health, economics, sociology, geography, urban design, urban or regional planning, and environment and transportation engineering. The growing literature on this topic is reinforced by the increased availability of road travel data and geo-referenced data on the supply and demand of products and services, and the computational advances of the geographic information systems (GIS).

The usefulness of the assessment of access to supply points relies on the fact that this kind of evaluation produces valuable results for the assistance (i) on firms' strategic market entry decisions and (ii) on public policies to improve the geographic distribution of supply resources for disadvantaged populations.

Competition assessment is important for firms to know their competitors and the markets, in which they compete, or in which they plan to enter. In strategic and business plans, competition assessment is carried out to identify and choose the most attractive geographic areas for a successful and sustainable market entry.

As far as public administrations are concerned, competition assessment is relevant and useful to identify geographic markets and regional differences in the competition between firms. The identification of regional differences in competition is important, for example, when public administrations wish to promote public bids for contracting firms to provide products and services to the populations. This way, they can target competitive markets to explore competition and establish lower, more efficient prices for the contracted products and services.

Still, competition assessment is important for public regulatory agencies and competition authorities, who need to evaluate mergers and acquisitions, in order to estimate market structure changes and decide whether or not to approve these concentration operations.

The usefulness of combining the assessments of access and of competition can also be established for firms and public policies.

Regarding the utility of this integration for firms, the results of a combined assessment can help in aligning the firms' supply decisions with their strategic and market positioning objectives. The integrated assessment can enable the identification of target markets, where there is relatively more unmet potential demand, and where no competition hindrance by incumbent firms is expected.

The integrated assessment of access and competition is also relevant and useful for public policies destined to improve the distribution of supply resources, reduce access inequalities and improve the populations' health status. This is due to the fact that competition problems, consisting of lack of competition and abuse of a dominant position, can affect access negatively, for example, via discrimination of patients, excessively high prices, diminished supply and/or reduced quality of products and services (CC, 2013; EC, 2004; Motta, 2004; OECD, 2011, 2012a; OFT, 2007). Regarding the relation between competition and health care quality, competition has been proved to improve quality in health care provision, and quality of care is essential for an efficacious access to health care, to improve the patients' health status and to satisfy their health needs (Gaynor et al., 2012a; Ikkersheim and Koolman, 2012).

However, to the best of our knowledge, no model has been created so far that enables an integrated assessment of access and competition. This integration is proposed for the first time in this thesis, such that it constitutes its main innovation and scientific challenge.

1.1. Motivation

The integration of the assessments of access and of competition constitutes the main innovation of the thesis, but also the topmost motivation for this work.

In recent years there has been increased attention in the assessment of access for the detection of geographic areas with low access, where disadvantaged populations reside, and for the identification of opportunities for firms and public administrations to better distribute supply resources.

The proliferation of new access studies and research advances on this theme in the scientific literature is supported by the increasing availability of geo-referenced data sets and travel times data, and by the computational advances in GIS, as is recognized by some authors. "With increasing availability of geo-referenced individual-level data and improvement in the representational and geo-computational capabilities of Geographical Information Systems (GIS), it is now more feasible than ever before to operationalize time-geographic constructs" (Kwan, 2004: 277). "The increasing availability of geo-referenced data and analytic tools like GIS makes possible the improved reach and efficiency of community-based interventions" (Alcaraz et al., 2009: 56). In particular, regarding the health care industry, "the increasing abundance of digital data (for example, population data, street and road network, physician database) and advancement of GIS technology now make it possible to identify distributions of physicians and population at finer spatial resolutions" (Luo and Wang, 2003: 866).

The relevance of research on access to supply points contributes to the multiplication of studies on this theme. The importance of these studies is well established, since "an important objective of firms is improving the accessibility of a product to consumers, and improving the accessibility to distribution centers", and "an important public policy goal is to improve the accessibility of public services to constituents" (Burkey, 2012: 783). In industries in which consumers travel to supply points to obtain the needed products, improved access stimulated by a better geographic distribution of supply resources can be crucial for a firm to obtain a higher market share, or for public administrations to effectively address the populations' needs.

The health care industry is the industry that attracts most of the attention of researchers involved in access studies. There have been a number of studies concerned with developing measures of access to health care (Higgs, 2004). Research on access to health care has been focusing numerous kinds of health care, like primary care (Dewulf et al., 2013), cancer prevention exams (Lian et al., 2012), mental health care (Ngui and Vanasse, 2012), hospital care (Xu and Cui, 2012), dialysis (renal care) (D.-H. Yang et al., 2006), pediatric care (Cervigni et al., 2008) and maternity services (Song et al., 2013), to name a few.

However, access studies have been conducted also in other contexts, like, for example, public transportation and access to bus stops (Langford et al., 2012), quality of life and access to urban parks (Lee and Hong, 2013), education and access to public libraries (Higgs et al., 2013), food and access to grocery stores (Shaw, 2012), and veterinary services (Polo et al., 2013).

Still other kinds of access research, closely related to typical assessments of consumers' access to supply points is being carried out, including spatial crime analysis for investigation and prevention (Fahui Wang, 2012), studies for the identification and interpretation of spatial patterns of fire incidents (Corcoran et al., 2013), and studies for spatial tourism planning and policy (van der Merwe and van Niekerk, 2013).

These studies are being conducted by researchers from a variety of university departments, which reflects their multidisciplinary character. Access studies have been published in scientific journals by authors from departments of geography, medicine, public health, environmental sciences, epidemiology, veterinary medicine, physics, civil engineering, urban engineering, economics, social sciences, biostatistics, business, and management and industrial engineering, among others.

Nevertheless, even though there is a widespread attention on the access research theme, still some obstacles need to be overcome to better study access problems caused by the geographic asymmetries that arise from an inadequate supply distribution.

One obstacle is the scarcity of some data. While there are nowadays more geo-referenced data sets and travel times data, spatially disaggregated, updated and complete consumption or utilization data are still almost always unavailable. Utilization data for the local population is usually not available, even in highly developed countries like the United States (Lin et al., 2005). The lack of specific health care utilization data, sometimes due to confidentiality issues, hinders the application of analytical methods built to use these data.

Other reported problems include the complexity of some methods and the needed skills to apply a GIS, alongside with the high costs of some mainstream software applications, the software modules and the needed proprietary data. Unequal access to GIS and spatial data has the effect of excluding or disempowering stakeholders in the planning process (Olafsson and Skov-Petersen, 2013). The increasing reliance of the application of space-time accessibility assessment on the availability of specific expertise and GIS software has the potential to obstruct the dissemination of access assessment beyond the geography, GIS and transportation communities (Neutens et al., 2010).

Thus, considering these problems, we identified an opportunity to create an integrated access and competition assessment method based on a popular method that produces simple supply-to-demand ratios with restrained, but wide-ranging, publicly and easily available information. Furthermore, we decided to construct the method in a parsimonious way, avoiding unwieldy and impractical calculations that could make the method excessively complex and hinder its use and future dissemination. In addition, we defined that it should be possible to apply the method with a relatively inexpensive and easy-to-use GIS, as the one that we used in the applications presented in this thesis (Microsoft MapPoint 2009). These were the directions for our innovative efforts, to overcome the identified research obstacles and fulfill our innovation motivation.

Another motivation for this thesis was to construct the new method for the Portuguese Health Regulatory Agency (Entidade Reguladora da Saúde, ERS), which is the public institution responsible for the regulation of the health care industry in Portugal. ERS has as a duty guaranteeing equity in access to health care among patients and defending competition between health care providers. The motivation was, more specifically, to improve the identification of priority regions as targets for regulatory actions, to correct access and competition problems.

Accordingly, the innovations should enable a fine-resolution analysis. In particular, the innovations to the competition assessment methods, that we propose to integrate into our method, should be useful for ERS to identify differences in competition across and inside markets, with an improved geographic resolution. The competition assessment methods that are regularly used by regulatory agencies and competition authorities do not enable the identification of variations in the competitive pressure due to the patients' flows across the borders of the defined geographic markets.

These motivations are already being fulfilled. This is evidenced by the work that is being carried out at present by ERS, and by the fact that the works presented in this thesis have already been contributing to the scientific literature and to discussions in different scientific communities, with published papers and presentations in conferences.

ERS is already adopting some of the developments that are presented herein. For example, ERS studied the access of patients to long-term care using an innovation proposed in this thesis (ERS, 2013). Also, opinions on competition among health care providers, issued by ERS to the government and to the

Portuguese competition authority, as required by the Decree-Law N. 139/2013, of 9 October, and Law N. 19/2012, of 8 May, respectively, make use of the competition assessment steps of the method (ERS, 2014a, 2014b, 2014c).

Regarding the published papers, these are Polzin et al. (2013), Polzin et al. (2014a), Polzin et al. (2014b), and Polzin et al. (2015).

As far as the presentations are concerned, we presented developments of the method in the European Conference on Health Economics 2012, in Zürich, Switzerland, in the 4th Symposium on Industrial Engineering and Management 2013, in the city of Porto, and in the Advances in Business-Related Scientific Research Conferences 2013 and 2014, in Rome, Italy.

1.2. Framework for the assessments

Access and competition assessments share similarities, because both depend on the definition of the specific products or services that will be studied and on the relevant geographic areas that will be considered. Besides, the two kinds of assessment consider supply and demand and, in industries in which consumers need to travel to supply points to obtain the products or services, both have to consider the attraction of consumers to supply points.

Even though access and competition assessments share similarities, the separation of both is common in practice. One possible explanation for this separation is the fact that most institutions that conduct these assessments have different, specific objectives. Access assessment is carried out, for example, by public organizations that need to identify shortage areas, in order to direct resources to reduce the most severe inequalities in access to products or services. Competition assessment is conducted, for instance, by firms that need to monitor markets and their competitors, and to better define their marketing strategies, or for the elaboration of strategic or business plans that support market entry decisions.

Another explanation for the separation of access and competition assessments is the fact that there is no method that integrates access and competition assessments.

To integrate both kinds of assessment the method proposed in this thesis is the extended kernel density 4-step floating catchment area (EKD4SFCA) method. The EKD4SFCA method is based on an influential method that was created in 2003 and applied to the health care industry context, namely the 2-step floating catchment area (2SFCA) method of Luo and Wang (2003), which was in turn built on the spatial decomposition method created by Radke and Mu (2000).

The origins of the 2SFCA method trace back to 1997, more specifically to the work by Peng (1997) that introduces the notion of floating catchment areas to avoid the problem of measuring jobs-housing ratios on pre-defined and arbitrary jurisdictional or administrative boundaries.

Catchment areas are market or customer geographic areas that reflect the consumers' preferences and their buying patterns. The boundaries of the catchment areas are defined by the maximum distances or travel times that the consumers are willing to travel to reach the products and services that they need or desire. Floating catchment areas are, in turn, catchment areas that float, in the sense that a series of them are computed, for all the geographic units of analysis of a chosen study region.

There have been many developments on the 2SFCA method since 2003, and many new 2SFCA-based methods have been created, which have been applied also to other industries besides the health care industry.

As McGrail (2012) explains, the developments and improvements to the spatial access assessment with the 2SFCA method are mainly of two kinds, namely the addition of a distance decay function and the consideration of variable catchment areas sizes. A third kind of improvement is the consideration of non-spatial (or aspatial) variables, to enable a comprehensive access assessment, instead of only a spatial accessibility assessment.

As our own enhanced 2SFCA method, we present in this thesis the extended kernel density 2-step floating catchment area (EKD2SFCA) method (Polzin et al., 2014a). This method changes the distance decay function of the kernel density 2-step floating catchment area (KD2SFCA) method of Dai and Wang (2011) to properly model the potential travel patterns of patients to health care supply points. It also embeds two indices, in order to adjust the final access scores of the EKD2SFCA method for the populations' health care needs and the mobility of the populations when accessing health care.

The EKD2SFCA method is the basis for our EKD4SFCA method (Polzin et al., 2013, 2014b, 2015). As will be seen, the proposed EKD4SFCA method presents as its main innovation the incorporation into it of two new competition assessment methods. These proposed competition assessment methods were constructed based on two existing methods, respectively: the Herfindahl-Hirschman Index (HHI) and the method created by Melnik et al. (2008) that identifies market dominance.

The HHI is a well-known and widely adopted index used to assess market concentration and the relative positions of the firms in a market. Its calculation is based on the assessed market shares of the firms that are active in the market that is being studied, and it reflects the entire market share distribution of these firms.

The method of Melnik et al. (2008) is also based on the market shares of the firms. It calculates a dominance threshold to be compared to the market share of the main firm, that is, the firm with the highest market share: if the market share of this firm is higher than the threshold, than the firm dominates the market.

1.3. Research objectives

The main objective of this thesis is to propose a new method to assess access to health care and competition between health care providers in an integrated way. As noted before, this is the first time that access and competition assessments are integrated, which constitutes also the main innovation of the thesis and motivation for this work.

Supplementary objectives of this thesis are to introduce improvements in the access and competition assessment methods that are integrated with our four-step analytical approach, and to contribute to the literature on access to health care and competition among health care providers.

Given these objectives, we expect the following results in this thesis:

1. Theoretical background: identification of the main methods in the context of our research focus, indicating opportunities for enhancements and the need for a new, integrated method;
2. Access assessment: proposal of a new method to assess access to health care supply points by producing access scores, using a simple GIS, easily obtainable information on supply, and other data collected systematically by national statistics institutes;
3. Competition assessment: proposal of innovative competition assessment methods; and
4. Integration of the assessments of access and of competition: proposal of the new method, which will summarize access and competition assessment results, producing quantitative outcomes associated with geographic units.

1.4. Summary of the thesis

The thesis is structured as follows.

Chapter 2 is destined to present the theoretical background and the methods that are used to assess access and competition. The focus is on the health care industry, because the thesis concentrates its methodological developments and applications on that industry. This chapter reviews the literature, presenting state-of-the-art methods, and contributes to the literature on access to health care and competition among health care providers, consolidating relevant concepts. It identifies opportunities for methodological enhancements. We conclude chapter 2 indicating the need for the creation of our proposed method, considering the fact that there is no other method in the literature that integrates access and competition assessment, and the relevance and usefulness of this integration.

Chapter 3 is dedicated to the presentation of our proposed EKD4SFCA method, more specifically, its development and its constituent parts. We first present a new method, namely the EKD2SFCA method, which constitutes the first two steps of the proposed EKD4SFCA method. It was constructed based on a state-of-the-art method to assess access to health care. Thereafter, we present and explain steps 3 and 4 of our proposed method, namely the extended HHI and the extended dominance identification method. We also present two case studies for the access and the competition assessments, respectively, showing how our innovations produce improved results and how they are interpreted.

Chapter 4 is dedicated to the first full application of the EKD4SFCA method in this thesis. We demonstrate how the method is applied, and the logic of its application. Using data from the Portuguese hospital health care sector, we apply the method, with the objective of identifying target hospital health care markets for entry.

Chapter 5 is dedicated to the validation of the method – its applications and the results that it produces. We determine that our method attains content validity and present two case studies with validation exercises, in which we introduce changes to the data and assess how the method reacts to these changes. This way, we are able to confirm that the method captures changes in the data as expected, considering how it is constructed, the information that it uses and its objectives.

Finally, chapter 6 presents the concluding remarks, taking into account the research objectives of the thesis and the expected results. We also present paths for future research, indicating suggestions for further applications of the method and changes to the method that could be carried out and explored, in order to verify if further improvements could be accomplished.

2. Theoretical background and methods for the assessment of access and competition

As a theoretical underpinning to contextualize the research carried out for the creation of the EKD4SFCA method, this chapter begins with definitions of relevant concepts regarding spatial analysis.

Spatial analysis is executed with a collection of techniques and models that use spatial referencing of data (Goodchild and Haining, 2004). It focuses on the relative geographic location and distance between suppliers and consumers, and on characteristics of these two groups, among other factors (Gesler and Meade, 1988).

The theoretical pillar that supports spatial analysis and spatial models is Tobler's first law of geography, which states that "everything is related to everything else, but near things are more related than distant things" (Tobler, 1970: 236).

Jarvis' law is also an important contribution and it laid the foundation for contemporary spatial analysis. It postulates that admission rates to hospitals vary inversely with distance between the patients' residences and the hospitals (Sloas et al., 2011; Sohler and Clapis, 1972): "the advantages of any public lunatic hospital, however freely and equally they may be offered to all people of any State, are yet, to a certain degree, local in their operation, and are enjoyed by people and communities to an extent in proportion to their nearness to or distance from it" (Jarvis, 1866: 404).

Spatial regression, spatial autocorrelation and spatial interaction methods are examples of methods that are commonly adopted to carry out spatial analysis. Spatial interaction methods have been typically developed and applied to assess the potential access of consumers to supply points in industries, in which people have to travel to the supply points, in order to obtain the desired products or services. In such industries, these methods enable the identification of underserved geographic areas, where disadvantaged populations reside, i.e., populations that face access barriers. In these areas, the inhabitants have comparatively greater difficulties in accessing products or services, especially because they reside too far from the supply points. An example of an industry that is the focus of a multitude of developments and applications of spatial interaction methods is the health care industry. Acknowledging the relevance of spatial interaction methods in assessing access and the focus of their developments and applications on the health care industry, in this thesis we propose to develop a new spatial interaction method and we also focus on the health care industry.

Since this thesis focuses its methodological developments and applications on the health care industry and on access to health care, section 2.1 is destined to define health care and access to health care. It presents the theoretical spatial analysis approach that is adopted along this thesis, it explains access to supply points, and it presents the most relevant spatial interaction methods to assess access.

Thereafter, considering that competition between firms is another relevant theme explored in this thesis, as noted in the introduction, section 2.2 explains competition and its relevance. It defines the concept of relevant market, it discusses the assessment of competition, and it presents the main existing competition assessment methods.

2.1. Access to health care

The Directive 2011/24/EU of the European Parliament and of the Council of 9 March 2011 on the application of patients' rights in cross-border healthcare, defines health care as "health services provided by health professionals to patients to assess, maintain or restore their state of health, including the prescription, dispensation and provision of medicinal products and medical devices" (EU, 2011: 55).

Access to health care is a human right, as established by The Universal Declaration of Human Rights of the United Nations. Accordingly, everyone should have a standard of living adequate for the health of himself and of his family, including health care (UN, 1948). As noted by McGrail (2012), "access to health care is widely accepted internationally as a key goal in meeting the health needs of individuals" (McGrail, 2012: 1). Its importance resides on the fact that it is at the center of health systems analysis (Gulzar, 1999). However, McGrail (2012) notes that assessing the extent to which adequate access to health care is achieved is difficult. There is still no consensus on how access to health care should be defined, even though it has been studied for more than 50 years, during which hundreds of publications were produced. Norris and Aiken (2006) alerted that access to health care has been used as a "catch all" term, while the meaning can vary substantially across studies. The authors affirmed that "it is as if everyone is writing about 'it' but no one is saying what 'it' is" (Norris and Aiken, 2006: 60).

2.1.1. The concept of access to health care and its relation with need and utilization

The concept of access to health care is frequently explained together with two other related concepts, namely need and utilization of health care. According to McLafferty (2003), issues of access, need and utilization come together in the evaluation and planning of health care.

A model that helps to explain why access, need and utilization are interrelated is the behavioral model of health care utilization (Andersen, 1995). According to this model, need for care and factors that facilitate or hinder utilization are the central elements on which utilization of health care depends.

The concept of need is a core aspect to be considered when determining whether access to health services is equitable. As defined in Andersen et al. (1983), equity of access occurs when services are distributed according to people's need for them. Equitable access to health care can only be achieved if there is equal access for equal need, which is the definition of horizontal equity (Allin et al., 2007; Oliver and Mossialos, 2004; Rice and Smith, 2001) or formal equality (Hay, 1995).

Need influences people's utilization of health care and it is composed by two elements: perceived need and evaluated need (Aday and Andersen, 1974; Andersen, 1995). These elements are equivalent to two other terms, namely felt need and normative need, respectively (Joseph and Phillips, 1984; Oliver and Mossialos, 2004).

Felt or perceived need is the need that is perceived by the individual. It can be assessed with data collected in surveys on perceived health status, symptoms of illness and disability. Estimates of it can be obtained with variables such as education, occupation and ethnicity, because these variables represent social structure, which explains the social phenomenon of need (Andersen, 1995). Other proxies for perceived need are demographic variables, such as age and gender, which may imply the likelihood that people will need health services (Andersen, 1995). There is evidence for a U-shaped relationship between demand and age, such that younger and older people typically demand more health services, because they have higher needs. Women also need and demand more health services. The higher the health needs of a population, the greater is its demand for health relative to its size, and, *ceteris paribus*, the greater the demand, the worse is access, because supply will become insufficient (McGrail and Humphreys, 2009a).

Normative or evaluated need is the need assessed by health professionals. Indicators of it can be computed with data such as, for example, diagnosis classifications, physician-rated urgency of surgeries, or mortality and morbidity data (Graves, 2009; Oliver and Mossialos, 2004). According to the evaluated need definition, an individual with good health does not need a health service, even though he/she may feel that he/she needs it. Only a medical consultation can determine if there is need for health care. However, people do not always go to consultations following previous medical recommendations. Frequently, it is just the perceived need of an individual that induces him/her to go to a medical consultation, or to use other kinds of health care, such as exams. In these cases, only thereafter it will be possible for a health professional to establish whether the individual has good health, or whether more health care is needed.

Health care utilization is the act of using or consuming health care (according to Donabedian (1972), utilization is the proof of access). It is complex to establish the relationship between health care need (perceived or evaluated) and utilization, and the complexity stems from the fact that there is no direct correspondence between these variables (McLafferty, 2003). Besides need, many other variables can interfere and affect utilization. Nevertheless, it can be stated simply that, in general, need leads to utilization, or utilization partially reflects need.

The relationship between access, need and utilization can be better understood with the acknowledgment of the fact that underlying any effort to improve access is the belief that better access – that is, increased or more effective access – leads to an increase in appropriate utilization – namely utilization by people in need of health care (Thomas and Penchansky, 1984). In turn, the increase in utilization improves user satisfaction and health status, and promotes better health outcomes, that is, verifiable health improvements after the utilization of health care (Andersen, 1995; Gold, 1998; Gulzar, 1999).

Regarding the relationship between access and utilization, in particular, we note that some researchers use both terms as synonyms, and employ utilization as a proxy for access, such as Whitehead et al. (1997), for example. However, it is mostly accepted that access refers only to opportunities, while utilization is the materialization of these opportunities (Allin et al., 2007).

Access to health care has several alternative definitions, and its meaning can vary substantially across studies, as noted before, but some widely accepted definitions can be identified in the literature. Aday and Andersen (1974) and Andersen et al. (1983) indicate that access to health care is a set of dimensions that describe the potential and actual point of entry of individuals to the health care system. Penchansky and Thomas (1981) define access as a concept representing the degree of fit between patients and the health care system. Andrulis (1998) and Ricketts and Goldsmith (2005) mention a popular definition of access from the Institute of Medicine (IOM), which says that access is the timely utilization of health services to achieve the best possible outcome.

Since the method to assess access that is integrated into our proposed method aims to identify disadvantaged populations, i.e., populations that face access barriers, this thesis adopts a definition that mentions barriers explicitly. Thus, it considers access as the ability of a population to use health services, while access barriers (or facilitators) to the utilization affect this ability (Gulzar, 1999).

2.1.2. Potential access

Having defined access and its relationship with need and utilization, this subsection clarifies related access concepts and definitions found in the literature that are used in this thesis, including a number of access dimensions. Most importantly, as will be seen, it identifies two different approaches to assess access, namely the realized access approach and the potential access approach, and it establishes our focus on the potential access approach.

Access and accessibility are terms that are sometimes employed interchangeably. According to WHO (1978), accessibility is the continuing and organized supply of care that is geographically, financially, culturally and functionally within easy reach of people. However, the term accessibility is typically adopted in the literature as a reference to the ease of reach between two locations: a measurable, geographical dimension of access, which is also designated by the terms proximity and geographical accessibility (L. Wang, 2007).

Still, there are two other commonly used concepts that refer to accessibility, namely revealed accessibility and potential accessibility. These terms coin approaches that are typically adopted by researchers to assess patients' access to health care. One is the revealed accessibility approach, or realized access approach, which focuses on utilization patterns. The other is the potential accessibility approach, or potential access approach, which takes into account potential barriers to utilization and measures access as potential utilization (Cooper et al., 2009b; Joseph and Phillips, 1984).

The realized access approach requires the collection of utilization data, which can be useful, for example, to identify factors that determine health care utilization and to characterize the actual demand. However, such data are almost always unavailable (Lin et al., 2005).

When utilization data are not available, utilization can be measured by surveys that are applied to patients, who inform their utilization of health services. However, survey data are sometimes proprietary and costly, especially when up-to-date data for a wide range of health care providers have to be collected. Another problem is that these data may reflect survey bias and are usually not comparable to other surveys' data (Allin et al., 2007).

If utilization data are available, a likely problem is that they can conceal access barriers if factors that yield differences in utilization across regions are not accounted for. As demonstrated by Harrington et al. (2012) and Lewis and Longley (2012), a multitude of factors can produce differences in utilization among populations, and some factors are due to the complexity of patients' behavior.

As opposed to these two studies, and still other studies, like, for example, the ones of Clarke et al. (2002), which quantifies patterns of realized access using household expenditure data, and Birkin et al. (2005), which models variations in health care utilization, the work in this thesis focuses on potential access analysis.

Potential access refers to access barriers to the utilization and it is frequently defined as the presence of enabling resources, which are resources that are required for utilization to take place (Andersen, 1995; Graves, 2009; Gulzar, 1999; Higgs, 2004). The presence of enabling resources does not determine utilization or the actual demand. Thus, the main limitation of the assessment of potential access is the fact that it does not provide results that necessarily reflect the actual patients' utilization of health care. However, the assessment of access that considers enabling resources can identify the populations that face potential access barriers, i.e., the disadvantaged populations. These populations are the ones with the lowest potential access to health care, regardless of their utilization patterns. In fact, a disadvantaged population can be, for instance, a population with a relatively high utilization, but at an amount that is not sufficient to satisfy its needs.

Enabling resources include the availability of suppliers, and people's means to reach products or services and to make use of them. Some examples of enabling resources' metrics that can be used to assess potential access are availability of products or services, sources of supply, travel distances, and travel and waiting times (Andersen, 1995).

A set of access dimensions that comprises a broad range of enabling resources' variables are the five dimensions of Penchanksy and Thomas (1981), which represent closely related phenomena: accessibility, availability, accommodation, affordability and acceptability.

Accessibility and availability are the two spatial dimensions of access that together measure the consumers' spatial accessibility (Guagliardo, 2004; Luo, 2004). These two dimensions form the basis to address the other access dimensions. The accessibility dimension is, as mentioned before, also designated by the term proximity, or geographic accessibility, and it refers to the ease of reach between two locations measured in terms of distance or travel time (McLaughlin and Wyszewianski, 2002; L. Wang, 2007). The availability dimension refers to the available supply capacity and the supply points from which consumers can choose.

Accommodation, affordability and acceptability are non-spatial dimensions (Guagliardo, 2004). They are typically measured by demographic and socioeconomic variables that describe populations' characteristics.

Accommodation reflects the organization of the supplier's operation to meet the constraints and preferences of the consumers. The affordability dimension is determined by the relationship between the suppliers' fees for the products or services and the consumers' ability and willingness to pay for these services. Acceptability captures the extent to which the consumer is comfortable with the suppliers' characteristics, and vice versa, and examples of these characteristics are age, gender, social class, and ethnicity (McLaughlin and Wyszewianski, 2002; Thomas and Penchansky, 1984).

The five access dimensions of Penchanksy and Thomas (1981) are well-known and widely accepted in the literature, but they are not consensual, because there is a multitude of spatial and non-spatial dimensions of access (McGrail, 2012; Penchansky and Thomas, 1981; Fahui Wang and Luo, 2005). For example, Norris and Aiken (2006) proposed four access attributes, which they understood to be critical to the understanding of access: availability, eligibility, amenability and compatibility. According to Norris and Aiken (2006), availability comprehends geographic proximity and personal convenience, which considers factors such as office hours and waiting times. Eligibility is similar to the affordability dimension of Penchansky and Thomas (1981), because factors of eligibility include income and insurance coverage. Amenability is related to the willingness of the consumer to consume. Finally, considering the compatibility dimension, if the consumer and the supplier are incompatible, utilization will be rare. This dimension depends on cultural aspects.

Even though there are no consensual dimensions, attributes or subgroups of access, it is reasonable to classify variables and corresponding data collected for an access study in two broad groups: spatial and non-spatial access dimensions. Pragmatically, it can be said that spatial dimensions relate to proximity and availability, and non-spatial dimensions include supplementary factors that play a role in the consumers' potential access.

Finally, before passing to the methodological subsection regarding access, it is important to note the parsimony advantage of studying potential access to health care instead of realized access, even though potential access studies ignore the actual patients' utilization of health care. Contrary to the assessment of realized access, the assessment of potential access makes use of more easily obtainable data from official sources, which are sufficient to identify the populations that face potential access barriers. In fact, the potential access approach is the ideal approach to assess access to health care when utilization data are not available, and the scarcity of data is a real obstacle for the assessment of access to health care. Spatially disaggregated, updated and complete utilization data are almost always unavailable. In particular, utilization data for the local population is typically not available, even in highly developed countries like the United States (Lin et al., 2005).

Besides, another advantage of the assessment of potential access assessment is that it produces results that are valued by public policymakers to deal with problems of access (Lin et al., 2005). In particular, it is considered useful for policies aimed at reducing inequities in access to health care, i.e., to promote equitable access (Andersen, 1995; Andersen and Newman, 1973). The assessment of potential access enables the identification of the locations of the truly disadvantaged populations, such that the results that it produces can be regarded as the essential first step toward any meaningful and effective government intervention program (Gatrell and Wood, 2012; Luo, 2004).

The choice to focus on the assessment of potential access in this thesis stems from the consideration of these advantages.

2.1.3. Spatial interaction methods to assess potential access

This subsection serves the purpose to present the main spatial interaction methods to assess potential access in the context of our research.

Most spatial interaction studies in the scientific literature focus on measuring spatial accessibility, typically via the construction of indices and the computation of spatial accessibility scores. A groundbreaking study to assess spatial accessibility of patients to health care is the one of Luo and Wang (2003). This paper studied the variation of spatial accessibility to primary care in the Chicago region, in the US, and proposed a new spatial interaction method to assess spatial accessibility, namely the 2-step floating catchment area (2SFCA) method. Besides, it suggested that this new method might be used to improve the definition of the Health Professional Shortage Areas (HPSA) by the US Department of Health and Human Services.

Next, we focus on the influential 2SFCA method and its main predecessors – supply-to-demand ratios, gravity-based methods and kernel estimation methods –, and also on the enhanced, newer versions of the 2SFCA method. The 2SFCA method and 2SFCA-based methods are, at present, the most popular spatial interaction methods adopted for the assessment of potential access.

2.1.3.1. Supply-to-demand ratios

Typical spatial accessibility measures, which consider only geographical accessibility and availability metrics, are supply-to-demand ratios, also known as supply ratios:

$$R_k = \frac{S_k}{P_k}; \tag{1}$$

where R_k is the supply ratio at area k, S_k refers to supply resources, like the number of professionals working at the supply points located at area k, for example, and P_k is the population that resides at area k, as a proxy for demand.

These ratios are area-based metrics that are computed within bordered areas, such as administrative regions, using generally readily available data, like, number of physicians, clinics, or hospital beds for the numerator, and population size for the denominator (Guagliardo, 2004). In the health care industry context, they are also known as physician-to-population ratios or, more generally, provider-to-population ratios (or population-to-provider ratios, when the demand proxy is set as the numerator and the supply variable is considered in the denominator).

However, there are four limitations of the supply ratios, which constitute problems for the assessment of access:

(i) They do not consider border crossing of consumers. Access is assessed in each area considering only the population that resides in that area. However, many people have to cross borders when traveling to a supply point, in order to obtain the desired products or services.

(ii) They do not measure variability in access within the areas of aggregation. Supply ratios use large areas as geographic assessment units. However, by using large areas, it is not possible to detect access level variations inside these areas.

(iii) They do not incorporate any measurement of distance decay or impedance (Fortney et al., 2000; Guagliardo, 2004; McGrail and Humphreys, 2009a). The large areas used by the supply ratios treat large distances inside these areas the same way as small distances. For example, they do not differentiate a travel time of one hour from a travel time of five minutes. This is a limitation, since higher access levels shall be associated to smaller distances, in accordance with Tobler's first law of geography and Jarvis' law.

(iv) They are more vulnerable to the modifiable areal unit problem (MAUP), which is the lack of robustness that occurs because the choice of the boundaries of the geographic areas use to assess access influences the results (Chaix et al., 2005; Fotheringham and Wong, 1991; Guagliardo, 2004; Guagliardo and Ronzio, 2005; Higgs, 2009; Kwan and Weber, 2003; Openshaw, 1984). In fact, choosing different large areas to assess access with a supply ratio can produce considerable variations in the results for a specific population that resides inside all those alternative chosen areas (a population of a postcode area that is contained in those areas, for example).

2.1.3.2. Gravity-based methods

A first spatial interaction model that deals with these problems, which is commonly cited in most works on analyses of patients' access to health care, is the gravity-based method. As L. Wang (2011) notes, gravity-based methods are the most widely used, among the numerous types of accessibility methods. Gravity-based methods have been used in geographical and urban studies since the 1940s, and they were initially developed to predict retail travel, or to help with land use planning (Geurs and van Wee, 2004; Guagliardo, 2004). The gravity model enables the calculation of the potential interaction between populations and supply points.

As formulated by Hansen (1959), access at a population location 1 to a type of supply activity at an area 2 is directly proportional to the size of that activity and inversely proportional to a function of the distance between location 1 and area 2. The popular gravity method of Hansen (1959) can be described by the following formula:

$$A_i^{\mathrm{H}} = \sum_{j=1}^{n} S_j d_{ij}^{-\beta} ; \qquad (2)$$

where A_i^{H} is the gravity-based index, S_j is the number of professionals or other supply resources of the supply point S located at j, d_{ij} is the road travel time by car or distance between the population residing at i and j, β is the distance decay or travel-friction coefficient, which describes the travel-time effect between areas and points, n is the total number of supply points available for the population residing at i, and $d_{ij}^{-\beta}$ is the distance decay function: an inverse power function.

The A_i^{H} index produced by this gravity method is the result of the sum of the spatial accessibilities to each of the supply points (Hansen, 1959; Tanser, 2006). The nearer the supply points are to a specific population area, and the more supply resources are available, the higher is the spatial accessibility associated to that area.

Some drawbacks of the gravity model can be identified. One of the drawbacks is that there is no consensus on the choice of the distance decay function and of the distance decay coefficient. Many other distance decay functions are used besides the inverse power function, namely linear, Gaussian, logistic or exponential functions (Geurs and van Wee, 2004; Guagliardo, 2004; Hansen, 1959; Joseph and Phillips, 1984; Kwan and Weber, 2003). The choice of the function and the β coefficient can be adequate for a specific region and industry. Yet, the transferability of the chosen function and coefficient to other regions or industries may not be appropriate, because the pattern of demand-to-supply contacts can vary (Joseph and Phillips, 1984).

Considering the health care industry and the β coefficient, examples of data that could be used to estimate β are patient-doctor contacts data, which are utilization data. However, these utilization data can only be obtained (i) from physician records, which are generally not available due to confidentiality concerns; (ii) through surveys, which are time consuming and can be costly (and biased); or (iii) from insurance payment sources, which are very expensive (Luo, 2004). Section 1.1 and subsection 2.1.2 already mentioned these utilization data problems that may render impossible the application and the dissemination of the methods that require the collection of such data.

Guagliardo (2004) mentions two additional problems with the gravity model. A key problem is that the A_i^{H} index is difficult to interpret and is not considered intuitive for policymakers, who prefer to think of spatial accessibility in terms of the most simple access measures, such as supply-to-demand ratios.

Another problem noted by Guagliardo (2004) is the fact that this model considers only supply and does not adjust for the demand of each supply point. A high spatial accessibility may be associated to an area according to the A_i^{H} index, because there are many supply points nearby with a high number of supply resources. However, the potential access of the population residing in that area will be lower than expected if the neighboring areas are highly populated and their populations compete for the same supply points.

An enhanced gravity method that addresses this last problem is the gravity-based model adjusted for demand, i.e., the gravity method with integrated demand, which is presented in equation (3). Weibull (1976) introduced the notion of populations competing for the supply in his study of accessibility to employment opportunities, and Joseph and Bantock (1982) applied the method to assess health care accessibility.

Equation (3) is an improved equation (2):

$$A_i^{\mathrm{G}} = \sum_{j=1}^{n} \frac{S_j d_{ij}^{-\beta}}{\sum_{k=1}^{m} P_k d_{kj}^{-\beta}}; \qquad (3)$$

where A_i^{G} is the gravity-based index adjusted for demand, m is the number of population areas, P_k refers to the population size in area k, while the other variables are the same as in equation (2). Population is used as a proxy for demand, the same way as in equation (1) (the supply-to-demand ratio, presented in subsection 2.1.3.1).

2.1.3.3. The kernel density estimation method

Another known method to assess access is the kernel density estimation, or just kernel estimation method. It is a variant of the gravity model, and both models can be considered the two main spatial analytical methods for calculating spatial access to supply points (Schuurman, Bérubé, et al., 2010; D.-H. Yang et al., 2006). The roots of this method can be traced back to studies in mathematical statistics in the 1950s and 1960s (Porta et al., 2009).

The application of the kernel estimation method involves spreading a grid over the surface of the chosen study region and using a moving function to estimate the intensity of events at each grid cell (events are the supply points and the populations). The moving function weighs events within its region of influence according to their distance from the point where the intensity is estimated

(Borruso, 2005; Delmelle et al., 2010; Gatrell et al., 1996). The grids are arranged in rows and columns, and each grid cell has the same size (D.-H. Yang et al., 2006). Access is assessed with the intensity value that is calculated for each cell.

The following equation, that became popular in the literature, is the general form of the kernel density estimator described by Parzen (1962). It estimates the intensity for a vector location x:

$$\hat{f}(x) = \frac{1}{nh}\sum_{i=1}^{n} K\left(\frac{x - X_i}{h}\right); \qquad (4)$$

where $\hat{f}(x)$ is the intensity estimate, $K(\ )$ represents the weighting function, which is centered on x (a vector location of x,y coordinates), X_i is the location of the event i, and h>0 is the bandwidth parameter that determines the spread of the kernel function, defining the region of influence (Borruso, 2005; Gatrell et al., 1996; Parzen, 1962; Seaman and Powell, 1996).

The intensity estimate is graphically three-dimensional, forming a cone. The resulting area of influence is determined by the parameter h, which descends from the highest point in the center, on each location x, until the borders of each corresponding window. The value of h determines the degree of smoothing: the higher its value, the smoother, less disruptive is the spatial variation among the access values assessed for all locations inside the region of influence. Considering supply points as the events, peaks represent areas of a high density of supply points, and valleys correspond to areas of low density (McLafferty and Grady, 2004). When cones overlap, overlapping areas receive an access score that results from the sum of the contributions of the overlying cones (Guagliardo, 2004).

Like the gravity methods presented before (equations (2) and (3)), the kernel estimation method of equation (4) depends on the definition of a function and of a parameter.

A typical function is the quartic kernel, which is commonly used to model decaying spatial access to health care facilities (Cooper et al., 2009a). This function sums only values of x-X_i that do not exceed h, such that the resulting region of influence is a circle of radius h with center on x. It calculates at each point x on the grid the distances to each observed event X_i within the catchment area, which contribute to the intensity estimate at x with associated weights related to the proximity to x (Gatrell et al., 1996; Schönfelder and Axhausen, 2003). This creates a density of the events, which are discrete points, as a continuous field, and the access scores can be calculated dividing, for example, population density estimates by supply point density estimates (Porta et al., 2009; Spencer and Angeles, 2007).

However, the choice of the grid constitutes a disadvantage of the method, because of the lack of interpretability, since there is no association with existing geographic divisions, and due to the arbitrariness of the choice of the grid cells dimension, the orientation of the grid and the origin of the grid (Borruso, 2005).

The choice of the bandwidth can also be considered a drawback, because it influences results. It has a stronger impact on results than the choice of the function (Porta et al., 2009; Schuurman, Bérubé, et al., 2010).

Finally, a key drawback is that the method of equation (4) does not consider transportation access, because straight-line or Euclidean distances for the radius of the circles around each point are used. These distances can be considered an imperfect proxy for travel friction (Guagliardo et al., 2004; McLafferty and Grady, 2004; Schuurman, Bérubé, et al., 2010).

2.1.3.4. The 2-step floating catchment area method

The 2SFCA method produces simple supply-to-demand ratios. It is a special case of a gravity-based method (with a travel-friction coefficient of 1 inside the catchment area and of 0 outside). It is also similar to the kernel estimation method in that a window is moved across the study region (in fact, it is a catchment area that floats across the study region) (Luo and Wang, 2003; D.-H. Yang et al., 2006). Furthermore, it produces easily interpretable access scores and it can be applied with the consideration of transportation access and travel times, instead of straight-line distances.

In addition, the 2SFCA method was proved to perform better than the gravity-based method applied to health care access by Joseph and Bantock (1982), and the kernel estimation method applied by Yang et al. (2006) in a comparison of GIS-based methods to measure spatial accessibility to health services. In this regard, Luo and Wang (2003) note that the gravity-based method of Joseph and Bantock (1982) tends to give higher scores to areas with low accessibility and may conceal local pockets of poor accessibility. Moreover, Yang et al. (2006) show that the kernel estimation method's scores can form a disruptive pattern of access and are not useful in identifying low accessibility areas, as opposed to the smoothing pattern of the 2SFCA method's ratios.

The 2SFCA method and 2SFCA-based methods use floating catchment areas. As explained in the introduction, catchment areas are market or customer geographic areas that reflect the consumers' preferences and their buying patterns. The boundaries of the catchment areas are defined by the maximum distances or travel times that the consumers are willing to travel to reach the products and services that they need or desire. In other words, they are defined by the propensity of the consumers to travel for obtaining specific products or services. In turn, floating catchment areas are catchment areas that float, in the sense that a series of them are computed for all the geographic assessment units of a chosen study region, in order to produce a supply-to-demand access score for each unit.

The floating catchment areas are used in two steps. In the first step, the catchment areas are defined for the supply points' locations. In the second step, the catchment areas are defined for the populations' locations, namely the geographic units in which they reside.

The application of 2SFCA-based methods first requires the choice of the specific study region and of the small geographic units of that region that will be used for the assessment. The study region is the whole region of interest for the researcher, namely the region for which the researcher needs or wishes to identify differences in access among populations. Examples of study regions can be a county, a district, a state, a country, and a group of countries or a continent.

The reason for choosing small geographic units for the assessment is that, by using small units, it is possible to carry out a fine-resolution assessment, i.e., an assessment with good geographic resolution (Luo and Wang, 2003; Wing and Reynolds, 1988). The smaller the geographic unit, the finer is the geographic resolution. However, there is not a unique, ideal size. As Wing and Reynolds (1988) propose in their study of availability of physicians services, the chosen geographic units should be small enough to provide adequate resolution. Thus, it has to be acknowledged that the choice of the geographic assessment unit involves a certain degree of subjectivity and it depends on the researchers' sensibility of what fits best their analysis, considering their objectives. Examples of geographic units that are typically chosen are neighborhoods, cities, municipalities or postcode areas that can capture the distribution of the population across states or countries – and for which there are available data. The availability of data is essential for the choice of the unit, because any choice that does not meet this condition is impractical and renders assessment impossible.

Besides enabling a fine-resolution assessment, the use of small geographic units presents two additional advantages. Firstly, it avoids or alleviates the access level misrepresentation that can happen if one adopts large geographic assessment units (problem (ii) of the supply ratios – see subsection 2.1.3.1, page 24). This is because the adoption of large units will ignore the existing, possibly pronounced differences in access levels across the populations that reside in the enclosed small units, which together form the large geographic units. Figure 1 presents an illustrated example.

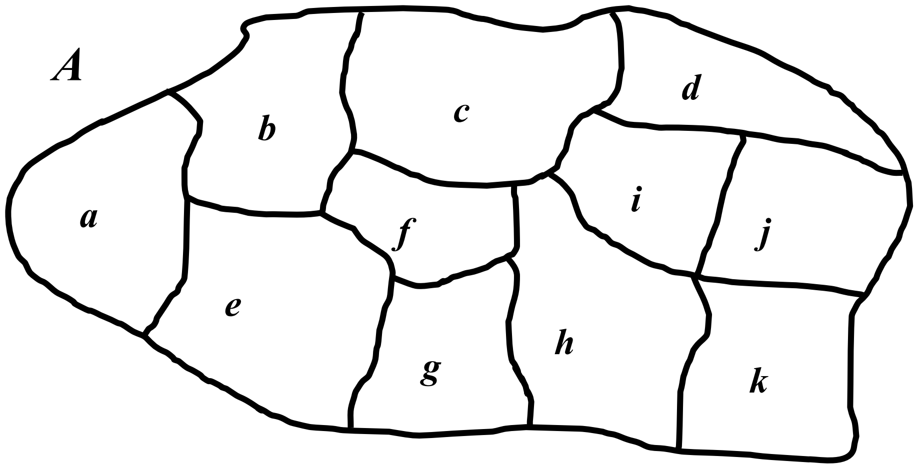

Figure 1 – The 2SFCA method uses small geographic assessment units

As it is illustrated in Figure 1, if a researcher studies the access level of the large study region *A* without considering probable access level variations in the small areas *a* to *k*, then he/she will misleadingly consider that access is homogeneous across the populations that reside in the small areas *a* to *k*. The access level variations can occur because of the different distances that the populations *a* to *k* have to travel to obtain the products and services, and because of different needs and other non-spatial characteristics that the different populations may have.

Secondly, an additional advantage of using small geographic units is that the fine-resolution results can be aggregated to any level deemed appropriate by the researcher (Love and Lindquist, 1995). For that purpose, the researcher can apply cluster analysis on the access scores, for example, as proposed by Makuc et al. (1991), such that the small-area results are aggregated and expressed in levels (for example, low, medium and high access). Alternatively, the researcher can use other thresholds for the identification of shortage regions, i.e., groups of small areas with low access results. As remarked by F. Wang and Luo (2005), shortage regions can be expanded or contracted, as desired. Besides, if fitted in well-known administrative regions, such as districts and municipalities, for example, it is easier to convey these results to policymakers, such that action can be taken to reduce access inequalities by promoting the redistribution of supply resources.

The 2SFCA method was first applied to the health care industry, but it also has been applied to other industries, as mentioned before. In fact, the first studies with floating catchment areas were assessments of job accessibility, such as Peng (1997) and F. Wang (2000). However, as noted in Luo and Wang (2003), one problem with these first floating catchment studies was that they erroneously assumed that all products provided by the supply points in a catchment area would be fully and exclusively available to the residents inside that area. Luo and Wang (2003) note that, instead, supply points near the catchment area's border may not be completely available to consumers residing in that area, if they also serve nearby consumers from outside (problem (i) of the supply ratios – see subsection 2.1.3.1, page 24).

Furthermore, in these first floating catchment area applications, people near the border limit can reside too far away from some supply points, even though all events are inside the same catchment. In fact, some consumers may reside at a distance higher than the defined threshold in relation to one or more supply points, such that these supply points should in fact be considered inaccessible for them. Therefore, the calculated supply ratios to assess spatial accessibility will be false.

Figure 2 explains these two problems of the early versions of the floating catchment area method with an example. In the figure, it is possible to see two catchment areas with a straight-line distance of 15 miles, from their centers until their border limits. One catchment area has as its center the (population-weighted) centroid (geographic midpoint) of census area 2, and the other, the centroid of census area 3. Considering one person residing at each centroid of the census areas, and one physician working at each of the locations *a*, *b* and *c*, we have physician-to-population ratios of 1/8 and 2/5 for the catchment areas with centers at census areas 2 and 3, respectively (R_2 and R_3). The small areas chosen as geographic assessment units in this example are census areas.

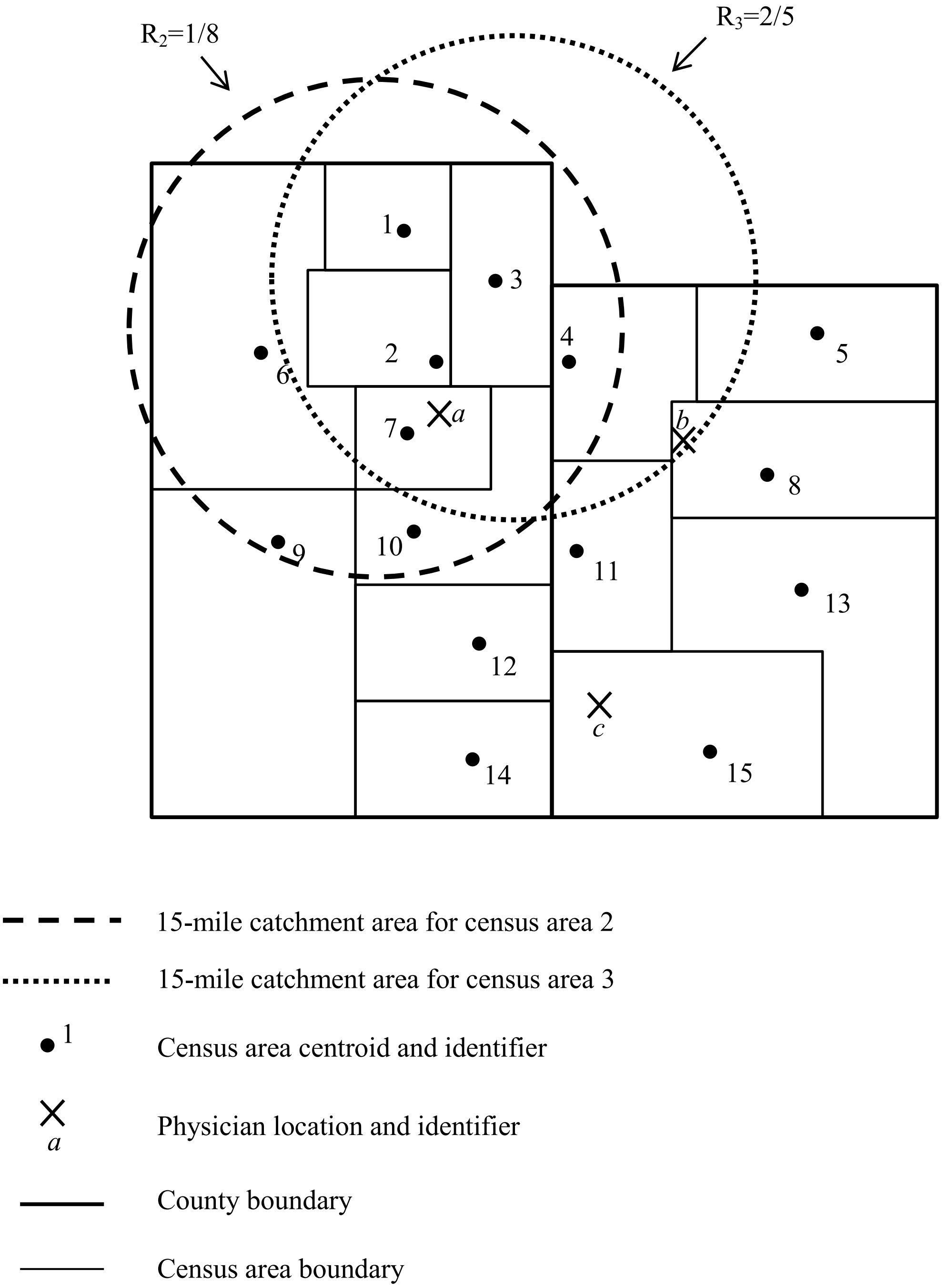

Figure 2 – An earlier version of the floating catchment area method, reproduced from Luo and Wang (2003)

Regarding the first above mentioned problem, Luo and Wang (2003) note that the physician at *b* may not be fully available to serve residents within the catchment R_3, because he or she will also serve nearby residents at 5, 8, or 11, who live outside the catchment.

As far as the second problem is concerned, we first recall that the boundaries of the catchment areas are defined by the maximum distances that consumers are willing to travel to reach the products and services that they need or desire: in this case, 15 miles. Accordingly, since the distance between the resident of centroid 1 and the physician at *b* is greater than 15 miles, as can be verified in the figure, the physician will be inaccessible for the resident. This way, the supply ratio of that catchment area with center at census area 3 (R_3) should not be 2/5 (but rather 1.8/5, because one of the physicians is available to only 4/5 of the patients that reside inside the catchment area).

The 2SFCA method of Luo and Wang (2003) corrects these problems of the early floating catchment area studies.

The first step of the original 2SFCA method considers the service catchment. It specifically models competition between populations for a supply point and is described by the following formula:

$$R_j = \frac{S_j}{\sum_{k \in \{d_{kj} \leq d_{\max}\}} P_k}; \tag{5}$$

where j and k are geographic locations, R_j is the supply-to-demand ratio for the supply point in j, S_j is the number of supply resources available at the supply point in j (physicians or beds, in the case of health care, for example), d_{kj} is the travel time between k and j, P_k is the population that resides in the geographic unit geo-referenced at k, and $d_{\max}$ is the time limit that populations travel to reach a supply point, which defines the boundaries of the catchment areas.

Thus, in this first step of the 2SFCA method a supply-to-demand ratio is calculated for each supply point of the study region by dividing the number of its supply resources by the sum of the size of the populations that it covers, i.e., the populations that reside inside the supply point's catchment area. This provides a preliminary insight of the spatial accessibility of the populations that reside within each supply point's catchment area.

From equation (5), it can be concluded that the 2SFCA method depends on the determination of $d_{\max}$. Luo and Wang (2003) chose 30 minutes as the reasonable or acceptable maximum travel time for primary care. However, other studies that applied 2SFCA-based methods to assess access to other kinds of health care chose different values for the threshold. Known orientations for the definition of $d_{\max}$ are 30 minutes for primary care, urgent health services or general care for adults and children, 45 minutes for obstetrical services or radiotherapy, and 90 minutes for general surgeries (Fortney et al., 2000; Hughes et al., 1981; UKDH, 2007). When the researcher is not able to collect and analyze travel behavior data (i.e., utilization data), in order to define the boundaries of the catchment area, it is admissible to use already established orientations (Bullen et al., 1996).

The geographic locations k and j can be, for example, centroids of the small geographic assessment units chosen by the researcher or, in the case of the supply points, their exact addresses (a very precise geo-referentiation of exact addresses depends, however, on the GIS that is adopted and the data that are collected).

Figure 3 illustrates a typical example of a step 1 catchment area. In this example, the catchment area has boundaries of 30 minutes travel times with center at the exact physician location c. It covers the populations residing at the census areas 10, 11, 12, 14 and 15, because the roads travel times from the centroids of these units until the physician location c are less than 30 minutes (these travel times are calculated with a GIS, usually considering minimum point-to-point travel times with average or maximum speeds).

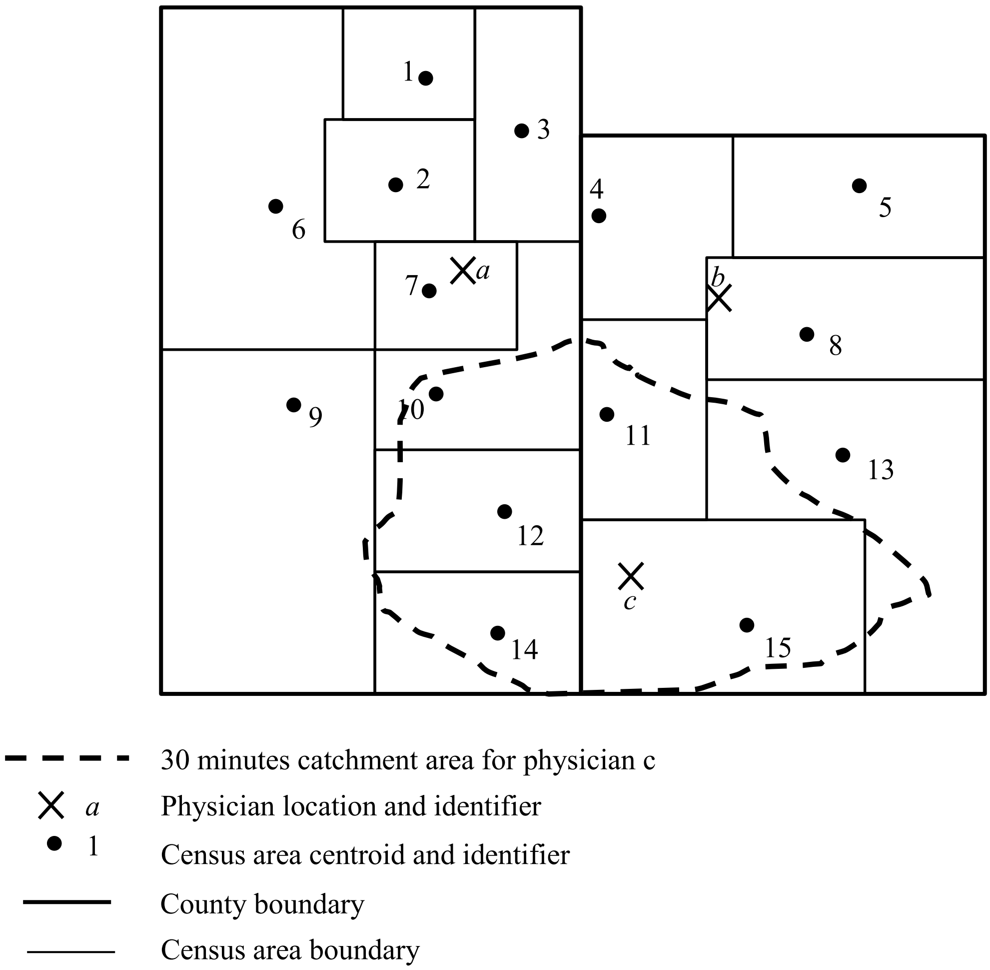

Figure 3 – Example of a 30 minutes step 1 catchment area

This example illustrates an approximated calculation, because it uses the centroids as references for the calculation of the travel times. We remark that with small geographic units this approximated calculation is considered to be admissible, because it is almost always impossible to have precise location data of all residences and supply points for an exact geo-referentiation using GIS.

The second step of the 2SFCA takes into account the population catchment. It accounts for the competition between supply points for populations and is represented by the following equation:

$$A_i = \sum_{l \in \{d_{il} \leq d_{\max}\}} R_l \, ; \qquad (6)$$

where A_i is the spatial accessibility score calculated for the population that resides at i, which is covered by the supply points located in l.

Figure 4 presents a typical example of a step 2 catchment area, with center at the population of census area 11 (the centroid of the census area).

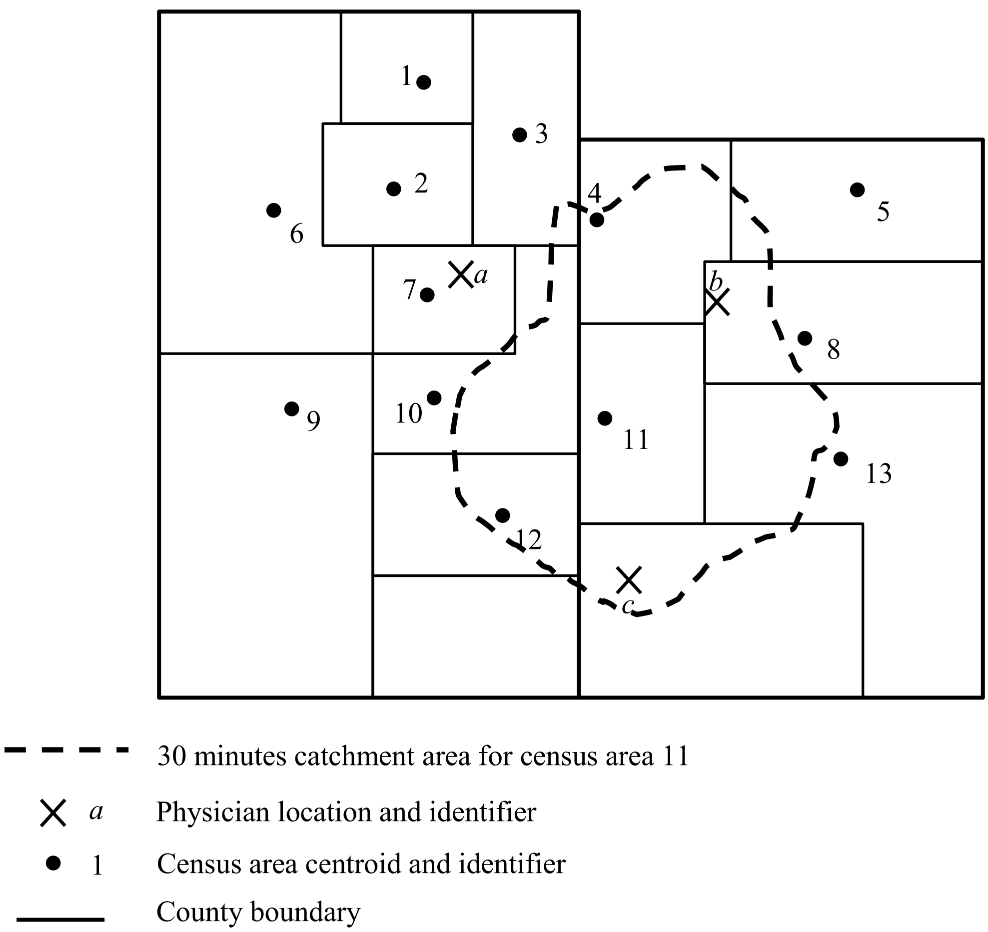

Figure 4 – Example of a 30 minutes step 2 catchment area

In this example, the step 1 supply ratios calculated for the physician locations *b* and *c* are summed up to produce, as a result, the spatial accessibility score for the population residing at the census area 11. Hence, in step 2, a score is computed for each geographic unit of the study region by summing all the step 1 supply ratios associated to the supply points that can be reached by the respective population until the defined maximum travel time. This score represents the spatial accessibility of each population.

Finally, it has to be noted that, while step two accounts for the competition between supply points for populations (equation (6)), the 2SFCA method does not include an assessment of competition between the firms that own the supply points. This is also valid for the enhanced versions of the 2SFCA method, which are presented next. The exception is the proposed EKD4SFCA method.

2.1.3.5. Enhanced versions of the 2SFCA method

The improvements to the 2SFCA method that were developed are mainly of three kinds:

(i) Non-spatial access: the consideration of non-spatial variables, to enable a comprehensive access assessment, instead of only a spatial accessibility assessment;

(ii) Variable catchment areas sizes: the consideration of variable catchment areas sizes, since, in urban areas, relatively smaller catchment areas would probably better represent the travel times of the populations when accessing health care; and

(iii) Distance decay function: the addition of a distance decay function, to measure distance decay or impedance, in order to differentiate higher travel times from smaller travel times, in line with Tobler's first law of geography and Jarvis' law (see problem (iii) of the supply ratios, in subsection 2.1.3.1, page 24).

The first, most relevant enhanced 2SFCA methods that introduced these improvements are presented next, beginning with enhancement (i), namely the consideration of non-spatial variables.

(i) Non-spatial access

After the 2SFCA method of Luo and Wang (2003), the first and most important development that can be found in the literature is the introduction of non-spatial variables by F. Wang and Luo (2005). F. Wang and Luo (2005) considered spatial and non-spatial factors in an analytical framework using the 2SFCA method. Based on a literature review, and particularly on Field (2000), the authors chose 11 non-spatial variables and grouped them into five classes: demographic variables affecting health needs; socioeconomic status; environment; linguistic barrier and service awareness; and transportation mobility. Considering that these variables might be correlated, the authors employed factor analysis to reveal the relevant non-spatial access dimensions. Principal component analysis was used as an initial step of the factor analysis to help determine the number of factors to be included in the analysis, in order to assess the relative importance of the variables. Then, with factor analysis, the 11 non-spatial variables were consolidated into three factors, for easy interpretation and mapping, labeled as:

- "Socioeconomic disadvantages" – the most important factor, capturing six variables;
- "Sociocultural barriers" – a less relevant factor that considered three variables; and
- "High healthcare needs" – which were found to be the least relevant factor, capturing two variables.

For the final assessment, the same steps 1 and 2 of the 2SFCA method were used (equations (5) and (6)). The spatial accessibility scores of the 2SFCA method and the three non-spatial factors' scores were analyzed conjointly, taking into account the following criteria:

- Areas with a spatial accessibility index of less than 1/3500 were considered "areas of poor spatial access";
- Areas with a spatial accessibility index of less than 1/3000, but greater than 1/3500, with healthcare needs scores one standard deviation above the mean value, were considered "areas of marginally poor spatial access with high needs";
- Areas with socioeconomic disadvantages scores one standard deviation above the mean value were defined as shortage areas and labeled as "areas of disadvantaged populations"; and
- Areas with socioeconomic disadvantages scores of less than one standard deviation above the mean, but greater than 3/4 standard deviations above the mean, and with sociocultural barriers scores one standard deviation above the mean were also defined as shortage areas and labeled as "areas of marginally disadvantaged populations with sociocultural barriers".

The thresholds of 1/3500 and 1/3000 were used because of the definition of Health Professionals Shortage Areas (HPSA) from the U.S. Department of Health and Human Services. According to such definition, HPSA are geographic areas with physician-to-population ratios of 1/3500 or less, or also with physician-to-population ratios close to 1/3000, if the population has unusual high needs.

However, we note that, in their study, F. Wang and Luo (2005) do not present an integrated method that incorporates the non-spatial components into the 2SFCA method. We understand that a further enhanced 2SFCA-based method should integrate spatial and non-spatial components into the model.

(ii) Variable catchment areas sizes

It is recognized in the literature that there are differences in travel behavior among urban and rural populations: urban populations generally do not travel to rural edging supply points, because they have a large choice of nearby options (more opportunities). This is in accordance with Stouffer's concept of intervening opportunities. This concept states that "the number of persons going a given distance is directly proportional to the number of opportunities at that distance and inversely proportional to the number of intervening opportunities" (Stouffer, 1940: 846). Thus, some applications of 2SFCA-based methods use catchment areas with varying sizes, i.e., with boundaries defined by different maximum travel times. Smaller catchment areas are defined in urban settings and larger catchment areas are set in rural environments.

The enhanced 2SFCA method that first introduced differentiation of catchment areas sizes is the one of McGrail and Humphreys (2009a, 2009b, 2009c). The authors introduced a capping function to the summation of the 2SFCA method's step 2, denoted by {100,10min}, such that each catchment area is capped to the nearest 100 supply points or 10 minutes for areas with high population density (McGrail and Humphreys, 2009a). This produces in practice smaller catchment areas sizes for urban areas with high population density, which have to be identified prior to the application of the method.

Another less arbitrary option shall be to consider adjustments in the access scores that take into account Stouffer's concept of intervening opportunities. For example, the variation of the mobility of the populations across the study region could be taken into consideration in the calculation of the access scores, with variables that take into account intervening opportunities in an approximated way. If the population of an urban area does not travel to alternative supply points in more rural areas, even though its catchment area covers these rural areas, then the access score of that population should be adjusted for this differentiated mobility. We understand that a further enhanced 2SFCA-based method could adopt this alternative.

(iii)Distance decay function

Finally, regarding the distance decay functions, these functions' rationale is Tobler's first law of geography, which states that everything is related to everything else, but near things are more related than distant things, as we recall from the beginning of this chapter. The use of distance decay functions is also supported by Jarvis' law, according to which patients will travel more to supply points that are nearer to their residences, than to supply points that are more distant. Therefore, it is important to differentiate distances inside the catchment areas, which is something that the original 2SFCA method does not do, because it considers uniform access within catchment areas. A distance decay function must be included in the access scores' calculations, in order to emulate distance decay inside the catchment areas and adjust the final access scores.

The enhanced 2SFCA (E2SFCA) method of Luo and Qi (2009) is the first example of 2SFCA-based method that introduces distance decay. Luo and Qi (2009) address the problem of uniform access within catchment areas by applying weights to three different travel time zones to account for travel impedance.

Step 1 of the E2SFCA method is represented by the following equation:

$$R_j = \frac{S_j}{\sum\limits_{k\in\{d_{kj}\in D_r\}} P_k W_r} = \frac{S_j}{\sum\limits_{k\in\{d_{kj}\in D_1\}} P_k W_1 + \sum\limits_{k\in\{d_{kj}\in D_2\}} P_k W_2 + \sum\limits_{k\in\{d_{kj}\in D_3\}} P_k W_3}; \qquad (7)$$

where d_{kj} is the travel time between area k and location j, D_r is the r_{th} travel time zone within the catchment area (r equals 1, 2 or 3), P_k is the population in area k, S_j is the number of health professionals at location j, and W_r is the weight derived from the Gaussian function for the r_{th} travel time. Three travel time zones with minutes' breaks of 0-10, 10-20 and 20-30 (zones 1, 2 and 3, respectively) are computed.

Step 2 of the E2SFCA method is modified accordingly, to take into account three different time zones:

$$A_i^{\mathrm{F}} = \sum_{j \in \{d_{ij} \in D_r\}} R_j W_r = \sum_{j \in \{d_{ij} \in D_1\}} R_j W_1 + \sum_{j \in \{d_{ij} \in D_2\}} R_j W_2 + \sum_{j \in \{d_{ij} \in D_3\}} R_j W_3 \, ; \qquad (8)$$

where A_i^{F} is the spatial accessibility score computed by the E2SFCA method.

Thus, Luo and Qi (2009) consider uniform access within each of the three travel time zones inside the catchment areas. A drawback of this is the debatable choice of division breaks, but more recent studies circumvented this by using a continuous function to calculate weights for all travel times inside the catchment areas.

One of the most recent improved 2SFCA-based methods that considers distance decay for all travel times inside the catchment areas is the one presented in the study of Dai and Wang (2011). This study synthesized two GIS-based accessibility measures into one framework, creating the kernel density 2-step floating catchment area (KD2SFCA) method, and applied the method to measure the spatial access to food stores in southwest Mississippi.

As mentioned in subsection 1.2, the KD2SFCA method of Dai and Wang (2011) is used as the basis for the creation of our enhanced 2SFCA-based method, which is presented in the next chapter. The KD2SFCA method includes a kernel density function in each of the steps of the original 2SFCA method to model the distance decay between the residents and the providers, being able to take into account the varying proximity within catchment areas (Dai and Wang, 2011). In step 1, the function rescales populations according to their travel time to each supply point and, in step 2, the function rescales the step 1 supply-to-demand ratios according to the travel time of each population to its corresponding available supply points.

Equation (9) summarizes the two steps of the KD2SFCA method into just one equation:

$$A_i^{\mathrm{K}} = \sum_{l \in \{d_{il} \le d_{\max}\}} \frac{S_l f(d_{il}, d_{\max})}{\sum_{k \in \{d_{kl} \le d_{\max}\}} P_k f(d_{kl}, d_{\max})}; \tag{9}$$

where $f(d_{il},d_{max})$ and $f(d_{kl},d_{max})$ refer to the distance decay function, and the other variables are roughly similar to those of the original 2SFCA method presented in equations (5) and (6) (see subsection 2.1.3.4).

The distance decay function chosen by Dai and Wang (2011) is the Epanechnikov function:

$$f(d_{ij}, d_{\max}) = \begin{cases} \frac{3}{4}\left[1 - \left(\frac{d_{ij}}{d_{\max}}\right)^2\right], & if \quad d_{ij} \le d_{\max}; \\ 0 \quad , & if \quad d_{ij} > d_{\max}. \end{cases} \tag{10}$$

The Epanechnikov function is a default function of the geographic information system application that the authors used (ArcGIS). However, we note that other functions could be adopted, and the best choice for a further enhanced 2SFCA-based method shall be a function that reflects the consumers' preferences when traveling to the supply points. This way, each industry will have its ideal function, since consumers in diverse industries will have different preferences and a different behavior regarding proximity and, accordingly, they will ponder the varying proximity within catchment areas differently.

2.2. Competition between firms

Competition assessment can be simply defined as the assessment of the firms that are selling products or services in a market, and the examination of their different performances and relative positions in the market. As explained in the introduction, competition assessment is important for firms to know their competitors and the markets, in which they compete, or in which they plan to enter. For public administrations, competition assessment is relevant and useful to identify geographic markets and regional differences in competition between firms. For regulatory agencies and competition authorities, who need to evaluate mergers and acquisitions, competition assessment is conducted to estimate market structure changes and decide whether or not to approve these concentration operations.

As mentioned in section 1.2, competition assessment resembles access assessment in some aspects. Both depend on the definition of the specific products or services that will be studied and on the relevant geographic areas that will be considered. Besides, the two kinds of assessment consider supply and demand and, in industries in which consumers need to travel to supply points to obtain the desired products or services, like the health care industry, both have to consider the attraction of consumers to supply points.

Early research on competition followed the structure/conduct/performance (SCP) framework, which is still at present an important analytical framework. Many SCP studies focus on the existing causal relationship between market structure, the conduct of the firms in the market, and market performance, measured by prices and/or profits, for instance. The main SCP studies seek to explain firm performance through market structure conditions, such as number and size distribution of firms and entry condition in the market (Halbersma et al., 2011; Hellmer and Wårell, 2009; Propper, 2012). Market structure can be defined as the way in which markets or industries are organized, and this depends on the number of the firms in the market or industry and on the extent of market control of each firm (Pulaj and Kume, 2013).

The main types of market structure are monopoly, perfect competition, oligopoly and monopolistic competition. Perfect competition and monopoly are two polar cases. Under perfect competition, each firm's production of a good is very small relative to the market, such that its sales do not affect the market price of the good, or other characteristics of the good (quality, for example). Under monopoly, there is only one firm in the market; hence, the market will have no competition and the firm will create entry barriers to guarantee its exclusive position (Motta, 2004; Pepall and Richards, 2002; Walker, 2006). In between perfect competition and monopoly is oligopoly, a market with a small number of dominating firms and entry barriers, and monopolistic competition, characterized by many firms in the market, but imperfect competition and some barriers to entry.

The competition assessment that follows the SCP framework considers that a more concentrated market indicates greater market power, that is, a higher ability of a firm to influence aspects of competition. This is because it can be established that in more concentrated markets firms usually have greater market power, that is, they face less competition (Nakamba et al., 2002). As a matter of fact, the highest market concentration can be identified in a monopoly market structure, which is the case in which there is no competition, because there is only one firm in the market. Accordingly, the SCP framework considers that more concentrated markets reflect less competition (CC, 2013; Halbersma et al., 2011; Hellmer and Wårell, 2009).

Thus, following the SCP line of research for competition assessment, the researcher assesses market concentration to conclude on the degree of competition in a market. Market concentration assessment metrics or methods depend basically on the number of firms that are active in the market and their respective market shares, i.e., the shares of the total production or sales. Low market concentration indicates that there is strong competition between firms, and high market concentration is an indicator of weak or lack of competition.

Many studies, reports and competition authorities' guidelines establish that low market concentration, reflecting a strong competition between firms, promotes productivity and efficiency. Strong competition has been also proved to spur innovation and to improve the quality of the provided products and services, among other benefits for industries, firms, markets and consumers (CC, 2013; EC, 2004; Gaynor et al., 2012a; Ikkersheim and Koolman, 2012; Motta, 2004; OECD, 2011, 2012a; OFT, 2007).

Conversely, excessive market concentration reflects weak or lack of competition, which may produce negative effects, such as, for example, excessively high prices, low quality products, diminished variety of products, restriction of freedom of choice, restricted access of firms to essential resources or infrastructure, and predatory pricing (CC, 2013; EC, 2004; Gaynor et al., 2012a; Ikkersheim and Koolman, 2012; Motta, 2004; OECD, 2011, 2012a; OFT, 2007).

With increasing market concentration, market entry and expansion becomes more difficult and costly (Pepall and Richards, 2002; Siegfried and Evans, 1994). In highly concentrated markets, entrants typically have to face a high initial investment in the supply capacity and sometimes also high advertising costs, in order to be able to compete with the well-established incumbents. Besides, these highly concentrated and weakly competitive markets can be dominated by incumbent firms, who can make entry unprofitable or unsuccessful.

A dominant firm is a firm with substantial market power, which does not face sufficiently strong competitive pressure from other firms. It can act independently of competitors and consumers and influence parameters of competition on the market, namely market prices, output, innovation and the variety or quality of goods and services, for a significant period of time (Melnik et al., 2008; OECD, 2006b, 2012b; OFT, 2004b).

Even though, according to Article 102 of the Treaty on the Functioning of the European Union, the abuse of market dominance is prohibited, market dominance itself is not (EU, 2012). Therefore, dominance is not a sufficient condition for violation of the law. It constitutes however a necessary condition, and when this condition is verified, it may generate an impediment to existing competition, producing the above mentioned negative effects of excessive market concentration or lack of competition.

Dominant firms can strengthen their dominant position by adopting a strategic behavior to raise entry barriers and deter possible market entrants, i.e., the potential competitors (firms that are not competitors presently, but that may decide to invest and enter the market after some time, like one year, for example). They can also create barriers to the expansion of their effective competitors (the actual, present competitors). An example of a behavior to raise entry barriers can be overinvestment in capacity, as a sign of willingness to respond to new competition by making entry unprofitable (Gaynor and Haas-Wilson, 1999).

Thus, market dominance represents a risk for dominated firms of having their market shares reduced or of being forced to exit the market. For potential entrants, it represents a risk of strategic entry deterrence. Firms that wish to survive, grow and obtain profits should therefore focus on markets that are not dominated, or niches with no competitors, or that have been ignored by other firms (Dalgic and Leeuw, 1994).

In turn, public administrations that wish to produce the benefits of competition should try to promote competition in dominated markets. Conversely, only in non-dominated markets, with low market concentration, it is possible for public administrations to explore the existing competitive pressure between firms in public bids. This way, they can obtain better conditions when contracting private firms to promote supply for disadvantaged populations. They are able to contract more supply to reduce access inequalities at lower prices.

Competition assessment can be used to evaluate if there are competition risks in a market, but it traditionally requires a previous market definition. The identification of the relevant market serves an analytic purpose by identifying the area where the actual competition occurs.

2.2.1. The relevant market definition

The relevant market has two dimensions, namely the product dimension and the geographic dimension, and its definition usually begins with the identification of the firms and the products or services that will be analyzed (EC, 1997; OFT, 2004b).

As far as the product dimension is concerned, the following examples of information can be collected to confirm the relevant product market that is defined based on the initially identified firms and products:

- Products' characteristics;
- Marketing studies, market analyses and business plans;
- Information on the human resources; and
- Information on production methods or processes.

The definition of the relevant geographic market considers the geographic area that can be differentiated from other areas because of considerably distinct competition conditions, namely the area where companies supply their products and where the competition conditions are sufficiently homogeneous (Alexiadis et al., 2003). The first area to be chosen, as a candidate geographic market, can be defined considering the geographic distribution of the firms' supply points or a first impression about the geographic catchment of the firms.

The candidate geographic market can be confirmed with the collection and analysis of information on supply and demand characteristics, to verify if firms located in different areas constitute alternative supply points to consumers.

Examples of such information are:

- Switching costs for consumers to products from other areas;
- Differences in prices, sales and marketing strategies between areas;
- Identification of legal, natural or strategically created barriers that can affect the acquisition of products from other regions; and
- Buying patterns of consumers and the identification of their preferences.

A catchment area can be chosen as the relevant geographic market, because it reflects buying patterns and it identifies the consumers' preferred geographic areas, namely the areas to which they prefer to travel when buying the products or services that they desire or need.

The adoption of catchment areas to define the relevant geographic market should take into account the flow of consumers to supply points and consists of identifying the areas where the majority of clients reside (CC&OFT, 2010). Thus, considering catchment areas defined by travel times and the hospital health care sector in an example, if the majority of the patients reside up to 30 minutes of travel time to hospitals, then catchment areas of 30 minutes to each of the hospitals can be set as the relevant geographic markets.

Notwithstanding, theoretical catchment areas can also be defined, based on benchmarks that have been set in the literature, considering maximum travel times or distances that consumers are willing to travel from their residences to reach the supply points that sell the products or services that they need. The maximum distance that consumers are willing to travel has been known in geography as *range*, according to the 1933 central place theory of the German geographer Christaller (Cromley and Albertsen, 1993; Yao et al., 2013). The use of theoretical catchment areas is providential, because data on the flow of consumers, like patients' flow data, i.e., data on patients' discharges from health care supply points and their residential postcode areas, are almost always unavailable. Some examples of maximum travel times set in the literature, mostly related to health care markets, are 30 minutes for food stores, primary care, urgent health services or general care for adults and children, 45 minutes for obstetrical services, and 90 minutes for general surgeries or hospital health care (Dai and Wang, 2011; Fortney et al., 2000; Hughes et al., 1981; Polzin et al., 2014a). As seen in subsection 2.1.3.4, these are also orientations that can be used to set the maximum travel times for the catchment areas of the 2SFCA method.

Having defined the relevant geographic market, its combination with the relevant product market allows the researcher to finally conduct his/her competition assessment. If the researcher chooses catchment areas as the relevant geographic markets in his/her competition assessment, he/she will be able to identify that a firm competes with another firm if their catchment areas overlap (Goody, 1993; Pavlyuk, 2009). The geographic areas covered by overlapping catchment areas of two firms indicate that the residents of these areas have two firms to choose from, when they decide to acquire their needed products or services.

2.2.2. Competition assessment methods

The simplest competition assessment methods are simple metrics, such as the number of firms and the market shares. These metrics capture structural characteristics of the market and may reflect changes in the market concentration as a result of firms' entries, exits and mergers and acquisitions.

The most popular competition assessment methods with calculations based on the assessed market shares of the actual competitors in a market are the concentration ratio and the Herfindahl-Hirschman Index (HHI). These methods are widely used by competition authorities worldwide (ACCC, 2008; CC, 2013; Chua et al., 2011; Nakamba et al., 2002; OECD, 2012b; OFT, 2004a; Walker, 2006). While there are many other less known and less adopted metrics to measure market concentration, most of these metrics were, however, found to be highly correlated (Chua et al., 2011; Wong et al., 2005).

2.2.2.1. The concentration ratio

As defined by Pulaj and Kume (2013), the concentration ratio reflects industry and market structures and indicates the concentration of production management in an industry. Concentration ratios are the sums of the largest market shares of the firms that are active in a market. The concentration ratio of the N largest market shares is denoted as RN (or CRN, or also C_N).

The concentration ratio R4 was an important concept in competition analysis until 1982, when it was replaced by the HHI in the US merger guidelines (OECD, 2012b). Its scale ranges from almost zero, indicating no market concentration, until 100%, which is interpreted as an indication of a high market concentration.

Concentration ratios present two drawbacks. One drawback is that they do not take into consideration the relative sizes of the leading firms, i.e., of the firms with largest market shares. As noted by Walker (2006), a market with four firms that have each 20% market shares will have the same R4 (of 80%) as a market with four firms with market shares of 55%, 20%, 3% and 2%. In the latter case there is a clear leader in the market, probably dominant, while in the first case there may be strong competition among the four leaders, and the concentration ratio disregards this information.

The second problem is that these ratios convey less information than the HHI. They do not consider information on all market shares of the competitors in the market. Hence, they are not used as frequently as the HHI in the scientific literature (Wong et al., 2005).

2.2.2.2. The Herfindahl-Hirschman Index (HHI)

The widely known and adopted HHI was developed by Hirschman and Herfindahl in their 1945 and 1950 works, respectively (Hirschman, 1964; Rhoades, 1993). The HHI enables the calculation of levels of market concentration giving an indication of the competitive pressure between firms that are active in a market. Instead of using only partial information of the firms' market shares, as the concentration ratio does, it includes all firms in the calculation. Besides, the HHI reflects the degree of market share inequality across the spectrum of firms that compete in a market: the influence of smaller firms is lessened, while the influence of larger firms is accentuated (Claudia, 2012).

As described in equation (11), the HHI is calculated as the sum of the squares of the firms' market shares in the market that is being analyzed by the researcher.

$$\mathrm{HHI} = \sum_{i=1}^{N} \mathrm{Q}_i^2 \; ; \qquad (11)$$

where Q_i are the market shares of the firms 1 to N expressed in a basis of 100. The market shares of the firms are customarily calculated with reference to their volumes of sales of the relevant product in the relevant geographic market.

As in all competition assessments, this calculation is carried out after the relevant geographic market is defined, considering all the firms that are active in the market and located inside the specific defined region. Geographic markets for the calculation of the HHI are usually administrative areas, like countries, states and districts.

Typical absolute HHI thresholds define three market concentration classes: low, medium and high, or not concentrated, moderately concentrated and highly concentrated, respectively. An HHI until 1000 indicates that market concentration is low, which is equivalent to say that the market is not concentrated. If the HHI is between 1000 and 2000, then market concentration can be classified as medium, and the market is moderately concentrated. An HHI above 2000 and until the maximum value of 10000 indicates that market concentration is high, i.e., the market is highly concentrated (CC, 2013; McIntosh and Hellmer, 2012).

Approximate associations between HHI values and types of market structure can be established. The maximum value of 10000 identifies the case of pure monopoly. HHI values above 2000, but below 10000, indicate that the market structure shall be an oligopoly. Yet, an HHI close to 2000 gives an indication that the market structure can be a monopolistic competition. Finally, an HHI that tends to zero gives an indication that the market structure tends to perfect competition.

Besides being the most known and adopted measure of concentration, the HHI often serves as a benchmark for other concentration indices (Pulaj and Kume, 2013).

However, the problem of the HHI lies on its calculation for administrative areas (countries, states, districts or metropolitan areas, for example). The typical choice of relatively large relevant geographic markets for the calculation of the HHI presents the four problems that were previously identified for the supply ratios (see problems (i) to (iv) in pages 24 and 25, on subsection 2.1.3.1), which we recall below, now adapted to the assessment of market concentration:

(i) Disregard of border crossing of consumers. Market concentration is assessed in each area considering only the supply points located in that area. However, many people cross borders when traveling to a supply point, in order to obtain the desired products or services.

(ii) No variability in market concentration within the relatively large areas chosen as the relevant geographic markets. By using large areas, the HHI of equation (11) does not enable the detection of market concentration variations inside these areas.

(iii) No incorporation of any measurement of distance decay or impedance. The large areas used by the HHI treat large distances inside these areas the same way as small distances.

(iv) Vulnerability to the MAUP. This is because, depending on the large geographic areas chosen to assess competition, results for a specific small area – a municipality, a postcode area or a neighborhood, for example –, may present considerable variations.

Another problem, as is seen next, is that the HHI is applied at the industry level, such that it cannot be used to assess market dominance, i.e., if a firm has a dominant position in the market.

2.2.3. Identifying market dominance

As, seen, the main competition assessment methods are the above presented concentration ratios and the HHI. They can give an initial indication of the competitive pressure in the defined relevant market.

When these methods indicate that the market concentration is high, reflecting a weak competition between firms, it is possible that the market has a firm with a dominant position, so that it can dominate the other firms that are active in the market and create entry barriers for potential competitors. A firm can only be dominant if it has substantial market power, which is generally associated with high market shares. Thus, a high market share gives an indication that a firm may be dominant.

Therefore, the most common method to assess market dominance is the single information on the largest market share, which is equivalent to the concentration ratio of just one firm, R1. The largest market share is then compared to a threshold. Many thresholds are used and they typically range from 33.3% to 60% (Schmid and Ulrich, 2013; Strong et al., 2000). If the largest market share is greater than the threshold, then the firm dominates the market, i.e., it has a dominant position.

In some cases, the HHI can also be used to identify market dominance. For example, it is clear that an HHI of 10000 indicates monopoly, which is the case of only one dominating firm in the market. Other HHI values very close to 10000 can also be interpreted as representing a market concentration of a dominated market. The HHI is, however, applied at the industry level, such that in practically all cases it will not be possible to use the HHI to determine whether the market is dominated. Only a method that can be applied at the firm level, such as the single largest market share (the concentration ratio R1), is able to determine if there is market dominance, i.e., to identify whether a firm has a dominant position in the relevant market (and also to identify precisely which firm is the dominant firm).

However, the problem of using only the largest market share R1 to determine that a firm has a dominant position in the market is that it does not consider the market shares of the other firms that are active in the market (this is one of the problems of using concentration ratios, as seen in the previous subsection). Another problem of using R1 to establish the existence of market dominance is the fact that, as seen, there is no consensus on the dominance threshold: they range from 33.3% to 60%.

A method that is also applied at the firm level, but that introduced an improved identification of dominance, by including the assessment of the market shares of all firms, is the groundbreaking method of Melnik et al. (2008). By improving the identification of market dominance, this method provides a more comprehensive information on the risk of abuse of market dominance and, accordingly, on the possible creation of entry and expansion barriers by incumbents. The method of Melnik et al. (2008) is a competition assessment method that enables the identification of market dominance by considering how existing competition limits the ability of a firm to dominate the market. As shown in equation (12), it considers the market shares of all firms in the industry to calculate a dominance threshold.

$$Q_d = \frac{1}{2}\left[1 - \gamma(Q_1 - Q_2)\left(1 - \sum_{i=3}^{n} Q_i\right)\right] = \frac{1}{2}\left[1 - \gamma\left(Q_1^2 - Q_2^2\right)\right]; \quad (12)$$

where Q_d is the dominance market share threshold and $\left(Q_1^2 - Q_2^2\right)$ describes in a condensed way, with the market shares of the two firms with the largest market shares (Q_1 and Q_2), how existing competition exerts pressure on the firm with the largest market share, i.e., firm 1. If firm 1 has a market share that is higher than the dominance threshold that is calculated with equation (12), then it has a dominant position (Hellmer and Wårell, 2009; Melnik et al., 2008).

The parameter γ represents industry-specific entry barriers and its purpose is to capture the possible existence of regulation, barriers to entry, powerful buyers, among other specificities of the industry that may harm competition. Lower values of γ represent lower barriers and increased competitive pressure of potential competition, limiting the ability of firm 1 to dominate the market. Typical values of γ are ½, for low barriers, 2, for high barriers, and 1, which can serve as a natural benchmark, or a base case (McIntosh and Hellmer, 2012; Melnik et al., 2008). Considering this base case, the lowest dominance threshold is zero, and this occurs for a monopoly market structure, in which there is only one firm with 100% of market share. The highest threshold is 50%, which occurs when the two firms with the largest market shares have equal shares, such as in a duopoly, for example.

Finally, as in the case of the HHI, the market shares in equation (12) can be calculated with reference to their sales of the relevant product in the relevant geographic market, which is a relatively large area, such as a country, a state or a district. Thus, the dominance identification method of Melnik et al. (2008) has the same four problems of the HHI of equation (11), namely problems (i) to (iv) identified in subsection 2.2.2.2 (see page 55).

Therefore, we understand that only different kinds of calculations of the HHI and of the dominance identification of Melnik et al. (2008), which we however did not identify in the literature, could circumvent these problems.

2.3. Conclusions

We presented in this chapter the theoretical background and the methodology that is adopted to assess access to health care and competition between firms, This way, we lay the theoretical, conceptual and methodological foundation for the presentation of our proposed method, and aid the explanation and the consolidation of access and competition related concepts. We also identified opportunities for innovations to improve the assessment methods, in addition to the innovative integration of the assessments of access and of competition into our proposed EKD4SFCA method.

Regarding the assessment of access (section 2.1), we defined the access concept and its dimensions, according to the literature. While a definitive and consensual set of access dimensions cannot be identified, it is clear that the dimensions and respective variables used for assessments can be divided into two groups: spatial and non-spatial. The spatial variables comprehend variables that take into consideration the geographic location of the supply points in relation to the demand, the distance or travel times until the supply points, and the supply capacity that is available to the populations. The non-spatial variables are needed for a comprehensive assessment of access, namely an assessment that is not restricted to spatial aspects.

We also established the difference between potential access and realized access. Potential access to health care is the focus in this thesis and refers to the health care supply that is potentially available to populations. As explained before, the assessment of potential access is used to identify shortage areas, with a relatively low available supply capacity, where disadvantaged populations reside. The potential access approach is the ideal approach to assess access to health care when utilization data are not available, and the scarcity of data is a real obstacle for the assessment of access to health care.

The main methods that are used to assess potential access were also presented. Their characteristics and mathematical formulas were identified, highlighting the important role of GIS to enable the computation of access scores. In particular, the 2SFCA method and enhanced versions of this method were given special attention, since these methods are the most popular spatial interaction methods to assess potential access, with easily interpretable supply ratios. Some of their features include the requirements to calculate travel times, to define catchment areas and to use small geographic assessment units. We also underlined the main developments introduced in the 2SFCA-based methods, including the consideration of non-spatial variables, the addition of a distance decay function, and the use of catchment areas of varying sizes. In addition, we identified opportunities to adapt and/or extend these developments for the creation of a further enhanced 2SFCA-based method. More specifically this further enhanced method (i) should have incorporated non-spatial components, (ii) should consider Stouffer's concept of intervening opportunities, and (iii) should have an adequate distance decay function for the industry that is being considered for the assessment.

Regarding competition assessment, as we explained, the relevant market definition plays a central role, as a first step to establish the foundation for the assessment. The role of the relevant market is linked to the essential character of market shares in competition assessments. The calculation of market shares depends on the relevant market definition, both in terms of the products that are being considered and of the geographic region that is relevant for the researcher. Thus, both dimensions of the relevant market have to be defined: the product dimension and the geographic dimension.

Market concentration reflects the actual competition in the relevant market and it is assessed with different methods that depend on the identification of the market shares of the firms that are active in the relevant market. High market shares of the leading firms will lead to a high market concentration, which will indicate that competition is weak in the market. A situation in which there are many firms in a market with low market shares will indicate that market concentration is low and, accordingly, that the competition between the firms is strong.

We identified the main competition assessment methods. Apart from the HHI, which is the most common method to assess market concentration, we also presented a relatively new method that was created to assess market dominance, i.e., to assess if the leading firm has a dominant position in the market. This method is the market dominance identification method of Melnik et al. (2008).

While market concentration assessment is made at the industry-level, market dominance is a firm-level assessment, since it has to be established whether the leading firm is dominant (or leading firms are dominant). Identifying market dominance is a way of identifying the most biased situation that can occur in a market, namely the situation in which one or more leading firms with very high market shares can act independently from the other firms, if they wish, abusing of their dominant position. As we explained in this chapter, the abuse of a dominant position leads to negative effects in the markets, not only for the actual competitors, but also for potential competitors. The dominant firms may create entry barriers to avoid potential competition. Dominance abuse can also produce negative effects for the consumers.

The HHI and the dominance identification method of Melnik et al. (2008) enable a comprehensive quantitative valuation with modest information requirements. This is especially useful for markets with scarce data, where typical competition assessments would require restricted and possibly unreliable secondary or proprietary data sets, with missing data and inconsistencies (Bernstein and Gauthier, 1998).

However, we were able to identify four problems that arise from the way that these methods are calculated:

(i) Disregard of border crossing of consumers;

(ii) No variability in market concentration within the relatively large areas;

(iii) No measurement of distance decay or impedance; and

(iv) Vulnerability to the MAUP.

Therefore, we understand that changes to the calculations should be promoted, in order to overcome or alleviate these problems.

In the next chapter we present the methodological improvements that we promoted to the access and competition assessment methods, following the opportunities for enhancements and corrections of problems identified in this chapter. We also present the development of our integrated method for the assessment of access and competition, i.e., our proposed EKD4SFCA method.

3. The Extended Kernel Density 4-Step Floating Catchment Area (EKD4SFCA) method

In this chapter we present the development and the constituent parts of our proposed EKD4SFCA method. In short, it is a new, enhanced 2SFCA-based method that incorporates into it innovative calculations of the HHI index and the dominance identification method presented in subsections 2.2.2 and 2.2.3, respectively. As noted in the introduction, the construction of this method is the main objective of this thesis. It also marks the first time that the assessments of access and of competition are integrated into a method, and it constitutes the main innovation of the thesis and motivation for this work.

The integration of the assessments of access and of competition is expected to be useful for different institutions and diverse applications. In fact, the results of an integrated assessment can help in aligning the firms' supply decisions with their strategic and market positioning objectives. Moreover, the integrated assessment can enable the identification of target markets for entry, where there is relatively more unmet potential demand, and where no competition hindrance by incumbents is expected. For public administrations, the integrated assessment of access and competition, applied to the health care industry, is expected to be useful when defining public policies destined to improve the distribution of supply resources, reduce access inequalities and improve the populations' health status.

The EKD4SFCA method consists of four steps: the first two steps refer to our new, enhanced 2SFCA-based method. The third and fourth steps refer to the competition assessment that is carried out with our innovative HHI and dominance identification methods.

As in all 2SFCA-based methods, a GIS is required for the application of the method. The GIS supplies the speed maps, the traffic rules, the subdivisions of the study region, and the centroids of the subdivisions, which are needed to compute the travel times and to define the span of the catchment areas.

3.1. The EKD2SFCA method (steps 1 and 2 of the EKD4SFCA method)

The first two steps of the EKD4SFCA method consist of an extended KD2SFCA (EKD2SFCA) method that we created. The KD2SFCA method was presented in subsection 2.1.3.5 (see pages 44 and 45), and its two steps are the following:

$$R_j^{\mathrm{K}} = \frac{S_j}{\sum_{k \in \{d_{kj} \le d_{\max}\}} P_k f\left(d_{kj}, d_{\max}\right)};$$

$$A_i^{\mathrm{K}} = \sum_{l \in \{d_{il} \le d_{\max}\}} R_l^{\mathrm{K}} f\left(d_{il}, d_{\max}\right).$$

We developed the proposed new EKD2SFCA method by incorporating two additional components into each of the two steps of the KD2SFCA method – a health needs index and a commuting index, respectively (commuting is defined as the individuals' daily movements between their residences and their workplaces or places of study). Furthermore, we changed the KD2SFCA method's distance decay function, in order to model decaying proximity to health care (Polzin et al., 2014a).

The adaptations and extensions that we introduced in the KD2SFCA method stemmed from the exploration of the previously identified opportunities to create a further enhanced 2SFCA-based method (see section 2.3, pages 59 and 60), which we now detail:

(i) Incorporation of non-spatial components into the model: in order to incorporate non-spatial components into the model and improve the assessment of access to health care, we propose the construction and incorporation of a health needs index with principal component analysis and non-spatial variables.

(ii) Scores adjustment according to Stouffer's concept of intervening opportunities: we propose to adjust the access scores computed by the method, taking into account Stouffer's concept (see subsection 2.1.3.5). For that matter, we innovatively propose to use commuting variables, as representative variables of the mobility of the populations when accessing health care, and incorporate this information into the model with a principal component analysis index.

(iii) Choice of an adequate distance decay function: we did not find in the literature an application of a 2SFCA-based method that explicitly uses a distance decay function to adequately emulate the behavior of patients regarding the distances and associated travel times to health care supply points. For our enhanced 2SFCA-based method, we propose to use such a distance decay function to improve the assessment of access to health care.

Figure 5 summarizes in a flowchart the sequence of the method's development phases.

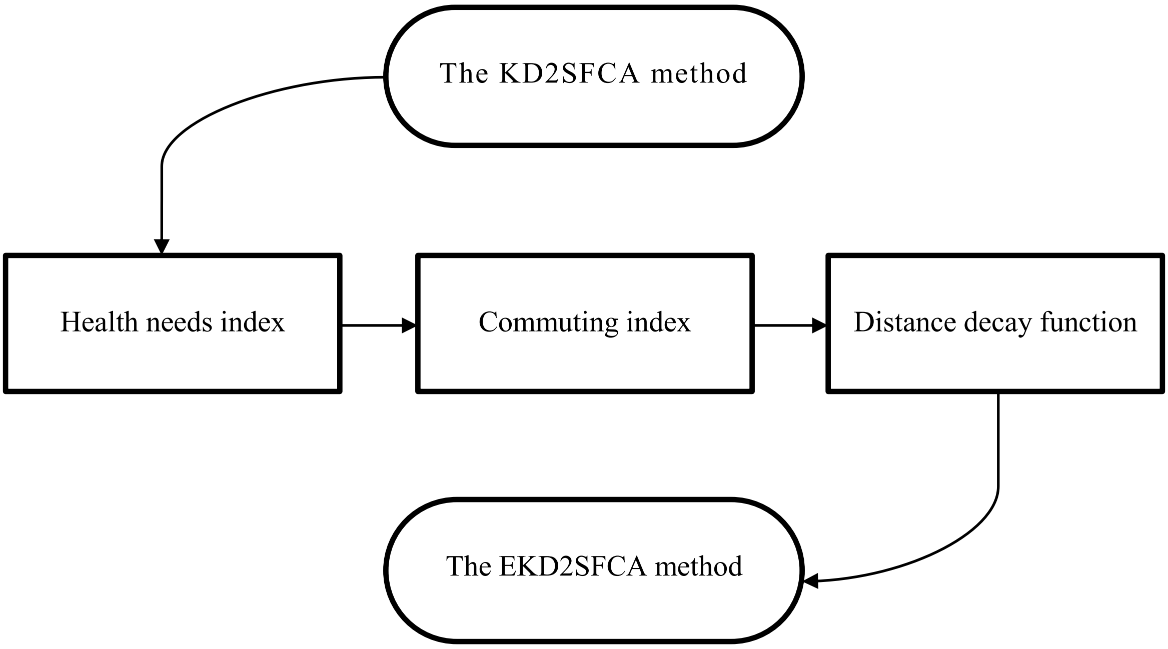

Figure 5 – The development process of the EKD2SFCA method

3.1.1. The health needs and the commuting indices

We propose to construct the health needs and the commuting indices in an original way, by means of principal component analysis (PCA), more specifically, with the first principal component of PCA.

3.1.1.1. Constructing an index with the first component of PCA

PCA is a statistical technique that is used to reduce the number of variables in a data set, in order to obtain a smaller number of dimensions, as components or indices (Jolliffe, 2002; Vyas and Kumaranayake, 2006). It synthesizes the information contained in a great number of variables by constructing new synthetic variables, namely the principal components, which are linear weighted combinations of correlated variables (Havard et al., 2008).

We propose the construction of each of the two indices – the health needs and the commuting indices – with the first principal component, following the methodology described, in particular, by Henry et al. (2003), Salmond et al. (2006) and Mooi and Sarstedt (2011). Such an index can be represented as:

$$PC_1 = a_{11}X_1 + a_{12}X_2 + ... + a_{1n}X_n\,; \tag{13}$$

where PC_1 is the first principal component, X_1, X_2 until X_n form the set of variables selected by the researcher, and a_{11}, a_{12} until a_{1n} are the coefficients, or weights, derived from the PCA (more specifically, from the correlation matrix of the data, or the covariance matrix, if the data are standardized) (Howe et al., 2008; Vyas and Kumaranayake, 2006).

The health needs and the commuting indices could be constructed with all components that result from the PCA. This would enable the consideration of more information to be incorporated into the model. Thus, the limitation of using only the first component is the consideration of less information than it would be the case if all components produced by PCA were used.

However, we identified advantages in constructing the indices with only the first principal component, which we understand that more than compensate the use of less information. The advantages are the following:

(i) Optimal summarized value: the first component of PCA is a linear index of all the variables selected by the researcher that attains the highest proportion of the explained variation of the original variables, i.e., the largest amount of information that is common to all the variables (Filmer and Pritchett, 2001; Neupane et al., 2015; Tareque et al., 2014). It can give the mathematically optimal value to summarize all variables selected by the researcher in a single combination (Schmidtlein et al., 2008). As it is acknowledged by different authors, in the cases in which the first component accounts for a large percentage of the variation in the variables, it can be used alone, without much loss of information (Manyong et al., 2006; Muchara et al., 2014).

(ii) Parsimonious construction: in accordance with the instructions of Humphreys (1998) for the development of an index, the construction of the index with just the first component of PCA is parsimonious and avoids unwieldy and impractical calculations. It is simpler to use just the first principal component.

(iii) Data-derived weights: the use of only the first component avoids the possibly arbitrary choice of the weights for computing an index with two or more components, because with one component there is no need to determine weights to combine components. The only weights in the index are the coefficients of the variables that are automatically data-derived via the PCA – from the correlation matrix of the data, or the covariance matrix, if the data are standardized (see equation (13)) (Bei and Cheng, 2012; Vyas and Kumaranayake, 2006).

(iv) Representation of the unique factor that the researcher wishes to include in the analysis: when the PCA produces more than one component, these components are uncorrelated with each other, and each will represent a different factor, as explained in subsection 2.1.3.5, when we described the study of F. Wang and Luo (2005) (see pages 40 and 41; in that case,

> the authors identified three factors, labeled as "socioeconomic disadvantages", "sociocultural barriers" and "high healthcare needs"). However, if the researcher uses PCA to produce only one significant component, he/she will be able to obtain the unique factor that he wishes to measure, by choosing variables that represent that specific factor (CGAP, 2000). In this case, any additional component will not be significant, since it will account for a very small percentage of the variation in the variables, and will be measuring a different factor, which was not planned by the researcher.

The process of construction of an index with the first component of PCA begins with the researcher selecting a comprehensive set of variables that are known to represent the desired factor in the scientific literature, or in econometric studies. This first phase is, thus, a screening phase to ensure that the selected variables present an association with the factor that the researcher wishes to measure (health needs or commuting, in our case), so as to not produce an index with a distorted measure.

Ideally, the selected variables should be balanced, to reflect different aspects of the desired factor, but there should be sufficient correlation among the variables (generally above 0.30), because PCA is based on the correlations between the items that are analyzed (Mooi and Sarstedt, 2011). Several variables reflecting similar aspects of the desired factor can be included in a PCA model, but a heavy concentration of similar variable types can inappropriately skew the resulting index to overstress one specific aspect (Henry, 2003). Therefore, if the correlations between two variables are too high (above 0.90, as a reference), it is advised to remove one of the variables, because they shall be measuring the same aspect.

As Henry et al. (2003) proposes, the initial set of variables should include at least 10, but no more than 20 variables, as a rule of thumb. In practice, only a small number of variables will be identified with PCA for a statistically valid index constructed with the first component.

Regarding the sample size required for statistically valid results, the number of observations for each variable should be at least ten times the number of variables used for analysis (Mooi and Sarstedt, 2011). Thus, if the researcher selects 10 variables for the initial set, the study region will have to consist of at least 100 small geographic assessment units, for which the data of the respective variables have to be collected.

After the researcher selects a suitable set of 10 to 20 variables, with adequate correlation between them and data with a proper sample size, the researcher has to follow an iterative approach with PCA and a continual scrutiny of the variables to determine which subset brings the best results, as described by Henry et al. (2003) and Salmond et al. (2006). The best results are the most statistically relevant results. This iterative approach consists of running a test model with PCA, using the selected variables, interpreting the results, and revising the model on the basis of the obtained results – reapplying PCA with subsets of the variables, until the results meet specific performance requirements (Henry, 2003). Thus, the approach consists of multiple consecutive runs of PCA and revisions, testing different subsets of the variables. The objective is to meet the performance requirements and to complete the process with the ideal subset of variables.

To assess whether the performance requirements are met, the researcher has to first assess the Kaiser-Meyer-Olkin (KMO) statistic, which is a measure of adequacy of the sample, and to conduct the Bartlett's Test of Sphericity, which tests the null hypothesis that the correlation matrix is an identity matrix (CGAP, 2000; Mooi and Sarstedt, 2011). These assessments are needed to ensure that the data are suitable for PCA. The KMO statistic is also known as the measure of sampling adequacy (MSA) and it indicates if the correlations between variables can be explained by the other variables in the dataset (Mooi and Sarstedt, 2011). The values of the KMO statistic range from 0 to 1, and the researcher should ideally have a KMO statistic close to one. Only KMO statistics above 0.50 are acceptable, although only values greater than 0.70 are considered to be adequate (Kaiser, 1974). Regarding the Bartlett’s Test of

Sphericity, it has to be rejected, because the PCA needs relatively high correlations between the variables (Mooi and Sarstedt, 2011).

Besides observing the KMO statistic and the Bartlett's Test of Sphericity, the researcher also has to observe the communalities, which represent the strength of the linear association between the variables and the component. This way, he/she is able to assess how much variance of each variable can be reproduced through factor extraction, i.e., how much of each variable's variance is captured or reproduced by the extracted factor. As it is the case with the KMO statistic, the values of the communalities range between 0 and 1. The greater their value, the greater the share of common variance explained by the extracted component (CGAP, 2000). Generally, at least 30% of a variable's variance should be accounted for through the extracted factors, which implies that the communalities should lie above 0.30. If the communalities are close to zero (for example, less than 0.10), then the variable in question is a candidate for exclusion in subsequent runs of the iterative approach (Mooi and Sarstedt, 2011).

Finally, the researcher has to observe the resulting factors' eigenvalues, which describe how much variance is accounted for by each factor, as a result of the PCA. More specifically, the researcher has to select a subset of variables that produces only one factor with an eigenvalue greater than one, following the rule known as Kaiser criterion, in order to define the number of components to retain for the construction of the index. Furthermore, the retained first principal component has to account for at least 50% of the total variance explained (Mooi and Sarstedt, 2011).

The main requirements for the construction of an index with the first component of PCA are summarized in Table 1, which describes the two process phases: (i) the selection of the initial set of 10 to 20 variables, and (ii) the iterative, trial and error approach for the selection of the relevant subset of variables: the ones that will produce the desired factor's scores.

Table 1 – Constructing an index with the first component of PCA

Process phase	Requirements
Selection of the initial set of variables	- Selection of variables that represent the desired factor; - Correlations among the variables > 0.3, but < 0.9; - 10 to 20 variables; - Sample size of at least ten times the number of variables.
Iterative approach to define the relevant subset of variables	- KMO statistic: > 0.5 is acceptable and > 0.7 is adequate; - Rejection of the Bartlett's Test of Sphericity; - Communalities > 0.3; - One factor with an eigenvalue > 1; - First component accounts for at least 50% of the total variance.

After defining the relevant subset of variables, the PCA can be applied to produce the scores of the index, for each of the geographic assessment units.

3.1.1.2. The health needs index

We propose to construct a health needs index with demographic and socioeconomic variables that are known to be associated with the health needs of the populations. As seen in subsections 2.1.1 and 2.1.2, demographic and socioeconomic variables that describe populations' characteristics can be used to measure health needs (perceived needs), and the non-spatial dimensions of access. This way, the incorporation of the health needs index into our method makes it possible to capture non-spatial access attributes in the assessment of access, such that the computed scores constitute comprehensive measures of the populations' access to health care – not just spatial accessibility.

The inclusion of a health needs index in our EKD2SFCA method is also appropriate, because the higher the health needs of a population, the greater is its demand for health relative to its size, and the greater the demand, the worse is access, because supply will become insufficient. This means that an unadjusted supply-to-demand ratio will indicate a higher access level than it should for populations with high health needs. Conversely, for populations with low health needs, an unadjusted ratio will indicate a lower access level than it should, because the ratio will not take into account that the population with low needs will demand less of the existing supply than the population size suggests. Therefore, when there are regional differences in the populations' health needs, the denominator of the supply-to-demand ratios that characterize 2SFCA-based methods should be adjusted to capture these regional differences. This can be achieved with the incorporation of a health needs index into the model, in order to correct the populations' sizes for their health needs and to thereby adjust the supply-to-demand ratios. In the geographic assessment units where their residents have high needs, their sizes, measured in terms of number of residents, should be increased by the index. In the units where the residents have low health needs, their sizes should be reduced by the index.

Many studies in the scientific literature identify demographic and socioeconomic variables that reflect populations' characteristics which relate to their health needs. Hence, such variables can be used as proxies for health needs. Alternatively, the researcher can promote a survey in order to obtain information from the populations on their health needs or conduct a study to identify representative variables. For instance, he/she can choose a health needs benchmark indicator and then identify a group of variables that is highly correlated with that benchmark indicator. However, these alternatives can be costly and time consuming, and choosing a benchmark indicator for a new study can be a debatable endeavor, because there is no consensus on such an indicator.

The following examples of demographic and socioeconomic variables associated with health needs can be identified in the literature:

- People with ages above 65 and ages 0-4, women with ages 15-44, population in poverty, home ownership, median income, and housing units lack of basic amenities, among others, as in the study of F. Wang and Luo (2005);

- Gender, unemployment (economically active people above age 16 who are also unemployed), single-parent households, and households where there is an average of more than one person per room, among others, as in the study Field (2000); and
- Employment, income, qualification, communication and housing, among others, like the ones chosen in the study of Bagheri et al. (2008).

After the researcher chooses an appropriate set of variables for the health needs index, and adopts the iterative approach to construct an index with the first component using PCA, the index produces standardized scores, i.e., scores with mean equal to zero and standard deviation equal to one. The scores are computed with a regression method, for example, using a statistical software package, such as SPSS, which is the software package used for the applications of our method in this thesis.

The incorporation of the health needs index into the model is described by equation (14), which is the first step of our proposed EKD2SFCA method:

$$R_j^E = \frac{S_j}{\sum_{k \in \{d_{kj} \le d_{\max}\}} P_k H_k g(d_{kj}, d_{\max})}; \qquad (14)$$

where R_j^E is the supply-to-demand ratio of the method for the supply point in j, S_j is the number of supply resources available at the supply point in j (number of physicians), P_k is the population that resides in the geographic unit geo-referenced at k, d_{kj} is the travel time between k and j, $d_{\max}$ is the maximum time that populations will travel to reach a supply point, H_k is the health needs index, representing the health needs of the population at k, and $g(d_{kj}, d_{\max})$ is the distance decay function, which we will specify in subsection 3.1.2.

As can be seen, the health needs index is multiplied by the population in the denominator in equation (14). In order to apply the standardized health needs scores in the equation, it is necessary to rescale them. To this effect, we have determined that, in accordance with the objective of the populations' size correction, the average health needs has to become equal to one (no adjustment of the supply-to-demand ratio), lower health needs have to take values below one, and higher needs have to be represented by values above one (Polzin et al., 2014a). This is a similar approach as the one adopted by McGrail and Humphreys (2009a) to incorporate a health needs index into their improved 2SFCA-based method, which was created to assess access to primary health care in rural areas.

To define appropriate scale ranges for the scores, we considered the definition of Health Professionals Shortage Areas (HPSA) from the U.S. Department of Health and Human Services, which was used by F. Wang and Luo (2005) in their application of the 2SFCA method (see subsection 2.1.3.5). HPSA are geographic areas with physician-to-population ratios of 1/3500 or less, but also with ratios close to 1/3000, if the population has unusually high needs. Accordingly, we defined the variation between these two ratios' values as the reference to be used for the adjustment of the population size, in order to account for health needs.

This way, we considered that populations with unusually high health needs should have their associated supply-to-demand ratios reduced by approximately 14.3 percentage points, i.e., the variation from 1/3000 to 1/3500. This is equivalent to an increase in the population size in equation (14) by about 16.67%, for a given number of physicians. In accordance therewith, defining unusually high needs as one standard deviation above the average health needs score, as proposed by F. Wang and Luo (2005), the positive standardized scores, between zero and one, have to be converted to the scale from one to 1.167, in a proportionate way. Also, considering a linear extrapolation of the scale, scores higher than one have to be transformed to values above 1.167.

Equally, for the negative scores, we considered that a supply-to-demand ratio for populations with very low needs, defined as one standard deviation below the mean, also has to vary 14.3 percentage points, which is the assumed necessary adjustment for health needs. This means that the population size has to decrease 12.4%, for a given number of physicians. Accordingly, the scores with one standard deviation below the average value have to be equal to 0.876, the negative scores between zero and minus 1 have to be converted to a scale from one to 0.876, in a proportionate way, and scores lower than minus 1 have to be rescaled to linearly defined values below 0.876.

Figure 6 illustrates the scale transformation that we propose.

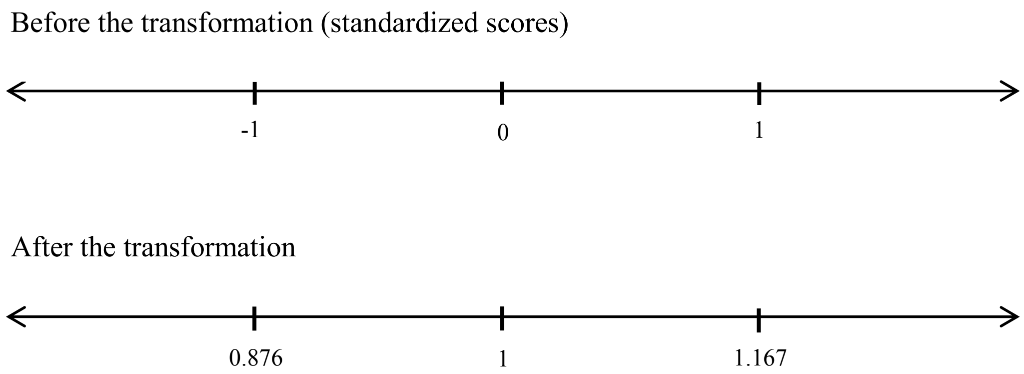

Figure 6 – Scale transformation of the health needs scores

The application of this scale transformation to the standardized health needs scores enables the researcher to properly incorporate the health needs index into the model, increasing the population size, when the health needs are high, and decreasing its size, when the health needs are low.

3.1.1.3. The commuting index

As we indicated in subsection 2.1.3.5, one enhancement to the 2SFCA method is the adoption of varying catchment areas' sizes, in order to take into account Stouffer's concept of intervening opportunities. The enhanced 2SFCA method that first introduced varying catchment areas' sizes, created by McGrail and Humphreys (2009a), uses a capping function to reflect differing mobility patterns of populations that reside in different urbanized settings (see subsection 2.1.3.5, page 42). Thus, with the adopted capping function, the catchment areas will be smaller in urban settings, because, generally, urban populations will travel less than more rural populations to reach supply points and obtain the products or services that they need. This occurs because there are more supply points, or "intervening opportunities", near urban residences. There is a higher concentration of available supply points in more urban regions.

However, as we noted in the presentation of the 2SFCA method, there are orientations that are used to define the maximum travel times that determine the boundaries of the catchment areas, such as, for example, 90 minutes for general surgeries. These are fixed orientations for the determination of catchment areas sizes and they depend on the products or services that are considered in the assessment. The subjacent reasoning for these orientations is that they are the maximum travel times that the consumers are willing to travel to obtain the products or services.

Hence, in order to maintain the original 2SFCA method's adoption of these orientations, and their subjacent reasoning, we propose to take into account Stouffer's concept using commuting variables, as representative variables of the mobility of the populations when accessing health care. Commuting variables are variables that characterize the regular movements of workers and students between their residences and their work and study locations. This way, we do not change the catchment areas sizes.

There are many regional differences regarding commuting, which give an indication that the mobility of the populations may vary substantially across regions. In fact, it is recognized, for example, that residents in urban areas usually commute less than in suburbs, i.e., less urbanized peripheral cities, where their residents traditionally travel longer distances to reach their work or study locations. Also, in many rural areas, located too distantly from larger cities and metropolitan areas, their residents do not commute much. They depend more on their neighborhoods and local economies. These commuting patterns reflect Stouffer's concept of intervening opportunities, considering intervening opportunities, in this case, as work and study locations.

Nevertheless, by definition, work and study locations are situated in activity spaces, which are the geographic areas that comprise the locations regularly visited in the individuals' normal daily activities, and paths to and from these locations (Cromley and Shannon, 1986). In turn, the activity spaces typically concentrate different kinds of supply points, like shopping centers and restaurants, and also health care supply points. Hence, it can be established that the commuting mobility patterns can matter for the assessment of patients' access to health care. In fact, it is recognized that people are more likely to use health services conveniently located in their activity spaces (McLafferty, 2003). Accordingly, people will generally travel smaller distances to obtain health care when activity spaces are small geographic areas, and longer distances when these spaces are large geographic areas.

In order to capture these mobility patterns in our EKD2SFCA method, we specifically propose to use commuting variables summarized in an index, in order to adjust the access scores of our method. The same way as with the health needs index, we use PCA to construct the commuting index with the first component. Thus, the same process summarized in Table 1 applies to the construction of our commuting index, such that the first construction phase refers to the selection of the initial set of variables.

The selection of variables for the initial data set needed for the PCA may include any of the commuting variables with data that are systematically provided by national institutes of statistics. Considering the Portuguese National Institute of Statistics, examples of variables are:

- Proportion of population that works or studies in another municipality;
- Population that works or studies in another municipality as a proportion of total employed or student population;
- Employed or student population with commuting of over 30, 60 or 90 minutes as a proportion of active population;
- Employed or student population that uses primarily bus, tram, metropolitan, train or car (as driver or passenger) in commuting as a proportion of active population;
- Employed or student population that uses primarily car (as driver or passenger) in commuting as a proportion of active population; and

– Average travel time in commuting (minutes) of the employed or student population.

After constructing the commuting index, the index is integrated into the EKD2SFCA method in its step 2, such that the access scores are adjusted for the mobility of the population of each geographic assessment unit, as approximated by the commuting index.

The integration of the index is shown in equation (15), which constitutes the step 2 of our proposed EKD2SFCA method:

$$A_i^E = \sum_{l \in \{d_{il} \leq d_{\max}\}} R_l^E C_i g(d_{il}, d_{\max}); \qquad (15)$$

where A_i^E is the access score computed by the method for the geographic unit i, $\sum_{l \in \{d_{il} \leq d_{\max}\}} R_l^E$ is the sum of the step 1 supply-to-demand ratios associated to the supply points that are within reach of the population at i, C_i refers to the commuting index, representing the mobility of the population at i when accessing health care, and $g(d_{il}, d_{\max})$ refers to the same distance decay function used in step 1 (equation (14)).

Before integrating the standardized scores of the commuting index computed with PCA into equation (15), they have to be rescaled. In order to rescale the scores, we propose that the same scale used by McGrail and Humphreys (2009a) for their mobility index should be adopted. As remarked by the authors, empirical evidence suggests that mobility can affect 30% of the population as a significant barrier (McGrail and Humphreys, 2009a). Accordingly, we established that the negative commuting standardized scores computed with PCA have to be transformed to the scale from 0.7 to one and, regarding the positive scores, we defined that they have to be equal to one, in a similar way as it is done with the mobility index of McGrail and Humphreys (2009a). Figure 7 illustrates the scale transformation that we propose.

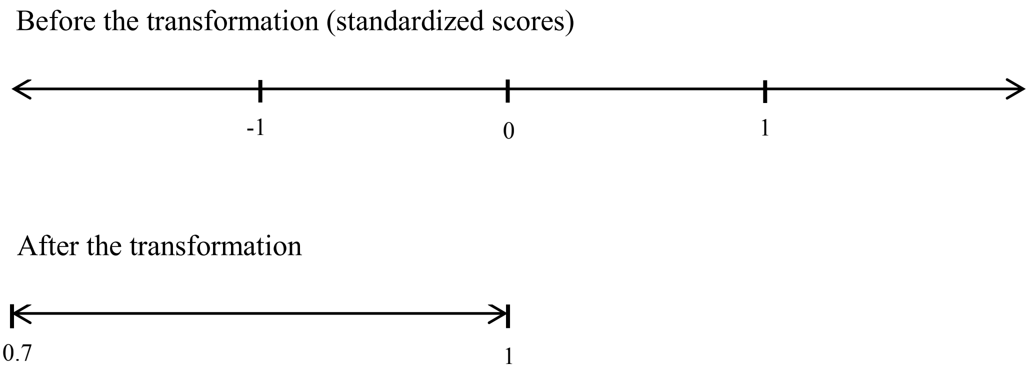

Figure 7 – Scale transformation of the commuting scores

This way, the commuting scores used by our extended method diminish overestimations of the patients' access to health care that would occur if the mobility of the populations when accessing health care was not taken into account. This correction considers that regions with relatively less significant commuting, measured by a set of commuting variables, are smaller activity spaces, where their residents disregard more the supply points that could be accessed in more distant geographic areas. In other words, the correction considers that the residents in areas with less commuting are less mobile residents, who rely more on their local supply and less on the available supply that could be additionally reached at longer travel times, yet still until acceptable maximum travel times. This way, these residents have, in the end, a lower access to the available health care supply than they would have otherwise.

Having completed the explanation regarding the health needs and the commuting indices, it is important to explain the change of the distance decay function that we promoted for our EKD2SFCA method.

3.1.2. Change of the distance decay function

Our proposed EKD2SFCA method is based on the KD2SFCA method of Dai and Wang (2011), which was explained in subsection 2.1.3.5. However, the KD2SFCA method was created to assess access to food stores. Its distance decay function is a default function of the geographic information system application that the authors used (ArcGIS). Since our focus is the health care industry, we sensed the need to change the distance decay function in our method, to one that would adequately represent the patients' travel behavior when traveling to health care supply points. Different functions model decaying proximity in diverse ways, as Dai and Wang (2011) recognize, and the best function should take into account the consumers' perception of travel impedance. Accordingly, each industry will have its ideal function, since consumers in diverse industries will have different preferences and behaviors regarding proximity, such that they will ponder the varying proximity within catchment areas differently.

The distance decay function in our EKD2SFCA method rescales populations according to their travel time to each supply point, and rescales the step 1 supply-to-demand ratios according to the travel time of each population to its corresponding available supply points (see equations (14) and (15) in subsections 3.1.1.2 and 3.1.1.3, respectively). Longer travel times are less valued than shorter travel times, receiving thus lower weights in the calculations. Many functional forms could be adopted, like, for example, inverse power, linear, Gaussian, logistic, exponential and quartic. Each form will produce a specific valuation or weighing of proximity, which may be adequate to represent the consumers' perception of distance decay in a specific market and corresponding industry, and, accordingly, the consumers' general behavior when traveling to supply points.

There are two main ways to identify the most appropriate functional form for the health care context. Firstly, a survey could be applied to patients in order to identify their perception of travel impedance, for example. In this case, patients would indicate their most preferred travel times until reaching the supply points, or their least preferred ones, due to high perceived travel costs and time. This way, the results of the survey would be used to identify weights that would represent the varying proximity inside the catchment areas. The researcher could then choose a functional form that would fit best the weighing of proximity.

An alternative way for choosing an adequate functional form, namely a form that properly represents the patients' behavior when traveling to health care supply points, can be the analysis of travel times data of the patients that were treated in the supply points of the study region during a specific time period. However, this kind of analysis may be biased, if patients are not completely free to choose the supply points, and if the supply is scarce and not evenly distributed across the region. In these cases, the data would not perfectly reflect the perception and preferences of the patients regarding travel times. For example, if there is only one supply point in the study region, all patients will travel to that supply point, even those patients that reside extremely far from it, because they have no choice. In addition, another drawback of this type of analysis is the fact that it depends on utilization data, which we already identified as a problem, due to its virtual unavailability.

Yet, an alternative to these potentially costly or time consuming studies is to identify in the scientific literature previous studies' results or consensual expert opinions that point to a specific choice of functional form for a determinate market or industry.

Following this third alternative for the choice of an adequate distance decay function for the health care industry, due to its simplicity, we were able to identify in the literature that a typical modeling of decaying proximity to health care makes use of the widely adopted quartic function (Cooper et al., 2009a; McLafferty and Grady, 2004). The quartic function can be thus considered to better reflect the behavior of patients regarding the distances and associated travel times to health care utilization. To this extent, it better represents the effect of proximity barriers on patients' access to health care.

Hence, for the proposed extended method we chose the quartic function, instead of the Epanechnikov function used by Dai and Wang (2011).

Figure 8 compares the quartic function with the Epanechnikov function of the original KD2SFCA method until the 90 minutes travel time limit, which is a reference for maximum travel times to surgical interventions.

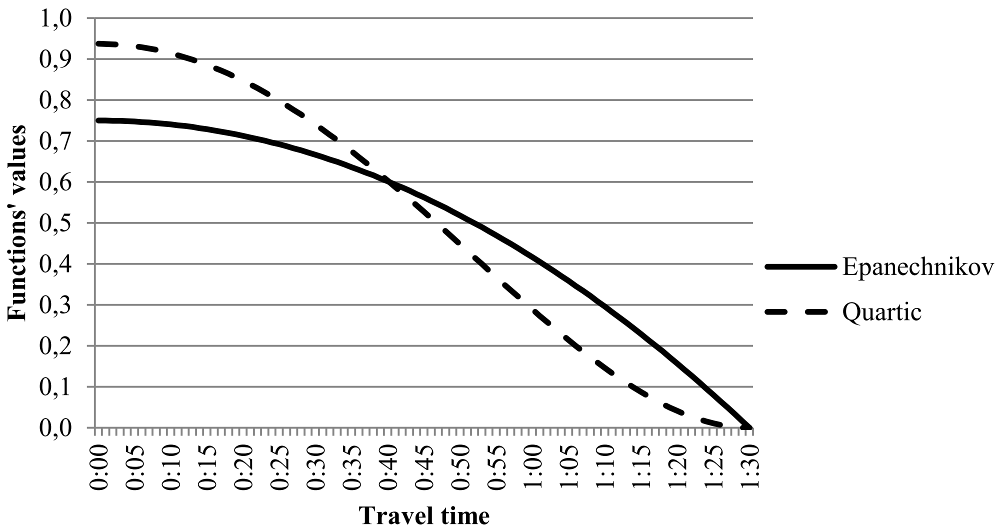

Figure 8 – Comparison between the quartic and the Epanechnikov functions up to the 90 minutes travel time limit

The functions' values in the figure are the weights or penalizations applied in steps 1 and 2 of the KD2SFCA method and, in the case of the quartic function, our EKD2SFCA method. As can be seen, the quartic function penalizes the shorter distances less and the longer distances more than the Epanechnikov function, and the penalization in smaller distances grows at higher rates as the travel time increases, while in longer distances the penalization augments at lower rates. These different weights of the quartic function can be considered to represent patients' behavior in a more accurate way when choosing and traveling to health care supply points.

The quartic function chosen for the EKD2SFCA method is the following:

$$g(d_{ij}, d_{\max}) = \begin{cases} 1 & ,if \quad 0 < d_{ij} < d_{init} \\ \frac{15}{16}\left[1 - \left(\frac{d_{ij}}{d_{\max}}\right)^2\right]^2 & ,if \quad d_{init} \leq d_{ij} \leq d_{\max} \\ 0 & ,if \quad d_{ij} > d_{\max} \end{cases} \quad ; \qquad (16)$$

where $g(d_{ij},d_{\max})$ is the quartic function (Kayri and Zırhlıoğlu, 2009; Tremblay and Merritt, 1995; H.-C. Yang et al., 2008), d_{init} is an initial catchment and $d_{\max}$ is the travel time that defines the boundaries of the catchment area. This function is thus the one that is included in equations (14) and (15), which respectively describe the steps 1 and 2 of our EKD2SFCA method. It is a continuous function that is applied in a discrete way, rescaling populations in step 1 according to their travel time to each supply point, and rescaling the step 1 ratios in step 2 according to the travel time of each population to its corresponding available supply points.

The initial catchment d_{init} is introduced as proposed by McGrail and Humphreys (2009a, 2009b, 2009c) and Schuurman et al. (2010), assuming that travel times below a determinate threshold do not present any sensible proximity impediment to utilization. This means that patients do not differentiate travel times until a specific threshold, after which the distance decay function applies. The choice of this initial catchment depends on the specific application of the method. We note that in the case studies presented in this thesis the threshold of 10 minutes is applied, assuming that in Portugal travel times below this threshold do not present any evident proximity barrier to health care utilization. However, we recognize that in other applications of the method, different definitions of initial catchment could be better fitted (Polzin et al., 2014a).

3.1.3. Assessment of the access scores of the EKD2SFCA method

As explained above, the EKD2SFCA method forms the first two steps of the EKD4SFCA method, and equations (14) and (15) present the formulas of these two steps, showing how the health needs index, the commuting index and the distance decay function are incorporated. The EKD2SFCA method enables the computation of patients' access scores for small geographic units of a study region, making it possible for the researcher to identify differences in access across the geographic units. The resultant supply-to-demand scores assimilate information from spatial and non-spatial access dimensions, and this makes the EKD2SFCA method a comprehensive method to assess access, complementing the spatial access assessment with relevant non-spatial information. This method also considers travel impedance, because of the distance decay function that differentiates travel times inside the catchment areas, and it takes into account Stouffer's concept of intervening opportunities. Hence, it considers all three main enhancements that were introduced by different 2SFCA-based methods to the original 2SFCA method, since its conception.

The access scores that are computed by the EKD2SFCA method can be used in different ways for the researcher to conclude on the access levels of the populations that reside in the study region and to identify shortage areas, where the disadvantaged populations reside. One way of carrying out this identification is by comparing the scores computed by the method to a benchmark value, such as the average value from a group of countries, quantiles, or average values of the scores of the whole study region. Also, a value determined by experts can be used as a reference. An example is the definition of health professionals' shortage areas mentioned before, as areas with physician-to-population ratios of 1/3500 or less. Geographic areas with access scores of below 1/3500 would, in this case, represent shortage areas.

Another way of assessing the access scores is by applying cluster analysis, for example, as noted before, in subsection 2.1.3.4. With cluster analysis, the small-area results are aggregated and expressed in access levels, like, for instance, low, medium and high. An advantage of cluster analysis is that the grouping of the access scores is determined by the data themselves, such that the level partitions are not arbitrarily determined by any researcher's choice.

Following this alternative, we propose to apply a k-means clustering algorithm to the access scores. This algorithm finds a partition such that the squared error between the empirical mean of a cluster and the points in the cluster is minimized (Jain, 2010). More specifically, the algorithm minimizes the sum of the squared error over the K clusters, as represented by the following formula:

$$J(C) = \sum_{k=1}^{K} \sum_{x_i \in c_k} \left\| x_i - \mu_k \right\|^2 ; \qquad (17)$$

where C refers to the set of clusters; k refers to the identification of a specific cluster (c_k); K is the total number of clusters; x_i is one observation of the data set (the set of access scores); and μ_k is the mean of the cluster c_k.

By means of the k-means algorithm, the access scores can be grouped in three access levels – high, medium and low –, and this classification can be useful for researchers to identify the shortage areas, namely the low access areas, where the disadvantaged populations reside.

Now, having explained the proposed EKD2SFCA method, the access scores it computes and how they can be used for the assessment of access to health care, we present next a case study, with an example of application of the method.

3.1.4. An application of the EKD2SFCA method

As an example of an application of the EKD2SFCA method we present a case study with an assessment of patients' access to hospital health care in continental Portugal. This example compares the EKD2SFCA method with the original KD2SFCA method of Dai and Wang (2011). The objectives of this comparison are: (i) to further explain the first two steps of the EKD4SFCA method, (ii) to assess how the results change after introducing our adaptations and extensions, and (iii) to show that the EKD2SFCA method produces improved results to identify disadvantaged populations, namely the populations with low access to hospital health care.

As explained in subsection 2.1.3.4, the application of 2SFCA-based methods first requires the choice of the specific study region and of the small geographic units of that region that will be used for the assessment. The study region is the whole region of interest for the researcher, namely the region for which the researcher wishes to identify differences in access among populations. In this case study, the study region is continental Portugal. The geographic assessment units are the 460 four-digit postcode areas of continental Portugal, which are presented in Figure 9.

We chose postcode areas as geographic assessment units, because we consider these areas to adequately reflect the distribution of the populations across the Portuguese territory. Besides, they are sufficiently small for the application of our extended method to the Portuguese hospital health care context, in order to identify with a fine resolution and enough detail the disadvantaged, less empowered populations to use hospital health care. In addition, they have been used extensively in health research (Luo and Wang, 2003).

These areas were identified and illustrated by Microsoft MapPoint 2009, which is the GIS that was adopted. This software application supplied the speed maps, the traffic rules, the postcode areas, as subdivisions of the study region, and the centroids of the postcode areas. This GIS has a simple interface, and enables an easy integration with Microsoft Excel, with Visual Basic for Application (VBA) codes.

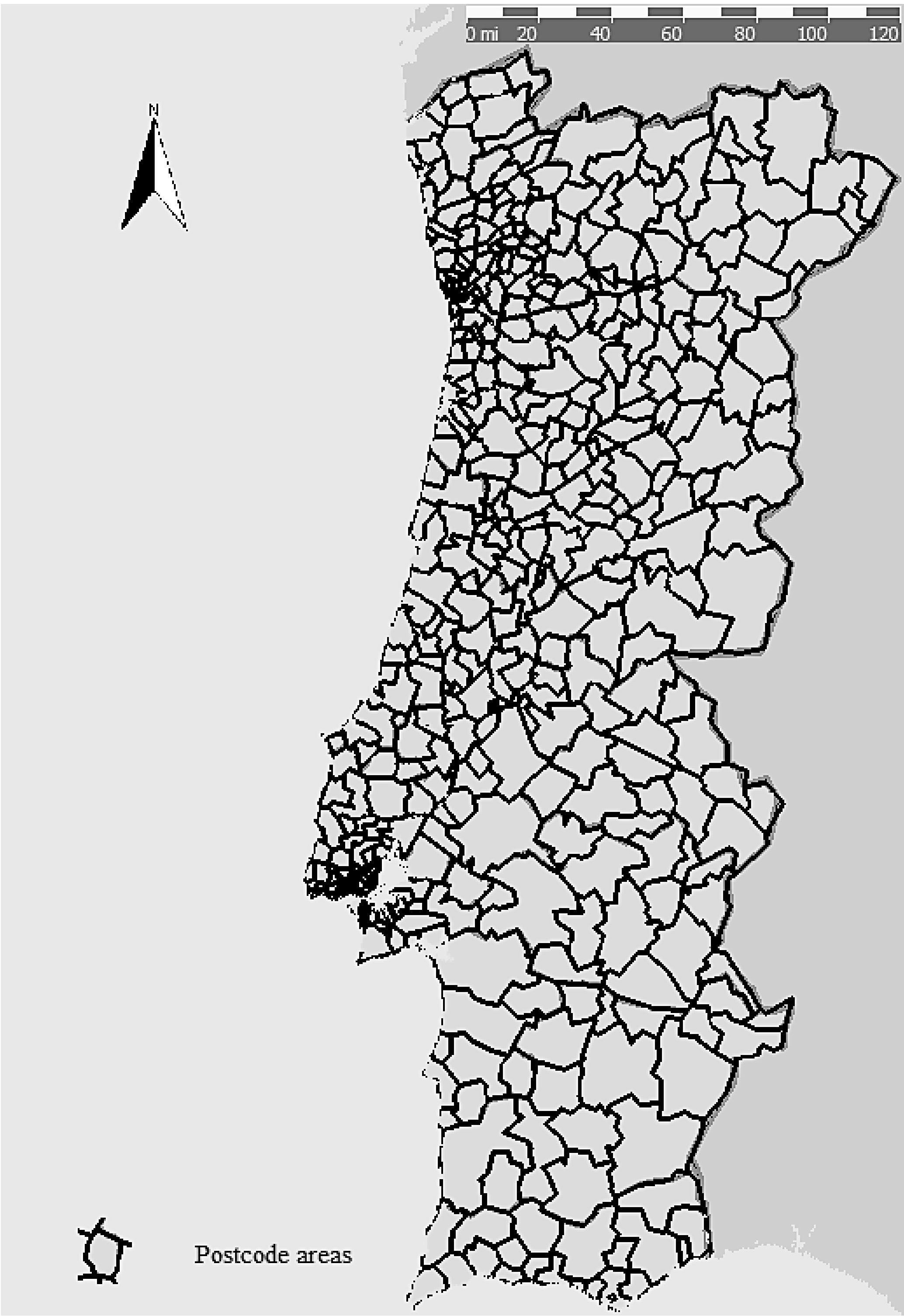

Figure 9 – The 460 postcode areas of continental Portugal

A specific VBA code (see Appendix A) was used to calculate all travel times from each of the 460 postcode areas until each of the 459 other postcode areas (211,140 calculations). This way, the travel time data needed for the calculations of steps 1 and 2 of the KD2SFCA method and of the EKD2SFCA method were exported to Excel, and all the calculations could be carried out in simple Excel sheets. The calculated travel time data with MapPoint resulted from the consideration of minimum travel times by car on motorways, roads and streets with average speeds. The centroids of the postcode areas were the points of departure and arrival used for the calculations, such that only the centroids were used for the geo-referentiation of the supply points and the populations (accordingly, the travel time of a population to a supply point in the same postcode area was considered to be equal to zero, which, however, for the sake of the calculations of the EKD2SFCA method, is equivalent to any other travel time until the initial catchment of 10 minutes – see subsection 3.1.2, page 83).

Apart from the definitions of the study region and the small geographic assessment units, the application of 2SFCA-based methods needs also the definition of the boundaries of the catchment areas, i.e., the catchment areas sizes. As explained before, catchment areas are geographic areas that reflect the consumers' preferences and their buying patterns. The catchment areas sizes are defined by the maximum travel times that consumers are willing to travel to reach the products and services that they need or desire. Known orientations for the definition of the maximum travel times are 30 minutes for primary care, urgent health services or general care for adults and children, 45 minutes for obstetrical services or radiotherapy, and 90 minutes for general surgeries. For instance, a catchment area for a primary care supply point can be set to cover all the populations that can reach that supply point by traveling 30 minutes or less by car.

Accordingly, to define the catchment areas we chose the 90 minutes travel time reference for general surgeries, since they are typically carried out in hospitals and the objective in the case study was to assess patients' access to hospital health care. Therefore, only the postcode areas with centroids that are located within the 90 minutes catchment were considered in each catchment area with center at a specific centroid. The only exceptions were the cases in which the catchment areas of public and publicly financed hospitals are officially defined, with travel times of less than 90 minutes. The official catchment areas defined by municipalities were obtained from the Portuguese Health Regulatory Agency and were converted to their respective postcode areas. As we could confirm, some of these official areas have boundaries of less than the maximum travel time of 90 minutes (in the cases in which the boundaries surpassed the 90 minutes reference, we cut them, in order to respect this limit).

The supply data used in the case study refer to December 2011. These data consist of the number of physicians and the locations of the existing 168 hospitals of continental Portugal (79 public, 25 private not-for-profits, of which 15 publicly financed, and 64 private for-profits). They were collected from the Portuguese Health Regulatory Agency registry of health services providers (Appendix B presents the list of the 168 hospitals).

Figure 10 presents continental Portugal, and the location of the 168 hospitals, considering the centroids of the postcode areas. The circles represent the number of hospitals that are located in each postcode area. The sizes of the circles vary according to the number of hospitals. The highest concentration of hospitals in one postcode area is of six facilities, and it occurs only once, in the north of the country.

The demographic data were collected from the Portuguese National Institute of Statistics, as well as the socioeconomic data gathered for the health needs and commuting indices. The data were Census data for the Portuguese localities (*freguesias*), which we summed up to the larger postcode areas level.

3. The Extended Kernel Density 4-Step Floating Catchment Area (EKD4SFCA) method
3.1. The EKD2SFCA method (steps 1 and 2 of the EKD4SFCA method)

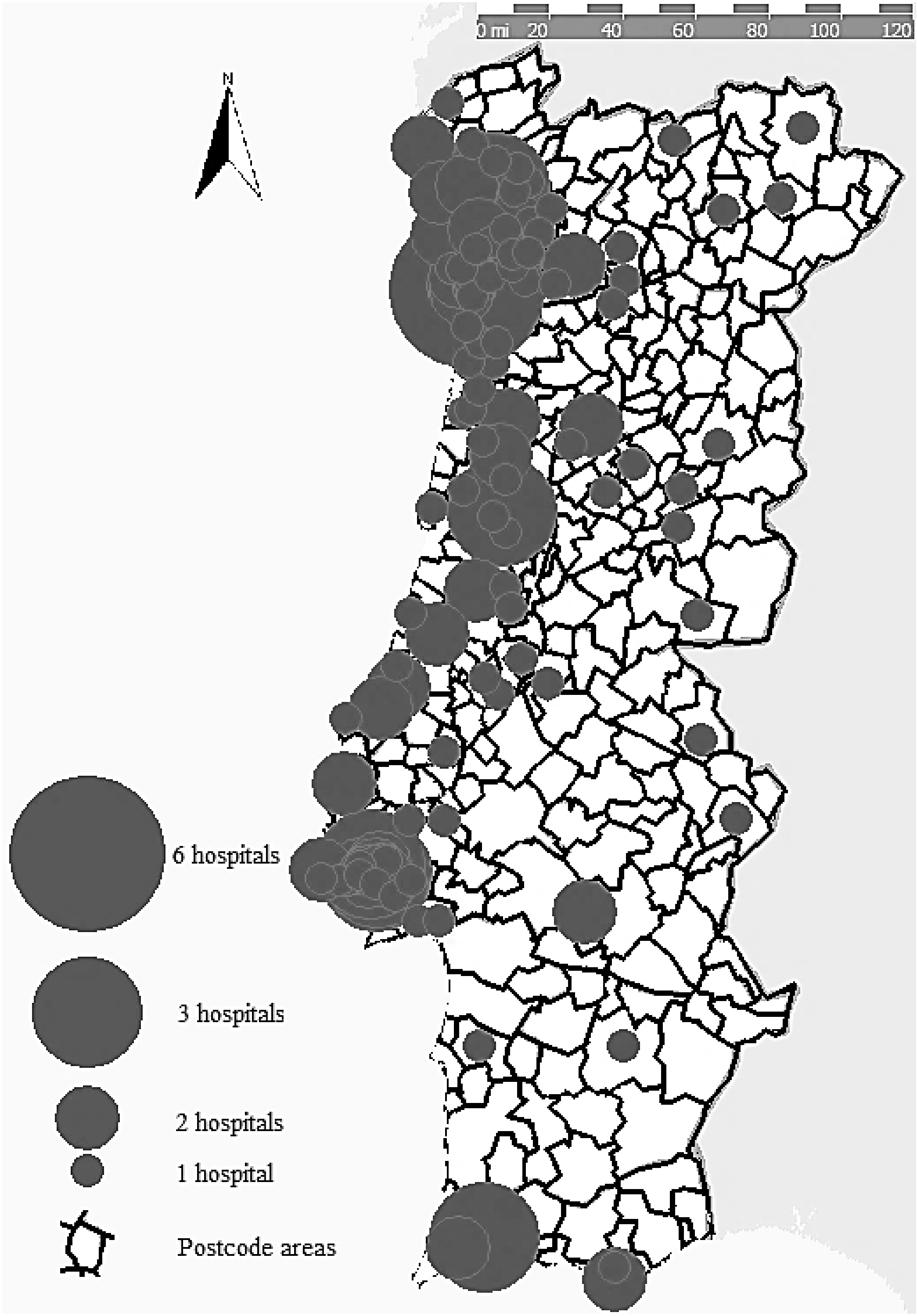

Figure 10 – The geographic distribution of the 168 hospitals

During the process of data collection, we note that we were able to confirm that the observation that Portugal lacks good information on hospital supply and populations' needs stated by Oliveira and Bevan (2003) remains valid. As far as the supply is concerned, the confirmation of the number of physicians and the collection of information regarding official catchment areas were hindered by the existing lack of easily available official information. Regarding not only the populations' health needs index, but also the commuting index, data from the previous 2001 Census had to be used, since the complete data of the 2011 Census were still not available at the time of the beginning of the study, in 2012 (Polzin et al., 2014a).

Next, before presenting and assessing the results of the application of the EKD2SFCA method, as compared to the results of the original KD2SFCA method, we detail how the health needs and commuting scores were computed.

3.1.4.1. Computation of the health needs scores

Following the instructions for the construction of the health needs index of the EKD2SFCA method with PCA (see subsections 3.1.1.1 and 3.1.1.2), we constructed the index H_k by first identifying an initial set of demographic and socioeconomic variables that are known to represent populations' health needs (Field, 2000; Field and Briggs, 2001).

We first chose a set of 14 variables to construct the health needs index for our case study, namely the following:

- Proportion of population aged 0 to 14 and 65 and higher, and women from 15 to 44;
- Proportion of population aged 0 to 14 and 65 and higher;
- Women aged 15 to 44;
- Population with a degree of disabilities (auditory, visual, motor, mental, cerebral paralysis, or other) of above 60%;
- Population with disabilities of 30% and above;

- Population with disabilities of above 80%;
- Unemployed as a proportion of total active population;
- Unemployed as a proportion of total population;
- Inactive population aged 15 and higher;
- Proportion of population that cannot read or write;
- Proportion of population aged 15 and higher that has not attended minimum obligatory education;
- Proportion of population that has not attended minimum obligatory education;
- Proportion of illiterate population over 14 years old or population that can read and write without any formal education; and
- Proportion of population without any education aged 15 and higher.

Thereafter, using SPSS software, we adopted the iterative approach described in subsection 3.1.1.1 to select, from the initial set the variables, the ones that were statistically more relevant. This way, we constructed the health needs index with the following four variables: age (proportion of population aged 0 to 14 and 65 and higher), recognized deficiency (population with disabilities of above 80%), economically inactive population (inactive population aged 15 and higher), and education (proportion of illiterate population over 14 years old or population that can read and write without any formal education).

The results obtained with the four selected indicators satisfied the necessary criteria to construct an index with PCA using the first component (Havard et al., 2008; Howe et al., 2008; Messer et al., 2006; Mooi and Sarstedt, 2011). In particular, we obtained one factor with an eigenvalue greater than one (2.741>1) with a high percentage of the total variance explained (68.535%>50%) and an adequate value for the KMO statistic (0.728>0.7) (see Table 1, in subsection 3.1.1.1, with the reference values).

Applying the scale transformations to the 460 health needs standardized scores, as detailed in subsection 3.1.1.2, we obtained final health needs scores in a scale with a minimum value of 0.76 until a maximum value of 1.53.

3.1.4.2. Computation of the commuting scores

The initial data set chosen for the commuting index considered 17 commuting variables collected from the Portuguese National Institute of Statistics, namely the following:

- Proportion of population that works or studies in another municipality;
- Employed or student population with commuting of over 30 minutes as a proportion of the active population;
- Employed or student population with commuting of over 60 minutes as a proportion of the active population;
- Employed or student population with commuting of over 90 minutes as a proportion of the active population;
- Population that works or studies in another municipality as a proportion of the total employed or student population;
- Employed or student population that uses primarily bus, tram, metropolitan, train or car (as driver or passenger) in commuting as a proportion of the active population;
- Employed or student population that uses primarily car (as driver or passenger) in commuting as a proportion of the active population;
- Employed or student population that uses primarily bus, tram, metropolitan or train in commuting as a proportion of the active population;
- Proportion of employed or student population that uses primarily bus, tram, metropolitan, train or car (as driver or passenger) in commuting;
- Proportion of employed or student population that uses primarily car (as driver or passenger) in commuting;
- Proportion of employed or student population that uses primarily bus, tram, metropolitan or train in commuting;

- Proportion of employed or student population with commuting of over 30 minutes;
- Proportion of employed or student population with commuting of over 60 minutes;
- Proportion of employed or student population with commuting of over 90 minutes;
- Proportion of employed or student population that uses primarily car (as driver) in commuting;
- Employed or student population that uses primarily car (as driver) in commuting as a proportion of the active population; and
- Average travel time in commuting (minutes) of the employed or student population.

After following the iterative approach described in subsection 3.1.1.1, four variables were selected to construct the commuting index of the case study: proportion of population that works or studies in another municipality, proportion of employed or student population that uses primarily bus, tram, metropolitan, train or car (as driver or passenger) in commuting, proportion of employed or student population with commuting of over 60 minutes, and average travel time in commuting (minutes) of the employed or student population. Regarding the proportion of population that works or studies in another municipality, since we used the centroids of the postcode areas as our references for departure or arrivals, we considered this variable as being, rather, the proportion of population from the postcode area that works or studies in postcode areas of other municipalities.

These four indicators produced PCA results that satisfied the criteria to construct an index with the first component, like, for instance, one factor with an eigenvalue greater than one (3.028>1) with a high percentage of the total variance explained (75.699%>50%). We also obtained an adequate value for the KMO statistic (0.734>0.7).

Thereafter, as explained in subsection 3.1.1.3, in order to rescale the 460 standardized scores, we adopted the scale from 0.7 to one, transforming the negative scores to the scale from 0.7 to one and setting the nonnegative scores equal to one.

3.1.4.3. Results and conclusions

We applied the original KD2SFCA method (equations (9) and (10) presented in subsection 2.1.3.5) and our EKD2SFCA method (equations (14) and (15) presented in subsections 3.1.1.2 and 3.1.1.3) to assess access to hospital health care in continental Portugal and compared their results in three ways, by:

(i) Analyzing plots of the supply ratios against the 45-degree line, in order to visualize the access scores differences between both methods;

(ii) Comparing the numbers of postcode areas of the three access levels (high, medium and low), the percentages of population residing in the three access levels regions, and the three access levels scores' means, in order to conclude on the differences across the levels; and

(iii) Representing the high, medium and low access levels in maps of continental Portugal, in order to visually detect the differences.

For the analyses (ii) and (iii), we adopted the k-means clustering algorithm to group access scores in the access levels, as indicated in subsection 3.1.3 (R software was used to apply k-means). In order to fix the access scores' bounds of each access level and enable the comparison between the results of the two methods, the k-means algorithm was applied only to the KD2SFCA method's results. We took logarithms of the scores to avoid the bias that would result due to the existing large public central hospitals and their effect on some of the access scores.

Beginning with analysis (i), we present the results of the EKD2SFCA method, compared to the KD2SFCA method, by plotting the line of equality (the 45-degree line) with the dispersion of their corresponding access scores, i.e., the numbers of physicians per 100,000 inhabitants, for the 460 geographic units. As it can be seen in Figure 11, our EKD2SFCA presents generally lower scores in geographic units with low KD2SFCA scores and higher scores in geographic units with high KD2SFCA scores.

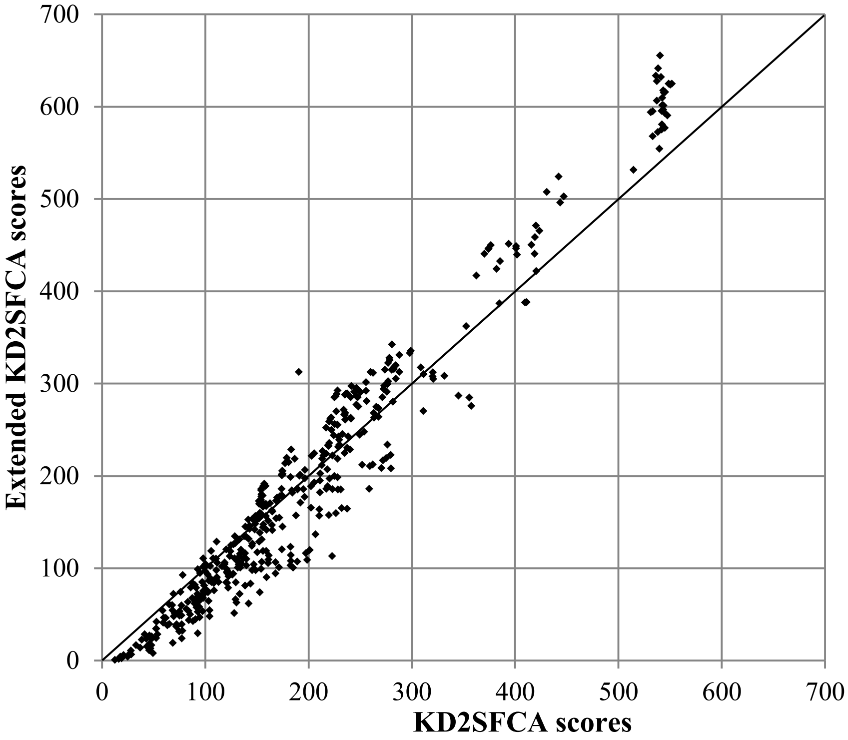

Figure 11 – Line of equality with the dispersion of the access scores

These scores differences are also noticeable in Table 2, which presents the results of analysis (ii). For each access cluster (high, medium and low) and each method's results, Table 2 indicates the number of postcode areas, the percentage of population covered and the access scores' mean.

Table 2 – Comparison between the KD2SFCA and the extended KD2SFCA results

Statistic	**Method**	**Cluster**		
		Low	**Medium**	**High**
Number of areas	KD2SFCA	57	218	185
	Extended KD2SFCA	95	187	178
Percentage of population covered	KD2SFCA	4.6%	40.8%	54.6%
	Extended KD2SFCA	8.1%	36.0%	55.9%
Mean access score	KD2SFCA	53.3	137.9	313.9
	Extended KD2SFCA	43.0	132.1	341.2

To convey an idea of the magnitude of the scores, we note that the average of physicians employed in hospitals per 100,000 inhabitants from the EU27 countries weighted by population is approximately 190, considering 2009 data from Eurostat (with the exception of Luxembourg, Slovakia, Sweden and the United Kingdom, due to lack of data). The two bounds that we fixed to define the three access levels are 78.2 and 198.7 physicians per 100,000 inhabitants. Thus, for access scores until 78.2 we identify a low access level, for access scores between 78.2 and 198.7 we have medium access, and for access scores of over 198.7 we have high access.

We conclude that, compared to the KD2SFCA method, our EKD2SFCA method identifies more areas and populations with low access, and with a lower average access score. This can be considered an advantage for researchers that focus their analysis on the identification of disadvantaged populations.

Finally, we present the results in maps (analysis (iii)). Figure 12 and Figure 13 present the spatial distribution of the postcode areas' scores grouped in the three access levels, for each of the two methods, respectively.

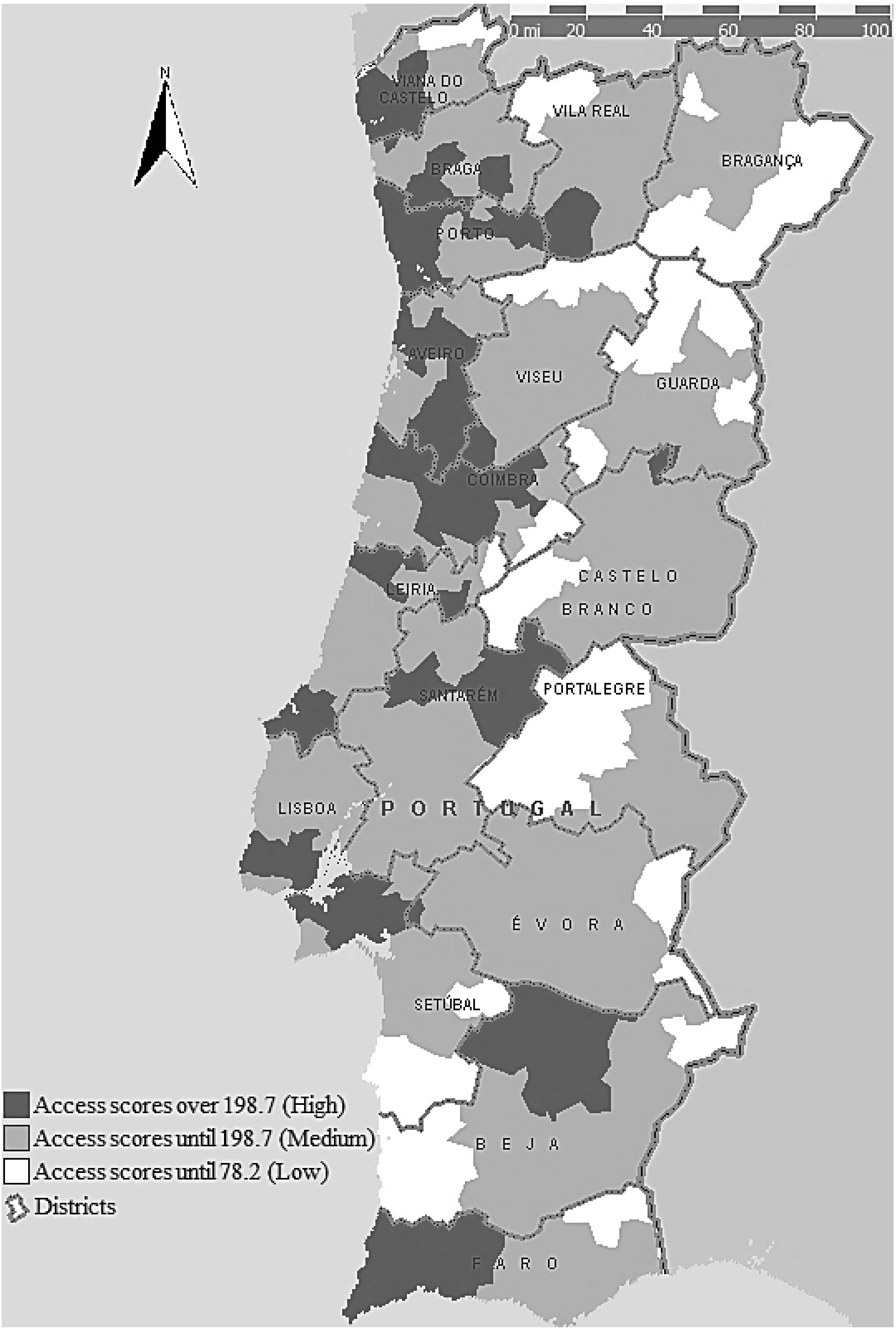

Figure 12 – Spatial distribution of the KD2SFCA high, medium and low access levels

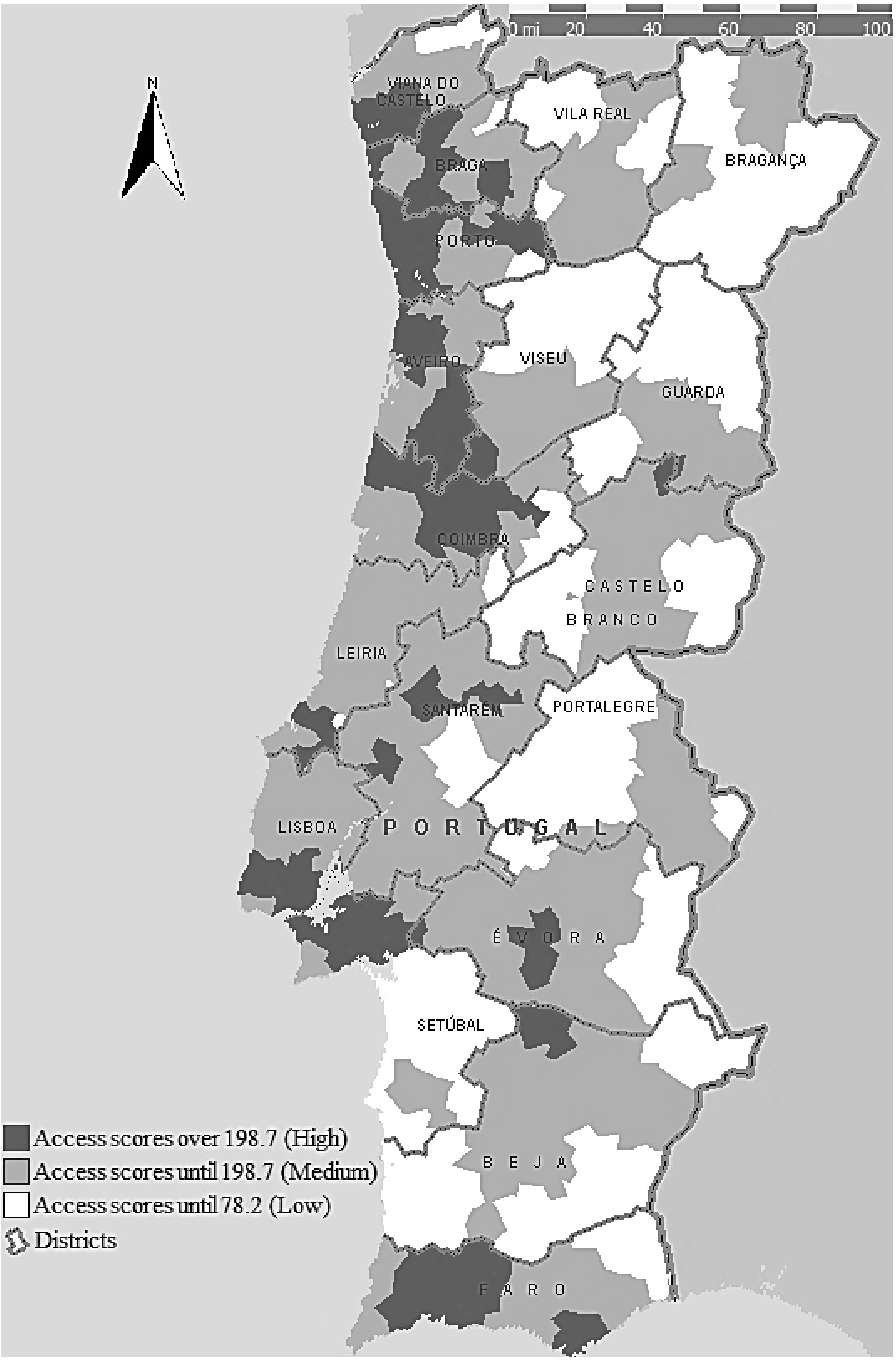

Figure 13 – Spatial distribution of the EKD2SFCA high, medium and low access levels

The results presented in both Figure 12 and Figure 13 give an indication of what is commonly acknowledged, namely that physicians are mostly concentrated in the coastal areas, where there is a higher concentration of central hospitals, and that Portugal has pronounced supply imbalances (Correia and Veiga, 2010; Oliveira, 2010; Oliveira and Bevan, 2003). This means that the two methods produce results that are apparently valid, reflecting largely what was expected and what the methods were supposed to measure.

Nevertheless, a more detailed comparison of the maps shows that our EKD2SFCA method changes some of the KD2SFCA method's results. These changes stem from the enhancements that we introduced into the model, namely: (i) the incorporation of non-spatial elements into the model, (ii) the consideration of Stouffer's concept of intervening opportunities, and (iii) the choice of an adequate distance decay function. Thus, we conclude that the changes are corrections of underestimations and overestimations. These corrections are summarized in Table 3, which presents the number of postcode areas that have their access level changed when we apply our method instead of the KD2SFCA method.

Table 3 – Access level changes after applying the extended KD2SFCA method

Access level changes		**Number of postcode areas**
Underestimations	Low to medium	1
	Medium to high	15
	Total	**16**
Overestimations	High to medium	22
	Medium to low	39
	Total	**61**

As regards the overestimations corrected by our method, and comparing again Figure 12 and Figure 13, we note that Figure 13 presents more low access areas, which can be found in 16 of the 18 districts, instead of in 13 districts, which encompass the low access areas in Figure 12. As can be seen, the EKD2SFCA method also identifies low access areas in the districts of Braga, Porto and Santarém, which do not have low access according to the original KD2SFCA method. This occurs because, as identified in Table 3, when our method is applied, 39 postcode areas with medium access, according to the KD2SFCA method, become low access areas, particularly in rural districts like, for example, Bragança and Beja.

In fact, the populations in some areas of these two districts are very distant from the regions where there is a high concentration of hospitals. Furthermore, they have high health needs and low mobility, as indicated by our indices. For example, we can find that postcode area 5320 in Bragança presents proportions of 40.2% of population aged 0 to 14 and 65 and higher, of 66.1% of inactive population aged 15 and higher, and of 43.5% of illiterate population over 14 years old or population that can read and write without any formal education. In addition, only 3.1% of its population works or studies in another municipality, and the average commuting time of the employed or student population is around 15 minutes. That postcode area has a health needs score of 1.29 and a commuting score of 0.74.

The medium to low access level changes occur in 15 districts in total, and produce as net results the identification of additional disadvantaged populations amounting to 3.5% of the total population. These changes make the results more in accordance with the existing supply imbalances, the health needs and the mobility of the populations, and also the behavior of the patients regarding the distance decay when traveling to hospitals. Consequently, this achievement enables an improved identification of the disadvantaged populations.

The health needs and the commuting indices contribute conjointly to 34 of the 39 medium to low access changes (with health needs scores between around

1.02 and 1.47 and commuting scores between 0.7 and approximately 0.93). The remaining five changes are due to unique contributions of the health needs index or the commuting index (two have health needs scores greater than one – of around 1.04 and 1.11 – and three have commuting scores of less than one – between approximately 0.87 and 0.94).

Finally, as regards the low to medium access level change (see Table 3), we do not find any contribution from the health needs or the commuting indices. This is a change that occurs close to the upper bound of the low access level and is due only to the introduction of the quartic function and its application after the initial catchment with no proximity barrier. In fact, the access score of that area is highly dependent on the supply of a hospital that is located in the same area. Hence, while the Epanechnikov function of the original KD2SFCA method reduces this score, such a reduction does not occur when our EKD2SFCA method is applied.

3.2. Extended competition assessment methods (steps 3 and 4 of the EKD4SFCA method)

After presenting steps 1 and 2 of the EKD4SFCA method in section 3.1, we dedicate this section to the presentation of the extended versions of the HHI index and the dominance identification method of Melnik et al. (2008), which were presented in subsections 2.2.2 and 2.2.3, respectively. These extended competition assessment methods constitute the steps 3 and 4 of our proposed EKD4SFCA method.

Before we begin the explanation of steps 3 and 4, we comment in subsection 3.2.1 on the relevant specificities of competition in the health care industry and on health care markets, since the proposed new method focuses on the health care industry.

3.2.1. Competition in the health care industry

Health services have special features, producing market imperfections that differentiate them from other products or services (OECD, 2012a). These features need to be explained because they create market failures that produce adverse effects on competition.

The main feature is the existence of information asymmetries in health care markets, namely between the health care provider and the consumer (patient), due to the imperfect information on the health services that is provided to patients.

3.2.1.1. Competition over quality and location

Health services are considered credence services, i.e., services whose quality, their necessity or the extent to which they are needed cannot be easily and properly assessed by patients (or can, but only under very high costs), both before the utilization and after the utilization (Darby and Karni, 1973). Most patients do not have perfect and complete information on health services' quality. Hence, they choose a specific health care provider following advice from his/her general practitioner or relatives and close friends who have had experience with that provider (OECD, 2012a). Or they choose based on their perception of the quality of the health services that are provided. Yet sometimes their choices are based on factors that may not represent real health care quality factors. For example, visual and comfort aspects of the waiting room of the facility, or information on waiting periods, on the room quality for inpatient care, or on the quality of the food that is served to patients (when this kind of information is available).

Health care quality is difficult to assess directly also because it is a complex concept, which is defined in many different ways. One of the main used definitions in the literature is the one by the US Institute of Medicine (IOM), according to which quality is the degree to which health services for people increase the likelihood of desired health outcomes and are consistent with professional knowledge (Goldenberg, 2012). However, this definition is not consensual, and one important factor that contributes to this lack of consensus is the multidimensional characteristic of the quality concept, which makes it difficult to grasp it. The World Health Organization (WHO) defines six quality dimensions, indicating that health services have to be effective, efficient, accessible, acceptable/patient-centered, equitable and safe (WHO, 2006).

Since most patients cannot directly observe the quality of the services, health care providers may compete with each other only over perceived quality aspects that increase patients' satisfaction. They will also compete over quality factors that affect indicators and other information provided by public authorities or accreditation/certification firms on the health services' quality, which constitute information that over time can develop reputation. Thus, even though information asymmetries and the credence characteristic of health services affect competition, it is possible for health care providers to compete over quality aspects used by patients to assess quality (Gaynor, 2006; Tay, 2003). It follows that in markets with strong competition, providers will strive to deliver high quality by trying to improve the mentioned quality factors or aspects, while the respective quality provided in markets with weaker competition will be lower.

A further cause of market failure in health care markets that has implications over competition is the moral hazard problem that occurs because most patients do not have to pay for all the health services they use. Most patients are insured with private and social insurances, as occurs in Bismarckian, i.e., social security health systems, or have the greatest part of their expenses covered by public funding, as in tax-funded Beveridge systems, namely National Health Services, like the Portuguese *Serviço Nacional de Saúde* (van der Zee and Kroneman, 2007). Hence, any price differential between health care providers is only fractionally paid for by the patients (Gaynor et al., 2014). This occurs at least in most countries members of the Organisation for Economic Co-operation and Development (OECD) and, as a consequence, most people do not feel financially hindered when consuming a health service, which makes the elasticity of demand with respect to price relatively inelastic. Accordingly, health care providers do not compete intensively with each other over price.

Thus, while on one hand there can be competition over quality factors in health care markets, on the other hand competition over price is typically minimized, and this is due especially because of information asymmetries. In fact, “if consumers have little information or don’t understand the information they have, they tend to rely on reputations, brand names, etc.”, and this “tends to decrease responsiveness of demand to prices or other factors and enhances firms’ market power” (OECD, 2012a: 353).

Finally, it is important to stress that, besides quality aspects, the geographic location of the clinics, hospitals or other health care supply points is another important variable used to compete for patients in health care markets. Health care markets’ structure is a function of the geographic configuration of patient location and health care providers’ sites (Gaynor et al., 2012b).

3.2.1.2. Health care markets

Relevant health care product markets can be public and/or private, and some examples are the following product markets: general hospital health care; long term care; primary care; specialties markets, based on medical specialties like ophthalmology, obstetrics, cardiology; and markets based on specific treatments, exams or consultations, like radiotherapy, chemotherapy, renal care/hemodialysis, rehabilitative treatments, general imaging exams, clinical analyses and dermatology consultations.

Some examples of relevant geographic markets are (even though not only for health care markets): census regions; administrative or political territories; metropolitan areas; countries; catchment areas based on distances; and catchment areas based on roads’ travel times by car (isochrones).

The combination of the relevant product markets with the relevant geographic markets allows the researcher to begin with his/her analysis of a concentration operation or any other kind of competition assessment, as the ones that can be carried out with steps 3 and 4 of our proposed method, which will be explained next.

3.2.2. The extended HHI

The original HHI developed by Hirschman and Herfindahl was presented in equation (11) (see subsection 2.2.2.2). As we recall, its application presents four limitations, which constitute problems for the assessment of competition:

(i) Disregard of border crossing of consumers. Market concentration is assessed in each area considering only the supply points located in that area.

(ii) No variability in market concentration within the relatively large areas chosen as the relevant geographic markets. By using large areas, the HHI of equation (11) does not enable the detection of market concentration variations inside these areas.

(iii) No incorporation of any measurement of distance decay. The large areas used by the HHI treat large distances inside these areas the same way as small distances.

(iv) Vulnerability to the modifiable areal unit problem (MAUP). Depending on the large geographic areas chosen to assess competition, results for a specific small area may present considerable variations.

Thus, in order to correct these problems, we introduce changes to the HHI and propose an extended HHI, which constitutes the step 3 of our EKD4SFCA method.

For the extended HHI (step 3 of our EKD4SFCA method), we adapt the HHI to be calculated with floating catchment areas and small geographic units, and extend it with the inclusion of a distance decay function, as described by equation (18):

$$HHI_i^E = \sum_{G=1}^{N}\left[\sum_{l\in\{d_{il}\le d_{\max}\}} \mathrm{Q}_{Gl}\, g(d_{il}, d_{\max})\right]^2 ; \qquad (18)$$

where HHI_i^E is the extended HHI calculated for the geographic unit i, G refers to groups of supply points, namely the firms or economic groups that own the supply points (G=1 is the largest group, G=2 is the second largest group, and so on, until the smallest group competing in the market i, namely G=N). Furthermore, $\sum_{l\in\{d_{il}\le \mathrm{d}_{\max}\}} \mathrm{Q}_{Gl}\; g(d_{il}, d_{\max})$ refers to the market share of the group G with supply points located at l that have catchments covering i, and $g(d_{il}, d_{\max})$ is the quartic distance decay function presented in equation (16) (see subsection 3.1.2).

This innovative way of calculating the HHI makes it possible to identify in a detailed way the competitive pressure between firms in each of the small geographic areas of a study region, taking into account the potential flows of the populations from their residences to the supply points. Problems (i) to (iv) above are this way addressed.

Regarding problem (i), this problem is addressed, because equation (18) takes into account all the potential flows of the populations from their residences to the supply points until a maximum travel time reference. It does not confine the assessment to specific large areas, such as districts or states, which are usually chosen as the relevant geographic markets in applications of the original HHI.

Problem (ii) is addressed by the use of small geographic assessment units, which makes it possible to detect market concentration variations inside the relatively large areas that are usually chosen as the relevant geographic markets.

Regarding problem (iii), it is addressed by the use of the distance decay function. The distance decay function applied to the calculation of the HHI has the objective of emulating the expected differences in the consumers' behavior due to the different distances that they have to travel from their residences to the supply points. Thus, it discounts the supply capacity of each supply point according to the travel time from the residences of the patients until the supply point. This way, a firm that owns a supply point that is far away from the geographic unit for which the HHI is being calculated will have a smaller market share than it would if the supply point was located nearer.

Finally, considering problem (iv), the MAUP is alleviated with the use of small geographic units, because, this way, the results do not depend on the choice of the relatively large areas, as it is the case when the original HHI is applied. For example, if the researcher identifies an HHI of 2350 in a district and an HHI of 3000 in a metropolitan region, and both regions comprise a specific postcode area, this HHI change does not occur if the researcher calculates the HHI for that specific postcode area with the extended HHI (HHI_i^E).

We note that the market shares for the computation of the HHI_i^E can be calculated with different kinds of information gathered from the firms, like volume of sales of the relevant product(s) or service(s). The supply capacity, represented by the number of employees (number of physicians), for example, can also be used for the calculation of the market shares. According to the European Commission, in its Notice on the definition of relevant markets, the capacity constitutes a valid parameter to be used in the calculation of market shares (EC, 1997).

Finally, the same interpretation of the results of the original HHI applies to the results of our HHI_i^E. Even though the calculation of the HHI_i^E depends on floating catchment areas, small geographic units and a distance decay function, the essence of the HHI is maintained: the HHI_i^E still represents fundamentally the sum of the squares of the firms' markets shares. Thus, the same HHI thresholds apply that divide market concentrations into three classes – low, medium and high (or not concentrated, concentrated and highly concentrated, respectively). An HHI_i^E until 1000 indicates that market concentration is low. If the HHI_i^E is between 1000 and 2000, then market concentration can be classified as medium. An HHI_i^E above 2000 and until the maximum value of 10000 indicates that market concentration is high.

3.2.3. The extended dominance identification method

Step 4 of the EKD4SFCA method is based on the dominance identification method of Melnik et al. (2008), which was presented in equation (12) (see subsection 2.2.3, page 57). As in the case of the HHI, for the calculations of equation (12) the relevant geographic market is a relatively large area, such as a country, a state or a district. Thus, the dominance identification method of Melnik et al. (2008) presents the same four problems of the HHI, namely problems (i) to (iv) identified in the previous subsection.

In order to correct these problems in step 4 of the EKD4SFCA method, we propose to adapt the method of Melnik et al. (2008) of equation (12), such that it can be used with floating catchment areas and small geographic assessment units, and to extend it with a distance decay function, as it is described by equation (19):

$$\mathrm{Q}_i^{D^E} = \frac{1}{2}\left\{1 - \left[\left[\sum_{l\in\{d_{il}\le d_{\max}\}} \mathrm{Q}_{1l}\, g(d_{il}, d_{\max})\right]^2 - \left[\sum_{l\in\{d_{il}\le d_{\max}\}} \mathrm{Q}_{21}\, g(d_{il}, d_{\max})\right]^2\right]\right\}; \quad (19)$$

where $\mathrm{Q}_i^{D^E}$ is the extended calculation of the dominance threshold for the geographic unit i, and the other variables, the distance decay function and $d_{\max}$ are the same ones as in equation (18), of the extended HHI.

The parameter γ of equation (12) does not appear in equation (19), because we propose to use the natural benchmark γ =1, which is the base case established by Melnik et al. (2008), Hellmer and Wårell (2009) and McIntosh and Hellmer (2012). Thus, the same value limits as in equation (12) apply here, namely the lowest dominance threshold is zero for monopolies, and the highest threshold will be 50% in the cases where the first and the second largest competitors have the same market share.

Equation (19) follows the same reasoning as equation (18), of the extended HHI, and the calculations are similar. The two competition assessment methods are applied with floating catchment areas and a distance decay function to obtain results for all small geographic units of the study region. Thus, the same solutions to problems (i) to (iv) above apply to both equations.

The practical utility of using equation (19) to assess dominance is the comparison that is carried out in this assessment, namely between the dominance thresholds and the highest market shares that are calculated for each of the small geographic units. This enables the detection of the dominated geographic units. If the market share of the firm with the largest market share is greater than the dominance threshold, then the firm has a dominant position. As proposed by Knoche and Thöni (2011) and McIntosh and Hellmer (2012), this complements the competition assessment initiated with the HHI.

3.2.4. An application of the competition assessment methods of the EKD4SFCA method

Steps 3 and 4 of the EKD4SFCA method represent the incorporation of the competition assessment into our integrated model for assessing access and competition. They can be used in a similar way as proposed by McIntosh and Hellmer (2012), namely considering the HHI over 2000 as a necessary but not sufficient condition for identifying the geographic markets with potential competition problems, and the identification of a dominant position as the necessary and sufficient condition. McIntosh and Hellmer (2012) showed that, depending on the data, the researcher can identify geographic units that are not dominated with an HHI over 2000, and he/she can also identify dominated geographic units with an HHI of 2000 or less (we also obtain such a result, as will be seen). However, only if the conditions are verified together, i.e., if the

HHI is over 2000 and the geographic unit is dominated, the market will have the highest risk of presenting competition problems. With this verification, the researcher ensures that the geographic areas that present high market concentration and market dominance at the same time are the ones with potential competition problems.

In order to better explain how steps 3 and 4 of the EKD4SFCA method are applied, we present a case study in this subsection with a practical example, focusing on the hospital health care market. For that matter, we use as reference an expert opinion issued by the Portuguese Health Regulatory Agency on a concentration operation that was notified to the Competition Authority (*Autoridade da Concorrência*) by a firm that owns hospitals, namely *José de Mello Saúde, S.A.* (JMS) (ERS, 2014c). As required by law (Law N. 19/2012, of 8 May), JMS informed that it was willing to acquire another firm that owns hospitals, *Espírito Santo Saúde – SGPS, S.A.* (ESS). JMS finally withdrew from the acquisition, after another firm made an offer with a higher price per share. Nonetheless, as it is legally required in such kinds of operations, the potential competition effects of this acquisition in the hospital health care market had to be assessed by the Portuguese Health Regulatory Agency.

The expert opinion presents the traditional competition assessment, with the application of equations (11) and (12), but also the innovative competition assessment carried out with our equations (18) and (19), as an alternative assessment of the potential competition effects. Therefore, with this case study, we are able to explain how the steps 3 and 4 of our proposed EKD4SFCA method are applied and the different results that they produce, as compared to the results of the traditional competition assessment methods.

We note that the hospitals and the numbers of physicians required for the application of the different competition assessment methods were identified at the Portuguese Health Regulatory Agency registry of health services providers in September 2014, such that the data refer to that month, since the registry is regularly updated. The numbers of physicians were used for the calculation of the market shares.

As another relevant introductory note, we remark that the Portuguese health system is a Beveridgean type of National Health Service (*Serviço Nacional de Saúde*), characterized mainly by a public tax-funded health care provision guaranteed for all citizens. When accessing public hospitals for general hospital health care, patients pay at maximum relatively low user fees, but they have to be referred by the public primary care providers to a specific hospital (the only exception is urgent care, which can be accessed at any hospital). Thus, the Portuguese patients do not have freedom of choice, when they obtain general hospital health care in public hospitals, which means that public hospitals do not compete for patients. This is also the main explanation of the Portuguese Health Regulatory Agency to consider only private hospitals in its competition assessment (ERS, 2014c). There is only competition for patients in the private hospital health care sector, in which patients can freely choose private hospitals. The private hospital health care market is a market for the patients that wish to have freedom of choice. However, in order to freely choose their hospitals, and obtain care in convenient locations, with differentiated perceived quality aspects, they pay more – directly, with out-of-pocket expenses, or indirectly, via insurance.

We further note that the European Commission also separated the market for private hospital health care in the analysis of a concentration (Case N. COMP/M.4367 – APW / APSA / NORDIC CAPITAL / CAPIO). As remarked in that analysis, regarding the United Kingdom (also a Beveridgean country), private hospitals were considered to respond to a specific demand for hospital health care, which is different from the demand for public hospitals, and this was confirmed by the market investigation that was carried out (EC, 2007).

Thus, the relevant product market in our case study is defined as the market for hospital health care, which is, more specifically, the market for multi-specialty secondary care provided by the private hospitals, including care from medical, surgical and diagnostic specialties. The firms that compete in this relevant product market are the private firms that own the hospitals, i.e., the supply points, where the patients get access to hospital health care. This definition of hospitals includes the supply points that provide only outpatient care, when they are integrated in a multihospital organization with at least one supply point that also provides inpatient care.

3.2.4.1. The traditional competition assessment

The traditional competition assessment, which has been followed by the Portuguese Health Regulatory Agency in recent years, consists of applying equations (11) and (12) – the original HHI and the market dominance identification method of Melnik et al. (2008), respectively (see subsections 2.2.2.2 and 2.2.3).

The adoption of equations (11) and (12) to assess competition requires the definition of relatively large relevant geographic markets. The Portuguese Health Regulatory Agency typically defines the relevant geographic markets for hospital health care as the regions of the Nomenclature of Territorial Units for Statistics, level III (NUTS III), which divides continental Portugal in 28 parts (ERS, 2014c). The NUTS are statistical regions defined with the objective of grouping adjacent municipalities with similar socioeconomic problems, challenges and profiles. In Portugal, they were created by the Decree-Law N. 46/89, of 15 February, and they have been suffering changes across the years, introduced by the Decree-Law N. 163/99, of 13 May, the Decree-Law N. 317/99, of 11 August, the Decree-Law N. 244/2002, of 5 November, and Law N. 21/2010, of 23 August.

The 28 NUTS III of continental Portugal are illustrated in Figure 14.

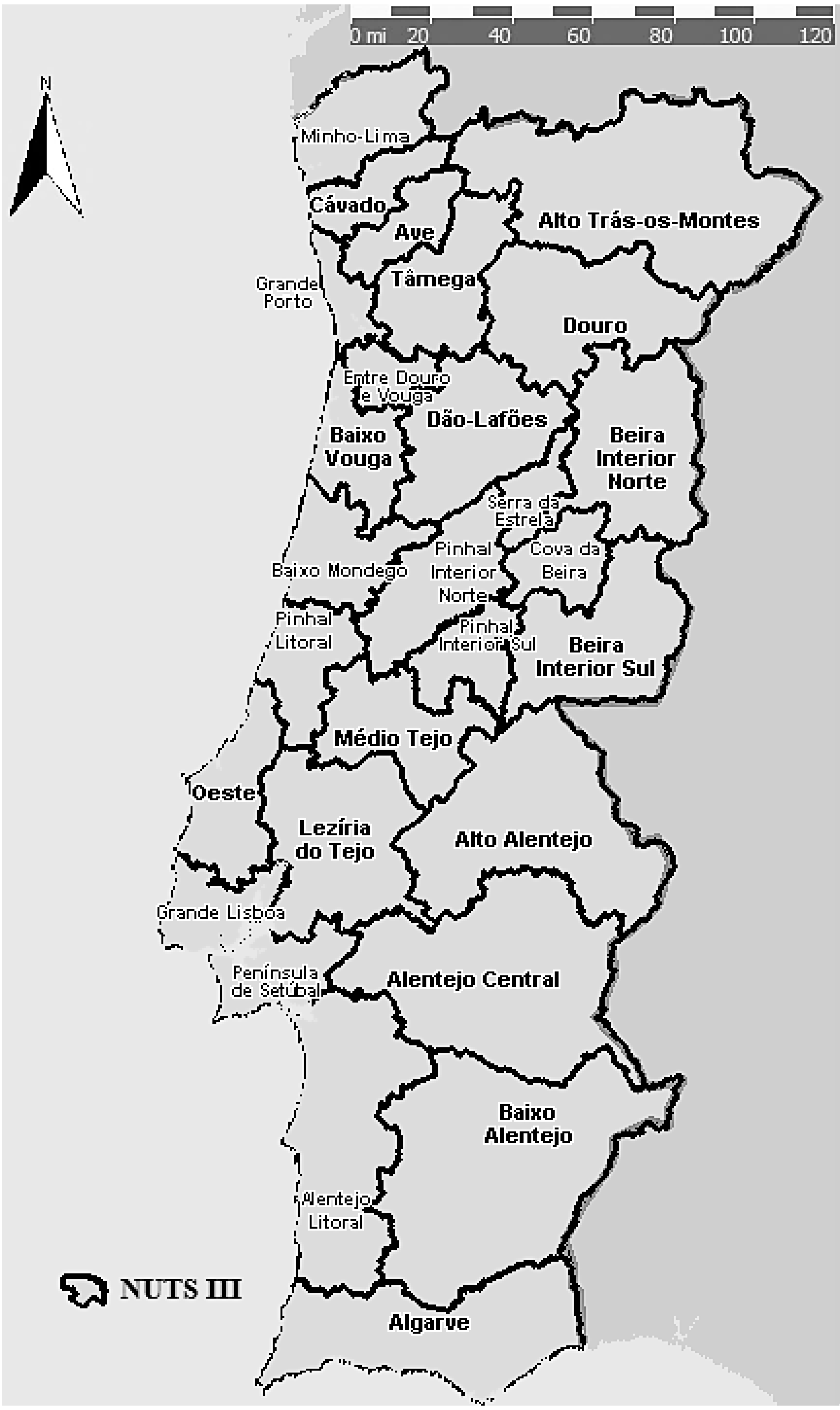

Figure 14 – NUTS III of continental Portugal

The competition assessment of the acquisition of ESS by JMS carried out with equations (11) and (12) focused on the regions with supply points from ESS and JMS, and this simultaneous presence of supply points of the two firms occurs only in Grande Lisboa and Grande Porto. This focus is justified because, if the acquisition would have been successful, only in these two relevant geographic markets there would be any change in terms of number of competing firms and market share, with variations in the HHI and the market dominance identification. While there are still other five NUTS III where ESS supply points are active, in these other markets there are no JMS supply points, such that a mere transfer of ownership would occur, with no other implication in terms of competition (ERS, 2014c).

The HHI assessment of the Portuguese Health Regulatory Agency followed the criteria that were defined by the European Commission for concentrations (EC, 2004). These criteria define that a concentration may produce competition problems, if:

(i) The HHI change is equal to or more than 250 points, when the HHI after the concentration in the relevant market is between 1000 and 2000; or

(ii) The HHI change is equal to or more than 150 points, when the HHI after the concentration in the relevant market is above 2000.

Regarding the market dominance assessment, the Portuguese Health Regulatory Agency applied equation (12) to verify if there would be market dominance after the concentration, which makes it possible to identify if the acquisition would create a dominant position of the acquiring firm, JMS. The European Commission considers that the creation or the strengthening of a dominant position is a primary form of the likely competitive harm that can be caused by a concentration (EC, 2004).

Both the HHI and the market dominance assessments assume that the concentration operation entails only the transfer of ownership of the supply points from ESS to JMS. The two firms become just one, namely JMS, the acquiring firm, such that, considering that the concentration would occur, their market shares are added in the *ex post* setting, i.e., after the concentration. As identified at the time of the elaboration of the opinion, in September 2014, there were four ESS supply points and nine JMS supply points in the relevant market of Grande Lisboa, and there were three ESS supply points and two JMS supply points in the relevant market of Grande Porto. This means that, in the *ex post* setting, JMS would have 13 supply points in Grande Lisboa (4 + 9) and five supply points in Grande Porto (3 + 2).

Table 4 presents the results that were obtained, namely the results *ex ante* and *ex post*, indicating, more specifically, the HHI – before, after and the variation that occurs –, and the market dominance identification – before and after. These are approximated results, as can be seen, but with sufficient detail to conclude, based on the adopted analytical criteria, whether competition problems would be expected to arise, as a result of the concentration. The fully detailed results cannot be disclosed, due to confidentiality issues.

Table 4 – Traditional competition assessment's results

NUTS III	HHI			Market dominance	
	Before	**After**	**Change**	**Before**	**After**
Grande Lisboa	<2.000	>2.000	>150	X	✓
Grande Porto	<2.000	>2.000	>150	X	X

As Table 4 indicates, in these two markets – Grande Lisboa and Grande Porto – the criteria adopted to identify potential competition problems with the HHI were met: after the concentration, the HHI would become higher than 2000 points and the change would be higher than 150 points. Besides, the analysis in Grande Lisboa identified that with the concentration JMS would obtain a dominant position. These results indicate that the concentration should be rejected, in order to avoid competition problems that could harm other firms and the patients. As we recall, competition problems, consisting of lack of competition and abuse of a dominant position, can produce entry barriers to potential competition, eliminate actual competition, and affect access negatively via discrimination of patients and the provision of reduced health care quality.

In order to complement this assessment, however, the Portuguese Health Regulatory Agency adopted our steps 3 and 4 of the EKD4SFCA method – equations (18) and (19), respectively –, and this assessment is explained next, together with the comparison of the results.

3.2.4.2. Application of the extended versions of the HHI and the dominance identification method

In this subsection we apply our extended versions of the HHI and the dominance identification method and their results are compared with the results of the original HHI and dominance identification methods, which were presented in the previous subsection.

As the Portuguese Health Regulatory Agency recognizes in its expert opinion, the competition assessment with our extended versions of the HHI and the market dominance identification methods of equations (18) and (19), respectively, enables an analysis that will better reflect competition in reality (ERS, 2014c). This is due to the fact that these new competition assessment methods overcome limitations of the original methods, namely the following four, as we recall:

(i) Disregard of border crossing of patients. Competition is assessed in each area considering only the supply points located in that area.

(ii) No variability in competition within the relatively large areas chosen as the relevant geographic markets. By using large areas, equations (11) and (12) do not enable the detection of variations in the competitive pressure inside these areas.

(iii) No incorporation of any measurement of distance decay. The large areas treat large distances inside these areas the same way as small distances.

(iv) Vulnerability to the modifiable areal unit problem (MAUP). Depending on the large geographic areas chosen to assess competition, results for a specific small area may present considerable variations.

The relevant geographic markets considered by the Portuguese Health Regulatory Agency in this complementary assessment were 90 minutes catchment areas, defined by the maximum travel time that the patients will accept to travel for general surgeries, which are typically provided in hospitals. As opposed to focusing only on Grande Lisboa and Grande Porto, this alternative assessment follows what was explained in subsection 2.2.1, namely that the firms compete when their catchment areas overlap. The geographic areas covered by the overlapping catchment areas of the supply points of the two firms – JMS and ESS – indicate that the residents of these areas have two firms to choose from, when deciding to obtain their needed care. Thus, we now focus on the overlapping areas of the supply points owned by JMS and ESS. These are the only geographic areas that would present changes in the competitive situation, as a result of the concentration operation.

We further note that the 460 postcode areas of continental Portugal were used as the small geographic assessment units, like in our application of the EKD2SFCA method, presented in subsection 3.1.4. As points of departure and arrival for the calculation of the travel times, the centroids of the postcode areas were also used. Likewise, only these centroids were used for the geo-referentiation of the supply points and the populations. The same GIS application was adopted, namely Microsoft MapPoint 2009.

Figure 15 presents the overlapping geographic areas, where there is competition between JMS and ESS in the *ex ante* setting, namely the regions formed by the 291 postcode areas that are covered by the catchment areas of the two firms, JMS and ESS, in a simultaneous way. They can be identified in 20 NUTS III, and in 14 of these 20 NUTS III, the competition areas cover more than 50% of their resident population.

Table 5 presents the 20 NUTS III and identifies the 14 NUTS III where JMS and ESS compete for more than 50% of their residents, who are potential patients to the hospital health care that the two firms provide.

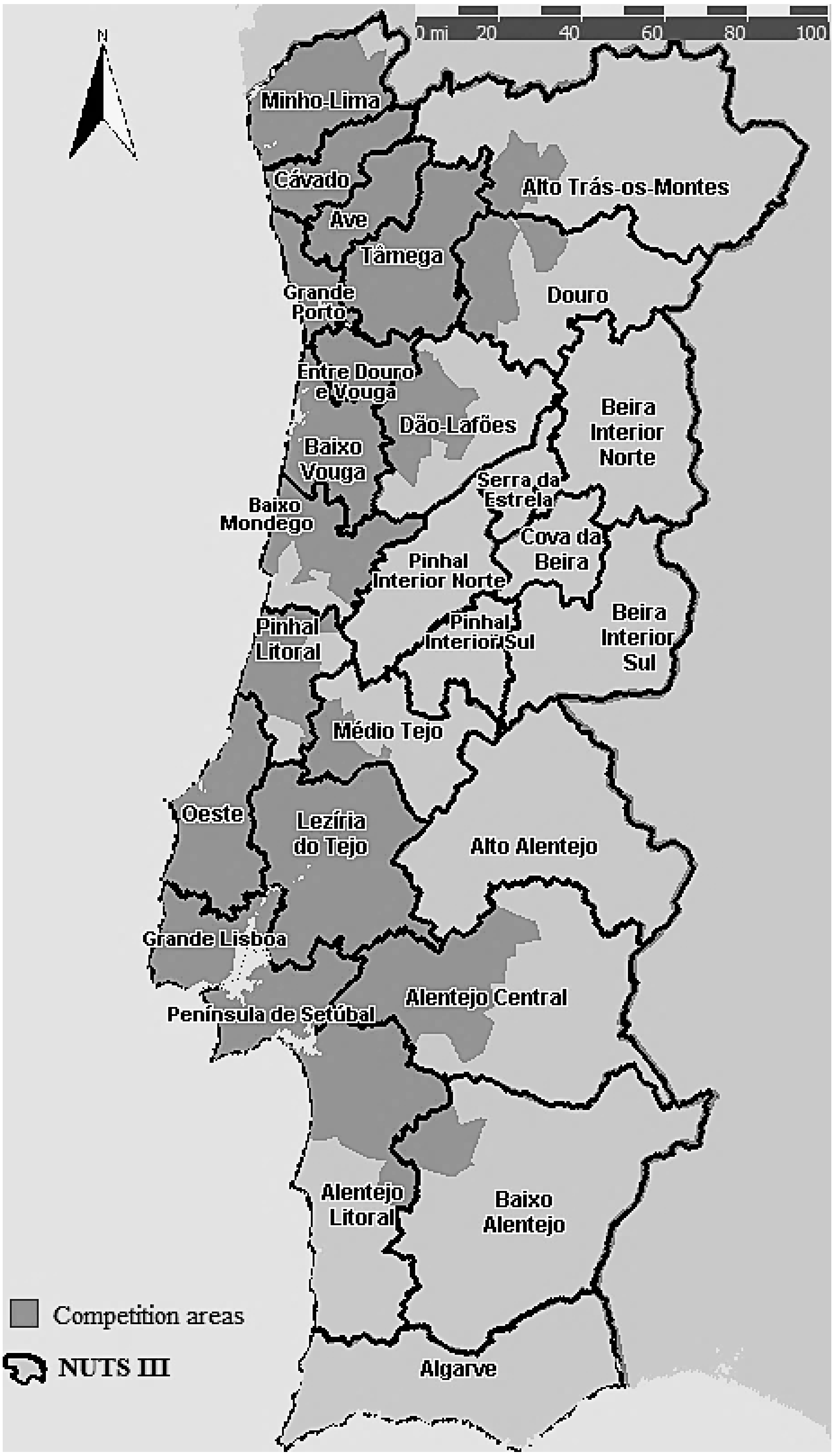

Figure 15 – Identification of the geographic areas where JMS and ESS compete for patients (competition areas)

Table 5 – NUTS III where JMS and ESS compete and identification of the proportion of the potentially affected population in each NUTS III

NUTS III	Population in the areas disputed by JMS and ESS	Total population	%
Alentejo Central	64094	166822	38%
Alentejo Litoral	31990	97925	33%
Alto Trás-os-Montes	21691	204381	11%
Ave	511737	511737	100%
Baixo Alentejo	8255	126692	7%
Baixo Mondego	284631	332326	86%
Baixo Vouga	390822	390822	100%
Cávado	410169	410169	100%
Dão-Lafões	124200	277240	45%
Douro	132522	205902	64%
Entre Douro e Vouga	274859	274859	100%
Grande Lisboa	2042477	2042477	100%
Grande Porto	1287282	1287282	100%
Lezíria do Tejo	247453	247453	100%
Médio Tejo	107595	227999	47%
Minho-Lima	235623	244836	96%
Oeste	362540	362540	100%
Península de Setúbal	779399	779399	100%
Pinhal Litoral	213416	260942	82%
Tâmega	550516	550516	100%

Thus, with our extended competition assessment methods we can conclude that there is not only competition between JMS and ESS in two NUTS III, i.e., Grande Lisboa and Grande Porto, as the traditional assessments indicate (see Table 4), but rather in 20 NUTS III, with substantial competition in 14 of them.

Furthermore, the traditional competition assessments identify competition only between a limited number of ESS and JMS supply points in Grande Lisboa and Grande Porto. In turn, our extended assessments identify that all the existing 14 ESS supply points and 12 JMS supply points in Portugal, located in different NUTS III, exert competitive pressure to each other. Therefore, while in the traditional assessment only 18 ESS and JMS hospitals are considered, all 26 ESS and JMS hospitals are taken into account when our extended competition assessment methods are applied. The additional supply points considered are ESS hospitals located in Alentejo Central, Baixo Vouga, Minho-Lima, Península de Setúbal and Tâmega, and a JMS hospital located in Oeste. The competition exerted by these hospitals is ignored in the traditional competition assessments.

Now, considering the areas with competition problems after the concentration, it is possible to identify with the application our extended competition assessment methods that around 81% of the 291 competition areas become problematic areas with the concentration: 237 postcode areas in total.

Figure 16 illustrates these 237 postcode areas grouped in three kinds of regions with competition problems: (i) regions constituted by 130 areas identified solely under the HHI level and change criteria ("HHI"), (ii) one region constituted by a unique postcode area with the exclusive verification of market dominance ("Dominance"), and (iii) regions formed by 106 areas that meet the HHI and the dominance criteria simultaneously ("HHI and dominance"). These latter regions with competition problems can be identified in 18 NUTS III.

Table 6 presents the 18 NUTS III with competition problems after the concentration. It also indicates the proportion of population covered by the regions with competition problems in each of the NUTS III, pointing out the NUTS III where the competition problems are more extensive. In 12 NUTS III the competition areas with problems after the concentration cover more than 50% of their residents.

Therefore, while the traditional assessments only identify problems in two NUTS III, as a result of the acquisition of ESS by JMS, the application of our extended assessment methods identifies problems in 18 NUTS III with the analytical criteria adopted by the Portuguese Health Regulatory Agency. Besides, in 12 NUTS III these problems are extensive, considering that the affected areas cover more than 50% of their residents: 10 additional NUTS III, besides Grande Lisboa and Grande Porto.

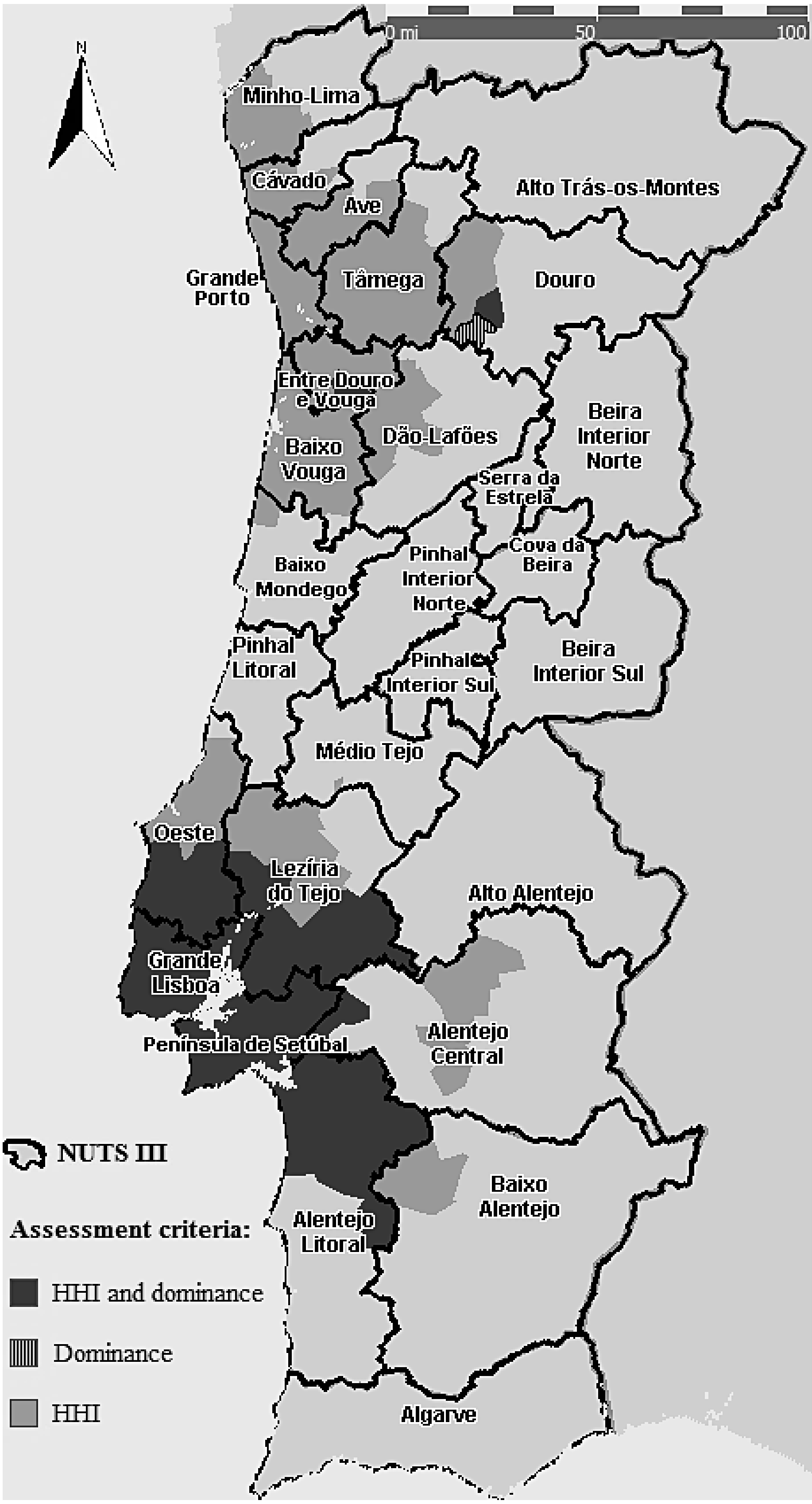

Figure 16 – Identification of problems with the extended competition assessment methods

Table 6 – NUTS III with competition problems identified with the extended competition assessment methods

NUTS III	Population in the areas with competition problems	Total population	%
Alentejo Central	46657	166822	28%
Alentejo Litoral	31990	97925	33%
Ave	442517	511737	86%
Baixo Alentejo	8255	126692	7%
Baixo Mondego	12465	332326	4%
Baixo Vouga	370394	390822	95%
Cávado	310527	410169	76%
Dão-Lafões	67839	277240	24%
Douro	121806	205902	59%
Entre Douro e Vouga	274859	274859	100%
Grande Lisboa	2042477	2042477	100%
Grande Porto	1287282	1287282	100%
Lezíria do Tejo	231868	247453	94%
Médio Tejo	20206	227999	9%
Minho-Lima	158160	244836	65%
Oeste	355944	362540	98%
Península de Setúbal	779399	779399	100%
Tâmega	519769	550516	94%

Finally, we mention some observations of the Portuguese Health Regulatory Agency in the opinion regarding travel times, which indicate that our extended competition assessment methods make more sense than the traditional assessment methods:

- In the NUTS III of Grande Porto, there are 40 firms competing for their residents, with some supply points located in adjacent NUTS III – Ave, Cávado and Tâmega – that are accessible in less than 30 minutes travel time to the residents of Grande Porto.
- There are JMS supply points in Grande Porto that compete for patients in Baixo Vouga: part of the residents of Baixo Vouga travel around 30 minutes to reach these supply points (much less than the 90 minutes maximum travel time).

- There are JMS supply points in Grande Porto that compete for patients in Minho-Lima: part of the residents of Minho-Lima travel less than 40 minutes to reach these supply points.
- There are JMS supply points in Grande Porto that compete for patients in Tâmega: part of the residents of Tâmega travel less than 30 minutes to reach these supply points.
- There are JMS supply points in Grande Lisboa that compete for patients in Península de Setúbal: part of the residents of Península de Setúbal travel less than 30 minutes to reach these supply points.

3.3. How to apply the EKD4SFCA method

In section 3.1, we present steps 1 and 2 of our EKD4SFCA method, and the assessment of patients' access to health care that can be carried out with these two steps. In section 3.2, in turn, we present steps 3 and 4 of our EKD4SFCA method, and the assessment of competition between health care providers that can be carried out with these two steps. In these two sections, we show that our innovative, extended access and competition assessment methods enable improved assessments. We concluded this after comparing our respective assessments with the ones carried out without our innovations, in two case studies (subsections 3.1.4 and 3.2.4). However, these two assessments were partial assessments, since they are not integrated.

In an integrated access and competition assessment with our EKD4SFCA method, the researcher will apply the EKD2SFCA method first, as shown in subsection 3.1.4. Thus, he/she will first use steps 1 and 2 of the EKD4SFCA method to compute the access scores and apply a cluster analysis with a k-means algorithm, in order to identify the underserved regions, where the disadvantaged populations reside. Thereafter, the researcher will apply steps 3 and 4 of the EKD4SFCA method, in order to identify the regions where there are no competition problems. Consequently, in an integrated assessment, the researcher will be able to identify the underserved regions with no competition problems.

A full application of the EKD4SFCA method is presented in the next chapter, in order to demonstrate how the method is applied, and the logic of its application.

3.4. Conclusions

This chapter presented the development and the constituent parts of our proposed EKD4SFCA method. It is a new method that was developed to assess patients' access to health care and competition between health care providers in an integrated way. Methodologically, it is an enhanced 2SFCA-based method that incorporates innovative calculations of the HHI index and the dominance identification method.

The application of the EKD4SFCA method, which relies on a GIS, consists of four steps: the first two steps refer to the new 2SFCA-based method, which is an extended KD2SFCA method. The third and fourth steps refer to the competition assessment. Equations (14), (15), (18) and (19) indicate the required calculations in each of the four steps, respectively (see subsections 3.1.1.2, 3.1.1.3, 3.2.2 and 3.2.3, pages 73, 78, 107 and 109, respectively).

We presented two case studies, in order to explain steps 1 and 2, and steps 3 and 4, respectively, and how they are applied. In subsection 3.1.4, we presented an application of steps 1 and 2 of our method to assess the patients' access to health care, and in subsection 3.2.4, we presented an application of steps 3 and 4 of the method to assess the competition between health care providers. In these two case studies, we showed that our innovative, extended access and competition assessment methods enable improved assessments of access and of competition, with a more accurate identification of regions with low access and with competition problems, respectively. We achieved this by comparing the assessments conducted with our methods with the assessments carried out with present, commonly accepted and adopted kinds of methods. The differences in the results that were identified in these comparisons are explained by the facts that our methods integrate more relevant data for an improved assessment and overcome limitations of the present methods, as explained in the sections 3.1 and 3.2.

In turn, in section 3.3, we indicated how to combine the four steps in an integrated access and competition assessment. As noted in the introduction of this thesis and, in a summarized way, in the beginning of this chapter, the usefulness of the integrated assessment enabled by the EKD4SFCA method can be established for firms and public policies. Regarding the usefulness of this integration for firms, the results of a combined assessment can help in aligning the firms' supply decisions with their strategic and market positioning objectives. The integrated assessment can enable the identification of target markets, where there is relatively more unmet potential demand, as identified by steps 1 and 2, and where no competition hindrance by incumbent firms is expected, as indicated by steps 3 and 4. These market entry targets will be the undersupplied regions that do not present the risk of occurrence of competition problems. In these regions, market entrants will manage to enter a market with a relatively higher chance of success, i.e., of survival and growth. This is due to the opportunity of providing its products or services to underserved populations and the minimized risk of having to face strong competitive pressure or strategic entry deterrence by existing competitors. For the same reasons, it will also be relatively easier for market entrants to expand their market shares and establish their brand names in these markets without having to face a high initial investment and high market entry costs.

The integrated assessment of access and competition is also relevant and useful for public policies destined to improve the distribution of supply resources, reduce access inequalities and improve the populations' health status. Public administrations can achieve this by exploring competition with public bids to contract private supply, in order to increase supply directed to better serve the disadvantaged populations, and at the same time define efficient prices for the contracted care. Public bidding for contracting firms in competitive health care markets identified with the EKD4SFCA method can profit from the existing competitive pressure to establish lower prices than it would be possible without competition. Alternatively, public administrations can direct public financing for entrant firms with free adhesion contracts, in order to increase the supply and the competition in the underserved regions with competition problems. This is a way to attract new entrants to these geographic markets, in order to increase the competition, and thereby promote the benefits of the competition for the populations, and avoid the negative implications of lack of competition identified in chapter 2.

Finally, even though the EKD4SFCA method does not identify the optimal locations for the supply points to serve the resulting target geographic areas, it could play a role as a location-allocation model. As defined by Matsumoto et al. (2012), location-allocation models are methods for deciding locations for supply points and simultaneously assigning spatially-distributed sets of demands to these supply points. In this case, the sets of demands would be the target areas identified by the EKD4SFCA method, and the best locations for the firms would be the ones with catchments that maximize the coverage of these specific areas (F. Wang and Tang, 2013).

4. Supporting market entry decisions with the EKD4SFCA method

This chapter presents a full application of the EKD4SFCA method, and we use it as an example to illustrate how it should be applied and its logic. The method is applied to the Portuguese hospital health care sector, with the objective of identifying target markets for entry (Polzin et al., 2013, 2015). The target markets are defined as the underserved geographic regions, where patients have low access to hospital health care, in which the market concentration is low, with no incumbent firm dominating the market. Therefore, in these markets, entrants will have relatively more demand to compete for with the incumbent firms without facing strategic impediments that could occur if the markets were highly concentrated and dominated.

In this case study, the private hospital health care market is the relevant product market, and we use the maximum travel time reference of 90 minutes for the catchment areas (the relevant geographic market), as in the previous chapter. The information regarding the hospitals, their locations, their ownership and the number of physicians in each hospital, were collected from the Portuguese Health Regulatory Agency registry, in December 2011. It was possible to identify 89 private hospitals. Figure 17 illustrates the geographic distribution of these hospitals with circles that represent the number of hospitals located in each postcode area (the sizes of the circles vary according to the number of hospitals).

The demographic and socioeconomic data used for steps 1 and 2 were 2011 Census data collected from the Portuguese National Institute of Statistics for the Portuguese localities. As before, these data were aggregated to the postcode areas level, since the chosen geographic units were the 460 postcode areas of continental Portugal.

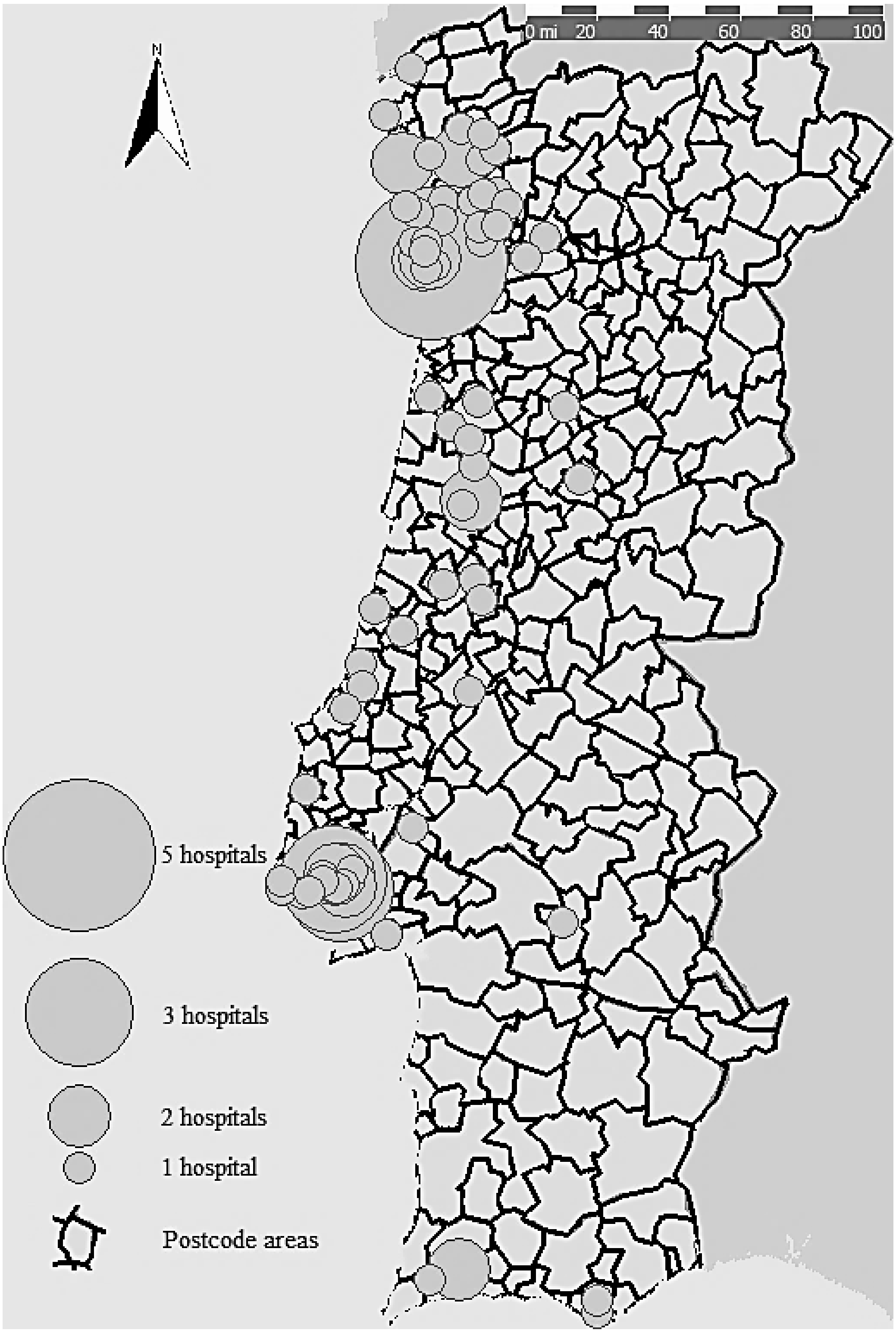

Figure 17 – The geographic distribution of the 89 private hospitals

4.1. Steps 1 and 2 of the EKD4SFCA method

For the health needs index in step 1, we used variables that are known to represent populations' health needs, which are associated with differences in health care utilization. Following the process described in subsection 3.1.1, we first chose an initial set of 10 variables and collected data for them. These initial variables were the following:

- Proportion of population aged 0 to 14;
- Proportion of population aged 0 to 10;
- Proportion of population aged 0 to 14 and 65 and higher;
- Proportion of population aged 0 to 10 and 65 and higher;
- Proportion of population aged 65 and higher;
- Proportion of economically inactive population;
- Proportion of illiterate population of over 10 years old;
- Proportion of population with no level of instruction;
- Proportion of population with no completed level of instruction; and
- Proportion of persons with one or more deficiencies.

Thereafter, following the iterative approach also shown in subsection 3.1.1, four variables were selected: the proportion of population with ages below 15 and over 65 years old, the proportion of population with one or more deficiencies, the proportion of economically inactive population, and the proportion of illiterate population of over 10 years old.

Regarding the commuting index in step 2, we first selected a set of 10 commuting variables, namely the following:

- Proportion of population that works or studies in another municipality;
- Proportion of population that works or studies in another municipality with commuting of over 60 minutes until 90 minutes;
- Proportion of population that works or studies in another municipality with commuting of over 90 minutes;

- Proportion of population that works or studies in another municipality with commuting of over 60 minutes;
- Proportion of population that works or studies in another municipality and uses primarily car as driver in commuting;
- Proportion of population that works or studies in another municipality and uses primarily car as driver or passenger in commuting;
- Proportion of population that works or studies in another municipality and uses primarily car or bus in commuting;
- Proportion of population that works or studies in another municipality and uses primarily car, bus or metropolitan in commuting;
- Proportion of population that works or studies in another municipality and uses primarily car, bus, metropolitan or train in commuting; and
- Mean travel time in commuting of the employed or student population.

After applying PCA and the iterative approach, the chosen variables were: the proportion of population that works or studies in other municipalities, the proportion of population with commuting of over one hour travel time to other municipalities, average commuting time, and the proportion of population that uses car, bus, subway or train as the main transport in commuting.

We first applied steps 1 and 2 of our method (equations (14) and (15)) to assess the patients' access to hospital health care and defined three access levels using a k-means algorithm: high, medium and low (as indicated in subsection 3.1.3). Table 7 presents the results obtained after applying the first two steps of our EKD4SFCA method, indicating for each access cluster the number of areas covered, the percentages of population covered and the access scores' means.

We note that, as a result of the first two steps, the low access level included 38 postcode areas with no associated access scores. These areas are not covered by the hospitals' catchment areas, as they are located more than 90 minutes away from a hospital, such that they can be considered unexploited markets. They have access scores equal to zero, which indicates that they belong in fact to the lowest level.

Table 7 – Access scores calculated with steps 1 and 2 of the EKD4SFCA method

Statistic	**Cluster**		
	Low	**Medium**	**High**
Number of areas	187	123	150
Population covered (%)	21.2	26.4	52.3
Mean access score	8.3	55.8	120.7

The two bounds of the access levels defined by the clustering algorithm were 32.0 and 87.9 physicians per 100,000 inhabitants. As shown in Table 7, there are 187 low access postcode areas, 123 medium access areas and 150 high access areas.

Figure 18 illustrates these results. More specifically, it illustrates the spatial distribution of the low, medium and high access regions.

As in the case of the assessment presented in subsection 3.1.4, Figure 18 also gives an indication that Portugal has pronounced supply imbalances. As can be seen in the previous figure (Figure 17), there is a higher concentration of private hospitals in the urban coastal areas. These are the largest private hospitals with the highest numbers of physicians. Therefore, the high access regions can be seen mainly in the districts of Lisboa, Porto and Braga, with the districts of Setúbal, Aveiro and Viana do Castelo also presenting high access areas. In turn, seven districts have only, or clearly predominantly populations with low access to private hospital health care, namely the districts of Beja, Portalegre, Castelo Branco, Guarda, Viseu, Bragança and Vila Real.

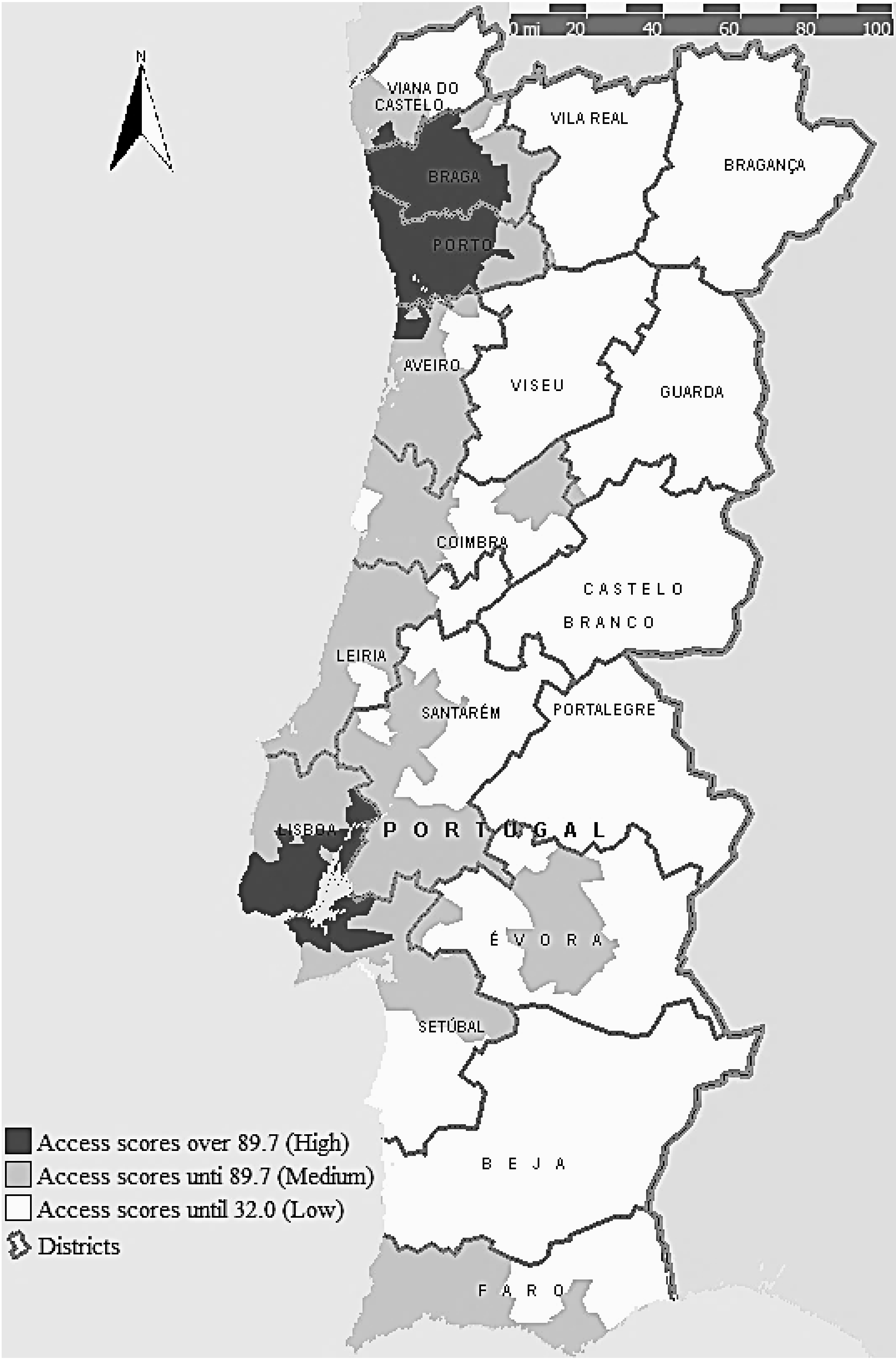

Figure 18 – Spatial distribution of the access levels identified with steps 1 and 2 of the proposed EKD4SFCA method

4.2. Steps 3 and 4 of the EKD4SFCA method

Having applied the first two steps of the proposed method, we now present the results of the application of step 3, which is the extended HHI (equation (18)). We applied step 3 considering the same supply data, maximum travel time reference and distance decay function as in steps 1 and 2, and calculating the market shares with the number of physicians. The method was able to identify 228 postcode areas with high market concentration, 194 postcode areas with medium concentration and no postcode area with low concentration. These results can be visualized in Figure 19, which presents the spatial distribution of the different hospital health care market concentration levels, showing the regions with each calculated market concentration degree: high and medium. The regions formed by 38 postcode areas that are not covered by the 90 minutes catchment areas of the supply points are also depicted (the white areas). The high concentration markets can be seen in 16 of the 18 districts, while the markets with medium concentration can be found only in coastal regions. These markets with medium concentration are mostly identified around the district of Lisboa and the northern districts, in or closer to regions where there are more hospitals and more alternatives from which the patients can choose.

The main contribution of step 3 in this case study is the identification of areas with medium concentration, namely the light grey areas in Figure 19. Together with the white areas, they form the regions where entrant firms shall enter the market and provide their hospital health care services with diminished risk of competition problems, as approximated by the market concentration levels. Considering the relation between market concentration and competition in the literature established in section 2.2, this means that these regions will not present the negative effects of weak competition (or lack of competition). Accordingly, entry and post-entry growth of firms serving these areas are expected to be less difficult and less costly.

Figure 19 – The market concentration levels obtained with step 3 of the EKD4SFCA method and the regions that are not covered by the catchment areas

The results of step 3 have to be seen in conjunction with the access clusters that are computed with steps 1 and 2, and the extended market dominance identification method's results that are obtained with step 4. In particular, with the fourth step of our method it is possible to identify a few highly concentrated geographic markets, which are also not expected to pose competition problems to entrants. These markets are the highly concentrated markets that are not dominated. As explained before, the extended dominance identification method presented in equation (19), in subsection 3.2.3, computes a threshold that represents the market share, above which a firm has a high enough market share to dominate the market. Hence, any firm that has a market share beyond this limit will not suffer sufficient competitive pressure of the competing firms in the relevant market. Step 4 enables the identification of all postcode areas that are dominated by a firm.

Accordingly, with the application of step 4 it is possible to identify 136 dominated postcode areas, which represent around 30% of all 460 postcode areas and about 50% of the 228 highly concentrated postcode areas (we did not find dominated areas that are not highly concentrated, in this case). The dominated private hospital health care regions of continental Portugal can be seen in Figure 20.

Figure 20 reveals an almost continuous coastal region from the district of Setúbal, in the south, until the district of Viana do Castelo, in the north, which is not dominated (the light grey areas). The other regions are dominated regions (the dark grey areas), or they are regions that are not covered by any catchment area of the supply points (the white areas, which represent the unexploited markets). These are partial results, however, because, as in the case of the results of step 3, they have to be seen together with the results of the other steps. Thus, they have to be assessed together with the results of steps 1 and 2 (the access levels illustrated in Figure 18) and of step 3 (the market concentration levels illustrated in Figure 19).

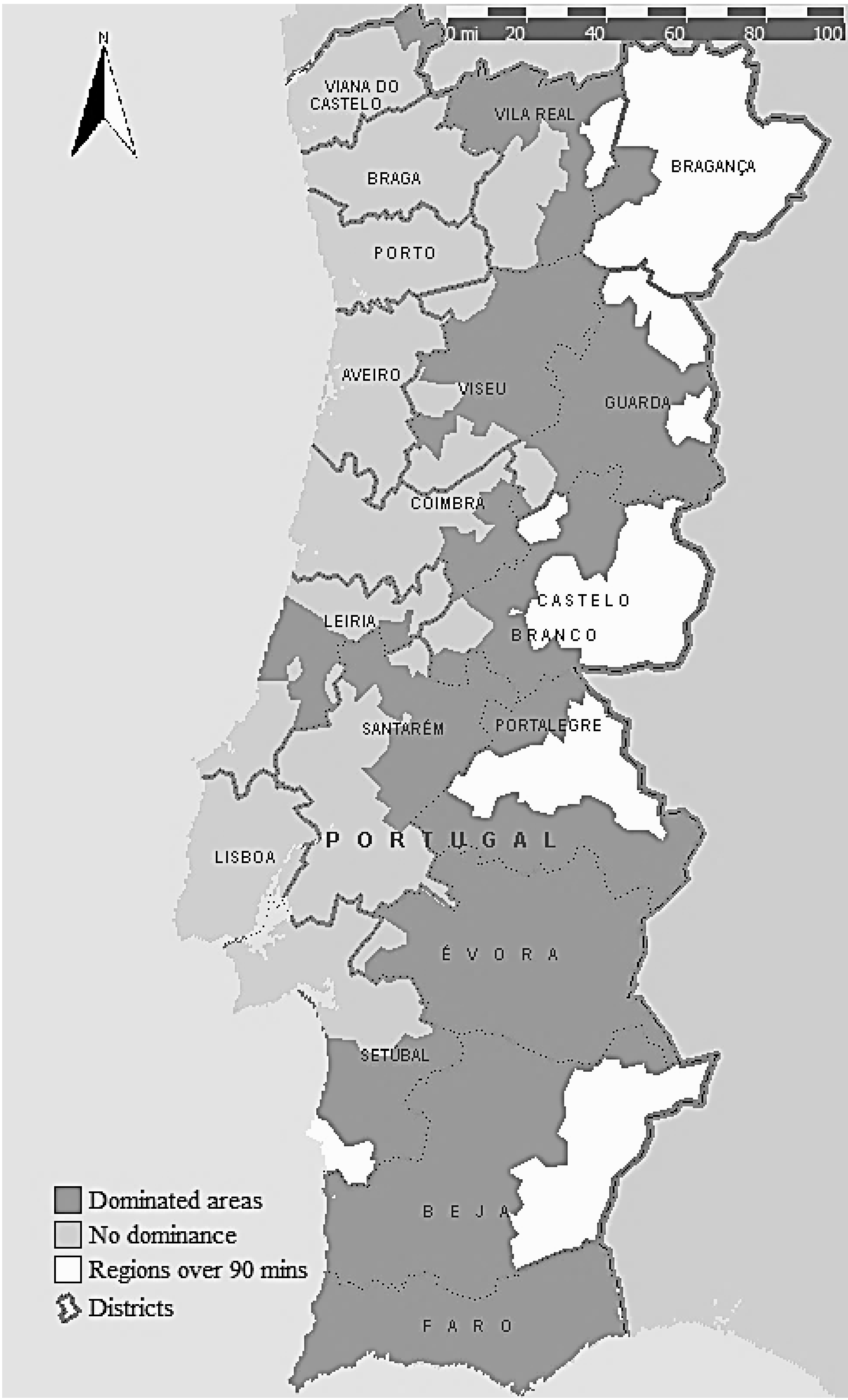

Figure 20 – Dominated private hospital health care regions identified with step 4 of the EKD4SFCA method

4.3. Final results

With the four steps of our EKD4SFCA method, it is possible to identify all low access postcode areas with the highest risk of occurrence of competition problems, i.e., with an extended HHI (HHI_i^E) higher than 2000 and the ones with a dominant position. Conversely, it is possible to identify all low access areas that present the lowest risk of occurrence of harmful competition problems – the not dominated low access areas with HHI_i^E equal to or lower than 2000 –, which we define as the market entry targets. Table 8 presents the final results that are obtained with the proposed EKD4SFCA method.

Table 8 – Identification of the market entry targets

Competition assessment	**Low access areas**	**Pop. covered (%)**
Areas with HHI>2000 and dominated areas	106	11.7
Market entry targets	81	9.5
Total	187	21.2

Focusing the assessment on the 187 postcode areas with low access, namely the relatively underserved areas, we identify 106 highly concentrated and dominated areas, as shown in Table 8. These 106 areas are low access, underserved areas that may present competition problems for market entrants, since they are highly concentrated areas (i.e., areas with an HHI of over 2000 points) and are dominated. The remaining 81 low access areas are the market entry targets, because these are areas that will not present the risk of occurrence of competition problems. These targets represent 17.6% of the 460 postcode areas and encompass a population that corresponds to 9.5% of the total population (around 955,000 residents). Thus, potential demand in the target regions sums up to about 955,000 patients.

Figure 21 shows these results presenting the dark shaded areas as the target regions formed by the 81 postcode areas, the light shaded areas as the relatively underserved areas with potential competition problems formed by the 106 postcode areas, and the white areas as the areas that do not present low access (formed by the 273 medium and high access postcode areas).

These results seem to be reasonable and valid, since the white areas are seen mainly in the urban coastal regions, where physicians in Portugal are mostly concentrated – especially in Porto and Lisboa, where there is a higher concentration of the largest hospitals. These are the regions, where the demand shall already be satisfied by the health care provision of the existing supply points, and where the incumbents shall have their brand and reputation well established. Consequently, in these regions, patients are not expected to change to new providers easily, because of the information asymmetries that are typical in health care markets (see subsection 3.2.1). This enhances the existing firms' market power and complicates the entry of new firms with new supply points.

In turn, the dark and light shaded areas are the low access, underserved areas, which can be identified in 16 of the 18 districts of continental Portugal. As can be seen, only the districts of Porto and Lisboa do not have low access areas. While the light shaded areas are areas that shall be difficult and costly for entry and expansion, because they are highly concentrated, dominated markets, the dark shaded areas are the areas to be targeted by new entrants. These are underserved areas that will not present competition problems. These target areas can be found in 14 of the 16 districts with underserved areas.

Hence, the best locations for the opening of new supply points, if a firm is considering to successfully enter the private hospital health care market in continental Portugal as a new competitor, will be in these dark shaded areas, or in locations close to these areas. They are mostly less urbanized, rural or semirural areas. This is in accordance with the study of Huiban (2011), which indicates that survival rates of entrants are significantly higher in relatively less urbanized, more rural regions.

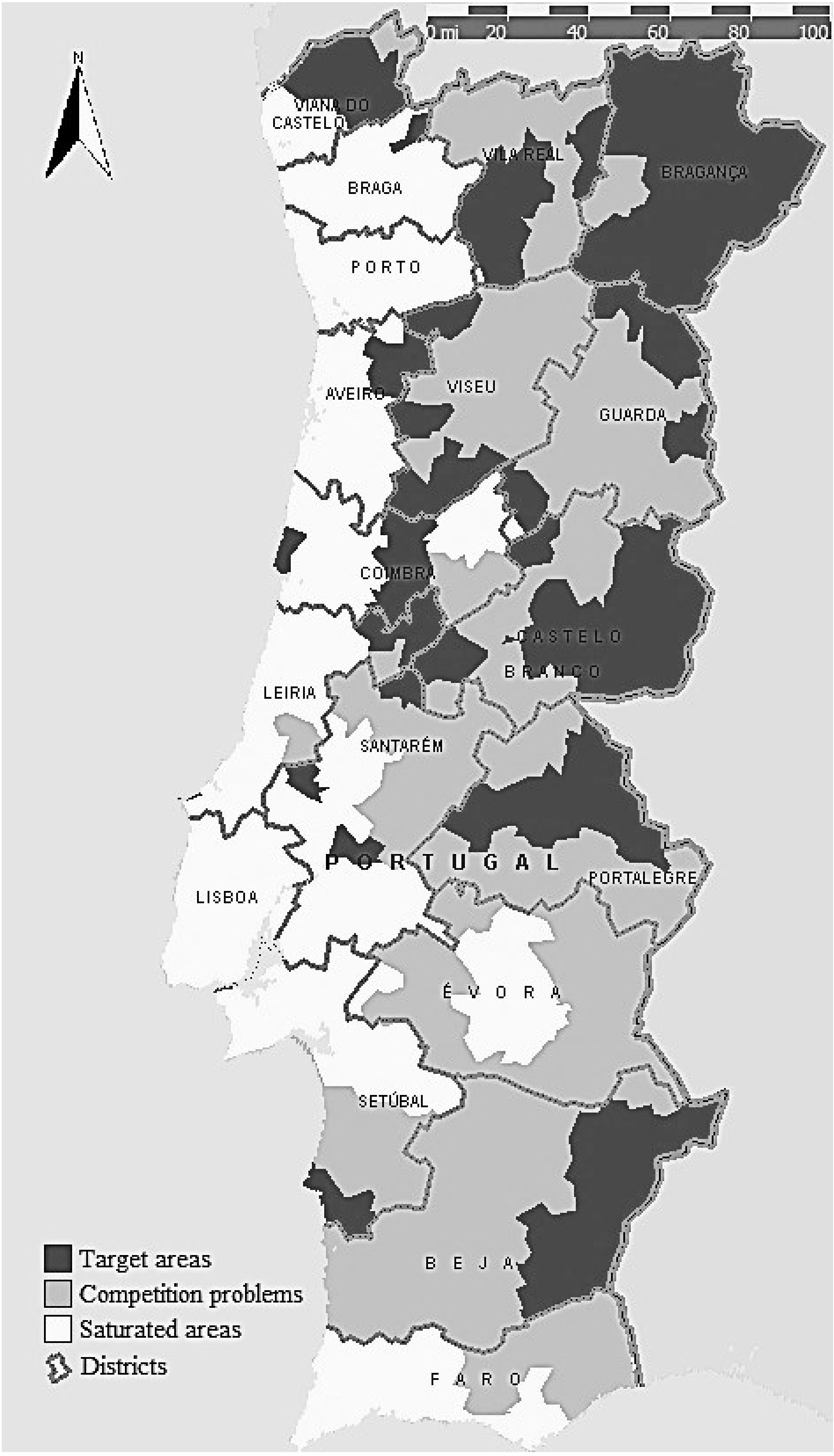

Figure 21 – Target hospital health care geographic markets for new entrants

These results can be useful for entrants when choosing a location for their new supply points, in order to better reach underserved markets, without the risk of having to deal with an adverse competitive situation that could render entry unsuccessful. The identification of specific target regions conveys an idea on where entrant firms should locate their new hospital so that it can maximize the respective target populations' coverage, given the projected investment. Accordingly, among alternative candidate hospitals' locations, the most adequate one can be chosen, for example, by considering its potential effects in terms of the target populations' coverage and their access scores increase. The greater and more widespread the access increase in the target region due to the new hospital, the higher the number of patients that the entrant will be able to attract and the higher will also be the entrant's market share.

It is important to conclude this case study, however, with the acknowledgement of a potential problem. Since our method identified some markets that are too small, relatively less urbanized or rural, and distant from a developed infrastructure, appropriate upstream suppliers and skilled labor, they are possibly not lucrative. Some examples of these geographic markets are the small target area in the district of Setúbal and the two isolated target regions in the district of Santarém, which can be identified in Figure 21. Notwithstanding, we note that entry in small rural or peripheral target markets can be regarded as strategic, in a two-stage entry strategy with a small initial investment, in which the entrant first enters a smaller target market to establish its brand and reputation, and then advances to compete stronger in other regions. Since reputation is an important competitive variable in health care markets (as seen in subsection 3.2.1), this could be a satisfactory entry strategy.

Also, it is important to note that the application of the proposed method to identify target markets clearly does not substitute the preparation of a thorough business plan that is needed to support entry decisions. A business plan will include the estimation of the potential demand, pricing, costs, the minimum profitable size and possible economies of scale, and the collection and analysis of information on the available local infrastructure, suppliers and labor.

4.4. Conclusions

This chapter presented a first full application of the proposed EKD4SFCA method, illustrating how it should be applied and the logic of its application, with the objective of identifying market entry targets. We explained how the results of the four steps of the method are assessed in an integrated way, in order to clarify how the method and its integrated assessment can be useful for firms.

However, we acknowledge that this is only one application of the method. As indicated in the previous chapter, the method can also be applied by public administrations, in the implementation of public policies destined to improve the distribution of supply resources, reduce access inequalities and improve the populations' health status. As remarked, public administrations can achieve this by exploring competition with public bids to contract private supply, in order to increase supply for disadvantaged populations, and at the same time define efficient prices for the contracted care. Alternatively, public administrations can direct public financing for entrant firms with free adhesion contracts, in order to increase supply and competition in the underserved regions with competition problems.

In the next chapter, the validity of the applications of the EKD4SFCA method and of their results is assessed with two validation exercises. The first exercise is an example of an assessment that can be useful for public administrations.

5. Validating the EKD4SFCA method

Having presented our proposed EKD4SFCA method in chapter 3, and a full application of it in chapter 4, in this chapter we focus on the validation of the method, with two validation exercises that assess the validity of applications of the method and their results, which we present in sections 5.1 and 5.2. As an introduction, we first discuss three common types of validity – (i) content, (ii) criterion, and (iii) construct – and we identify the specific type that is assessed with our two exercises.

Content validity is an important step in the development of a new method, because validating a method's content means linking abstract concepts with observable and measurable indicators (Haynes et al., 1995; Wynd et al., 2003; Yaghmaie, 2003). Addressing content validity is mostly achieved in the development phase of the method, with careful conceptualization and domain analysis (Polit and Beck, 2006). First the researcher needs to define the concept that he/she wants to measure (access or competition, in this thesis) and what is included in the definition (Mooi and Sarstedt, 2011). Subsequent efforts evaluate the relevance of the method's content (Polit and Beck, 2006). These efforts are mostly informal or subjective judgments based on literature review and expert review. In particular, the literature review of studies that are relevant to each component of the method plays an important role in guaranteeing content validity (OECD, 2006a).

Since the components of the EKD4SFCA method can be considered to be closely linked to the concepts that they are supposed to measure, with a theoretically sound basis, according to the literature review that we carried out, we understand that our method attains content validity. This conclusion is reinforced by the fact that we gathered feedback from experts in presentations of our work in national and international conferences, from the Portuguese Health Regulatory Agency, and from reviewers of our papers, and adjusted our work in accordance therewith.

Regarding criterion validity, a proposed method achieves criterion validity when its results correspond to the results of some other, superior method that accurately measures the concept or phenomenon of interest. This other method produces the criterion to be used as standard for comparison (Kaplan et al., 1976; McDowell, 2006).

It is mostly difficult to prove that the chosen criterion is itself valid. As Kane (2012) remarks, without some way to validate the criterion that does not involve another criterion, we face either infinite regress or circularity (Kane, 2012). Also, if the criterion is not a superior measure, then it can be flawed and insufficient as a reference for validity (Kaplan et al., 1976). When there is no criterion, it is not possible to assess criterion validity, because it will be insufficient to sustain any testing purpose (Messick, 1990).

We understand that it is not possible to assess the criterion validity of our EKD4SFCA method, because there is no criterion for it. To the best of our knowledge, there is no standard, superior measure for the identification of target regions based on an integrated assessment of access and competition, which our method enables. This shall be considered an evidence that the development of our method is justified, because, as remarked by Kaplan et al. (1976), the lack of a criterion is the first and foremost reason for developing a new method.

The alternative for assessing criterion validity is to assess construct validity, which is the degree to which the variance of the results of a method is consistent with predictions from the construct targeted by the method (Haynes et al., 1995).

Sensitivity to change analysis is the kind of analysis that can produce the most convincing evidence of construct validity (Burlingame et al., 2006; Vermeersch et al., 2000). The researcher is able to ascertain that a method attains construct validity, if he/she is able to show that the method detects the expected change in results, whether the change is induced experimentally by the researcher, or naturally (Hays and Hadorn, 1992).

Thus, the validation exercises that we present in this chapter are exercises that assess the sensitivity to change of our method. In these two exercises, we introduce changes to the data and assess how the method reacts to these changes, in order to confirm that it captures changes in the data as it is supposed to capture, considering how it is constructed, the information that it uses and its objectives.

5.1. Introducing a new market entrant with a new supply point

For the first validation exercise, we apply the EKD4SFCA method to data of the Portuguese long-term care sector, using the number of beds as the supply capacity variable (Polzin et al., 2014b). We use the same data set of August 2012 that was used in a study of the Portuguese Health Regulatory Agency, collected from its registry and from the public institution *Unidade de Missão para os Cuidados Continuados Integrados*. However, that study assessed only the patients' access to long-term care with an incomplete EKD2SFCA method. More specifically, a KD2SFCA method was adopted, complemented with the introduction of a health needs index (ERS, 2013).

In this section, we first apply our complete EKD4SFCA method, with the objective of identifying target regions for the financing of new long-term care supply points. After assessing its results, we introduce an artificial supply point in a target area, and reapply the method, to produce new results. Finally, we compare the results – before and after the new supply point –, assuming that no other changes occur. Hence, in this assessment, it is assumed that the only changes that occur are the effects on access and competition due to the introduction of a new supply point in the long-term care sector, owned by an entrant firm.

Before presenting the exercise, we first contextualize it, explaining what long-term care is, how it is organized in Portugal, characteristics of the demand and the supply.

5.1.1. Long-term care in Portugal

Long-term care can be defined as the set of sequential health or social support interventions, centered on the global recovery, understood as the active and continuous therapeutic and social support process (definition of the Law N. 52/2012, of 5 September). This kind of care is prescribed to promote autonomy by improving the functionality of the person that is in a situation of dependency, through his/her rehabilitation, re-adaptation and social and family reinsertion.

In continental Portugal, around 90% of all the people that are treated with long-term care, are treated in the supply points of the National Network of Integrated Long-Term Care (*Rede Nacional de Cuidados Continuados Integrados*, RNCCI) (ERS, 2011). The RNCCI was created in 2006 by the Decree-Law N. 101/2006, of 6 June, with the aim of making more dynamic the implementation of a financially sustainable supply, directed to the people in a dependency situation. This network includes contracted private and public providers and its objective is to improve access to technically and humanly adequate care provision of the citizens with loss of functionality, or in a risk situation of losing it.

Demand for long-term care in Portugal can be characterized by the most frequent demographic and socioeconomic patient profile that is treated in the RNCCI (ERS, 2013). Patients are mostly:

- Elderly;
- Women;
- Married or widows;
- People that attended only until six years of school;
- People that have worked as non-qualified workers;
- Dependent or incapable people in terms of physical autonomy;
- People supported by family, especially for hygiene and food;

- People that come from natural family or living alone; and
- People that have a medium or high risk of falling.

The strategic planning of the RNCCI was originally organized in three development phases. Phase 1, between 2006 and 2008, had the aim of covering 30% of the existing needs. The second phase, between 2009 and 2012, had the target of attaining 60% of coverage. The third phase, from 2013 and 2016, intended to attain the 100% coverage target, which shall be reached when the RNCCI has 15,308 beds in total (ERS, 2011). However, we note that in August 2012 there were still only 5,916 beds.

Thus, the development process of the network is still ongoing, such that our proposed EKD4SFCA method can be useful to identify underserved regions that should be prioritized, in order to improve the geographic distribution of supply along the process and better meet the populations' needs. Besides, since the EKD4SFCA method includes competition assessment, it can also indicate where the public financing could profit from public bids, to set efficient, lower prices for the contracted care. Until now, prices are administratively set (ERS, 2013).

While long-term care is provided by different kinds of supply units and teams, more specifically inpatient units, outpatient units, hospital teams and home visiting teams, we focus our assessment on the main long-term care supply of the RNCCI, which was constituted in August 2012 by 268 long-term care inpatient supply points. These supply points were geographically distributed as illustrated in Figure 22.

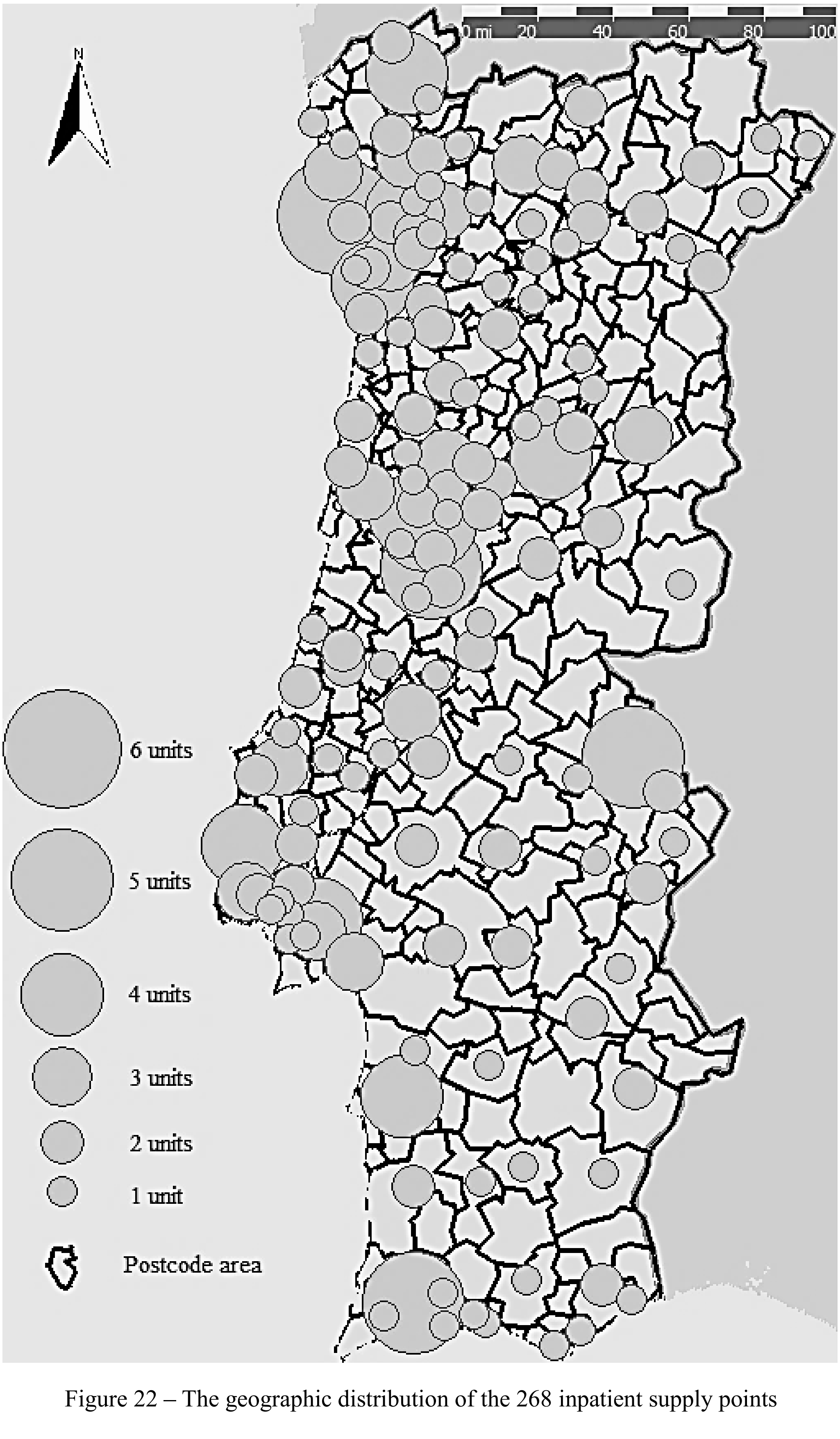

Figure 22 – The geographic distribution of the 268 inpatient supply points

5.1.2. Steps 1 and 2 of the EKD4SFCA method

First, we applied the EKD4SFCA method with the original long-term care data, prior to the inclusion of the new supply point. Regarding step 1 of the method, the health needs index was constructed considering variables that reflect the demand for long-term care, characterized by the most frequent patient profile described above. The following five variables were selected to construct the PCA index: proportion of people aged 65 years old and above, age dependency ratio (the relation between the people aged 65 and above and the people with ages between 15 and 64), female population aged 65 and above, the proportion of widowers, and the proportion of people aged 15 and above without any level of education. The results obtained with the five selected variables satisfied the necessary criteria to construct an index with PCA using the first component. In particular, we obtained one factor with an eigenvalue greater than one (4.254) with a high percentage of the total variance explained (85.077%), and an adequate value for the KMO statistic (0.798).

In order to rescale the standardized health needs principal component scores, we applied the thresholds presented in subsection 3.1.1.2. Accordingly, we converted the health needs standardized principal component scores from zero to one to the scale from one to 1.167, in a proportionate way, and the scores higher than one were transformed to values above 1.167, considering a linear extrapolation of the scale. The negative scores between -1 and zero were converted to a scale from 0.876 to one, and scores lower than -1 were rescaled to linearly defined values below 0.876. Applying these scale transformations, we obtained health needs scores varying from around 0.79 to 1.70.

Regarding the commuting index in step 2, we used the same commuting index that was used in the case study that was presented in section 4.1. Thus, the chosen variables were, again: the proportion of population that works or studies in other municipalities, the proportion of population with commuting of over one hour travel time to other municipalities, average commuting time, and the proportion of population that uses car, bus, subway or train as the main transport in commuting. Also, as indicated in subsection 3.1.1.3, the standardized scores were transformed to the scale from 0.7 to one.

Table 9 presents the distribution of the long-term care access scores calculated with steps 1 and 2 of the EKD4SFCA method across the three levels: low, medium and high. We used 60 minutes as the maximum travel time threshold, since this travel time is considered to be a reasonable maximum reference for palliative care, which is the kind of long-term care that values more proximity (Cinnamon et al., 2008; Schuurman, Crooks, et al., 2010). Also, this is the same reference that was used by the Portuguese Health Regulatory Agency in its access study that was carried out with the same data set (ERS, 2013).

Table 9 – Distribution of the long-term care access scores across the three access levels

Statistic	**Cluster**		
	Low	**Medium**	**High**
Number of areas	100	236	124
Population covered (%)	23.1	60.1	16.8
Mean access score	0.22	0.49	1.31

The two bounds of the three access levels defined by the k-means clustering algorithm are 0.31 and 0.77 beds per 1,000 inhabitants. Hence, for scores until 0.31 we identify a low access level, for scores until 0.77 we identify medium access, and scores above 0.77 are high access scores. As seen in Table 9, there are 100 low access postcode areas with 23.1% of all population, with a mean access score of 0.22 beds per 1,000 inhabitants (as a reference for comparison, we note that the targeted ratio of the RNCCI is 1.82 beds per 1,000 inhabitants).

5.1.3. Steps 3 and 4 of the EKD4SFCA method

In steps 3 and 4 we used the number of beds of the firms that own the supply points to calculate the market shares. It was possible to identify 309 postcode areas with high market concentration (HHI_i^E >2000), and 279 of them were dominated, including 44 of the 100 low access areas, as seen in Table 10. The low access areas are the selected areas to identify target markets, namely the low access areas where there is more competition between the long-term care providers.

Table 10 – Results of the EKD4SFCA method

Competition assessment	Low access areas	Pop. covered (%)
Areas with HHI>2000 and dominated areas	44	6.3
Target markets	56	16.8
Total	100	23.1

These low access areas with more competition are the 56 postcode areas identified in Table 10, which constitute the target markets for public administrations to promote financing for new supply points. These targets represent 12.2% of the 460 postcode areas, in 14 of the 18 districts of continental Portugal, and encompass a population that corresponds to 16.8% of the total population (almost 1,692,000 residents).

5.1.4. Results of the EKD4SFCA method

Figure 23 shows the results obtained with the EKD4SFCA method.

5. Validating the EKD4SFCA method

5.1. Introducing a new market entrant with a new supply point

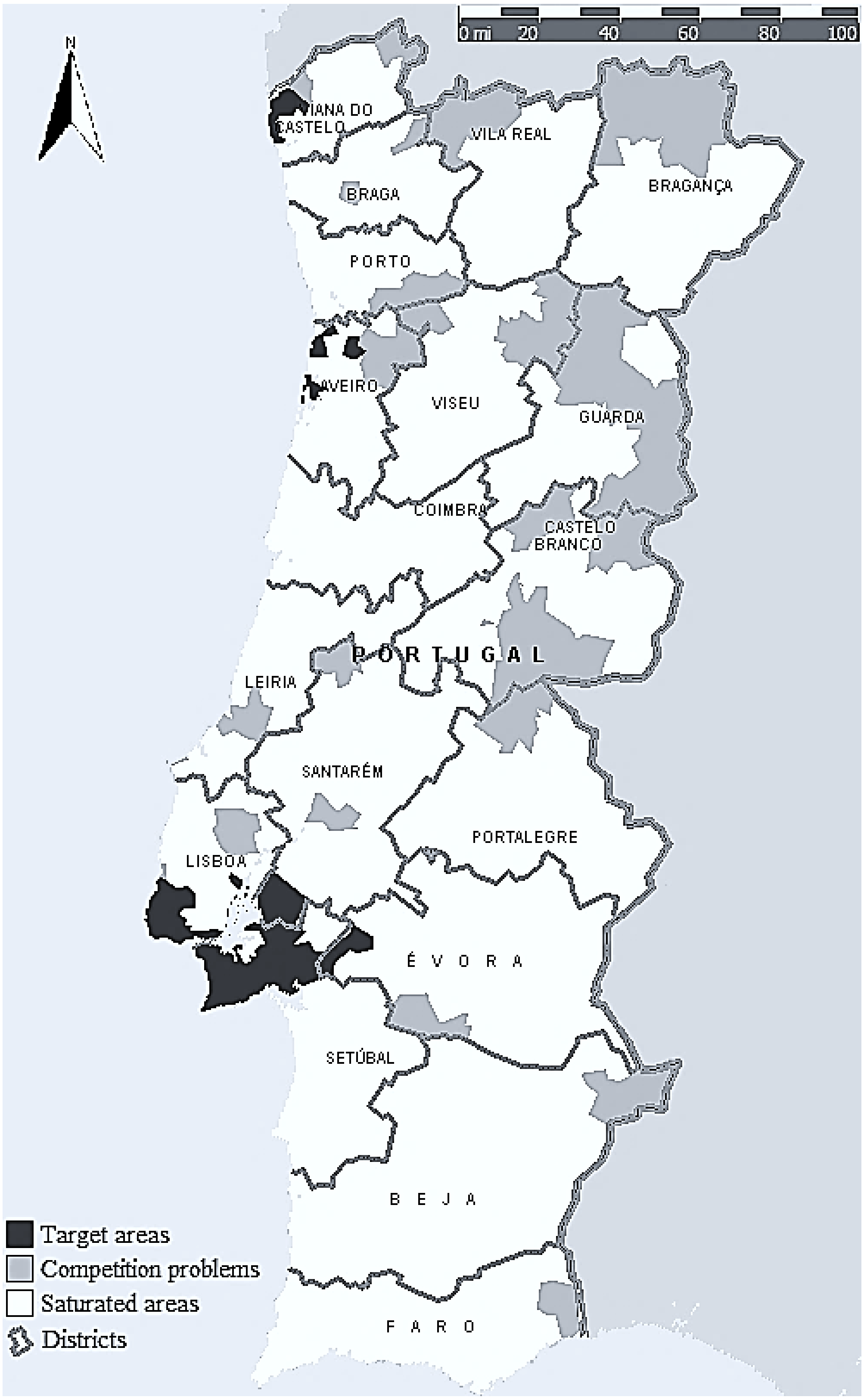

Figure 23 – Target geographic long-term care markets for promoting new financing

The dark shaded regions in Figure 23 are the target markets, the light shaded regions are the relatively underserved regions with potential competition problems, and the white regions are the regions where no disadvantaged populations reside. The target markets can be seen as target regions where public administrations can promote public biddings for contracting new providers and beds, i.e., new supply capacity for the disadvantaged, low access populations. Thus, the application of the EKD4SFCA method identifies where public administrations can explore competition between firms to obtain lower, more efficient publicly financed prices, while at the same time promoting new supply to the disadvantaged populations.

We note that public administrations can also target the light shaded regions, namely the low access regions with high risk of competition problems, and promote supply therein. Administrations can stimulate the creation of new firms to serve these regions by financing additional supply, offering free adhesion public contracts for new firms, with supply points that cover these areas. Public biddings would not work in these areas, because there is no competition to be explored to obtain efficient prices. However, free adhesion contracts for new firms that serve the populations in the light shaded regions can promote an increase in the access of their residents, reducing inequity in the development process of the RNCCI, and, at the same time, promote competition, by stimulating entry of new firms in regions with potential competition problems.

5.1.5. The introduction of a new supply point

For the validation exercise, to infer about the construct validity of the EKD4SFCA method, we introduce a market entrant with an artificial supply point with a supply capacity of 40 beds in a target area of the Évora district, and assess the changes that are produced to the results. The new supply point is introduced in a postcode area where previously there was not any supply point.

We can expect the following changes in the region where the new supply point will be active, considering the theoretical reasoning behind access and competition assessments:

(i) Access of the residents will increase; and

(ii) Competition in most areas will increase, but not in all areas: it will depend on the net result of two opposing effects.

The access is expected to increase, because more supply will become available to the populations of the region: there will be more available supply capacity in an additional supply point.

Regarding competition, the introduction of the new firm is expected to increase competition, because the number of firms competing in the region increases. In fact, with an increasing number of competing firms, the market structure moves in the opposite direction of the monopoly structure, in which the market has the highest concentration, with only one firm that dominates the market. However, depending on the supply capacity of the new supply point and on the competitive situation that existed in some areas, inside the region, prior to the entry of the new firm, competition is expected to become weaker. For example, if the competitive situation in an area is competitive, with low market concentration and no dominance, and the capacity of 40 beds is too high in relation to the capacities of the competing firms, we can expect that that area will become highly concentrated and dominated by the new entrant. Likewise, if the area of the new supply point is isolated, with weak competitive pressure exerted by other, distant firms, we can also expect that that area, and possibly the areas nearby, will suffer a decrease in competition, with the new firm increasing market concentration and probably becoming dominant.

We note that these changes are expected to be confined to the region where the new supply point will be active. Outside that region, no changes in access and competition are expected to occur, because patients in those areas will be too distant. They will not be traveling to the new supply point to obtain care, and other firms will be competing for those patients.

Also, we note that these are the only rational access and competition changes that are expected to occur. This means that, in our sensitivity to change analysis, we have to show that our method captures only changes that are in accordance with these expectations. If the method captures other changes, than its application and its results cannot be considered to attain construct validity.

We present next the results of our method, after the introduction of the new supply point that we carried out, and the sensitivity to change analysis. We begin presenting the access changes and conclude with the assessment of the changes in competition.

5.1.5.1. Access changes

After the introduction of the new supply point, the access scores increase in the 50 postcode areas that form the catchment area of the new supply point, across the districts of Évora, Santarém, Lisboa and Setúbal, as shown in Figure 24. These access changes occur as expected.

Figure 24 presents in grey the span of the catchment area of the new supply point, which was introduced in postcode area 7080. We remark that, even though postcode area 7050 is a neighboring area of the new supply point's postcode area 7080, that area is not covered by the 60 minutes catchment area's range (as we checked, the GIS application indicates that travel time between the centroid from postcode area 7080 until that adjacent area's centroid is one hour and four minutes).

The highest access increase occurs in postcode area 7080, where the new supply point is located, and it is an increase of around 0.46 beds per 1,000 inhabitants, from 0.15 to 0.61 beds per 1,000 inhabitants, approximately. As we identified, the higher the distance of the population from the new supply point, the less its access score increases due to the new supply point, and this is in accordance with the influence of the distance decay on access – duly captured by the distance decay function in steps 1 and 2 of the EKD4SFCA method.

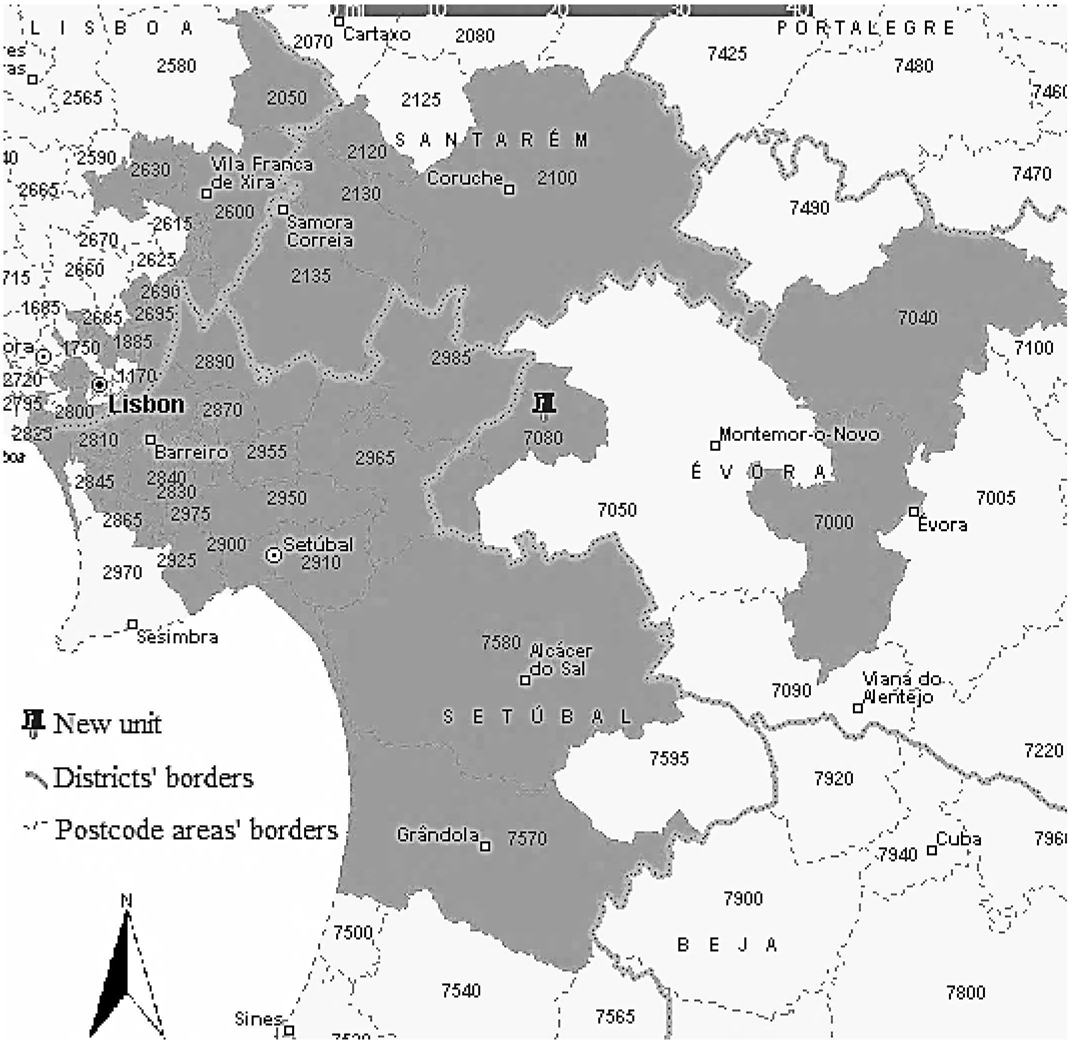

Figure 24 – Span of the new supply point's catchment area

Figure 25 and Figure 26 present the results of the assessment of the patients' access to long-term care in continental Portugal – before and after the introduction of the new supply point in the postcode area 7080.

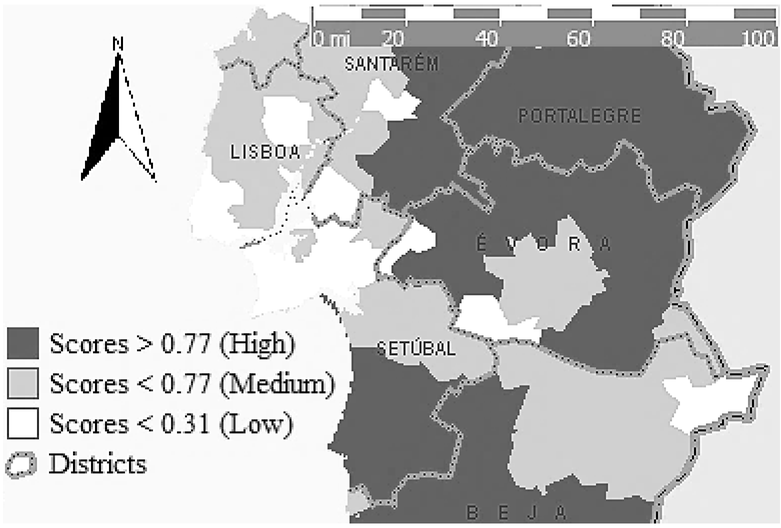

Figure 25 – Access levels before the new firm's entry

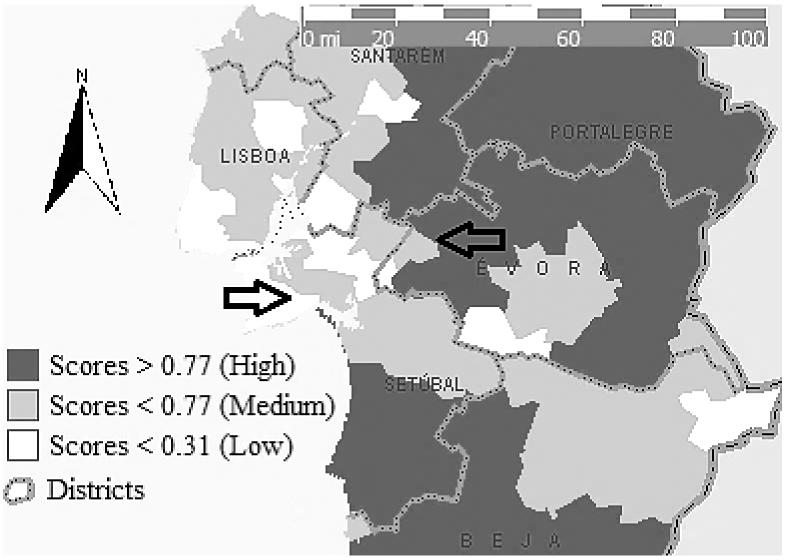

Figure 26 – Access levels after the new firm's entry

The changes are slight. In fact, only a few regions in the districts of Évora and Setúbal change from white to light shaded, because their access level changes from low to medium access (more specifically in the east of the district of Évora and in the north of the district of Setúbal, as the two arrows in Figure 26 indicate).

These regions are formed by six postcode areas: 7080, 2950, 2910, 2830, 2840 and 2800. The changes of the scores of these postcode areas and the identification of their corresponding district are presented in Table 11. As we recall from subsection 5.1.2, the threshold that separates the low access level from the medium access level is 0.31 beds per 1,000 inhabitants.

Table 11 – Low to medium access score changes (beds per 1,000 inhabitants)

Postcode areas	**District**	**Access scores**	
		Before	**After**
7080	Évora	0.154	0.612
2950	Setúbal	0.243	0.353
2910	Setúbal	0.231	0.311
2830	Setúbal	0.287	0.338
2840	Setúbal	0.307	0.358
2800	Setúbal	0.298	0.312

Considering these results, we verify that the method produces expected changes, such that it behaves as it should to measure the change to the potential access after the introduction of the new supply point. Because of the added supply capacity that becomes available to the populations residing in the affected areas, they have their access scores increased, such that some of them even have their access level increased. In particular, the change of access level in postcode area 7080 means that this area stops being a target area. This change is confirmed in the end of this validation exercise, after the assessment of the changes to the competitive situation.

5.1.5.2. Competition changes

As we could verify, there was not any supply point in postcode area 7080, in the Alentejo region, prior to the new firm's entry, and also no supply points nearby. Thus, it is expected that the new firm shall face weak competition from other firms after entry. This has to be confirmed with our assessment of the competition changes after the new firm's entry, which we now present.

Table 12 presents the firms that, before the introduction of the new supply point, compete for the residents of postcode area 7080, as potential patients. We identify the supply points that they own, the total number of beds of the supply points, and the postcode areas and districts of their locations.

Table 12 – Competing firms serving postcode area 7080

Supply point	Number of beds	Firm	Postcode area	District
Associação de Intervenção Social de Grândola	20	AISGRA	7570	Setúbal
Associação de Reformados e Idosos da Freguesia da Amora	30	ARIFA	2845	Setúbal
Centro Hospitalar Barreiro-Montijo	10	CHBM	2830	Setúbal
Hospital Residencial do Mar	38	Espírito Santo Saúde	2695	Lisboa
Hospital do Espírito Santo de Évora	19	HESE	7000	Évora
Centro Comunitário du Bocage	24	LATI	2910	Setúbal
Residências Montepio, Serviços de Saúde	119	Montepio Residências	2870	Setúbal
Rollar - Alojamento para Idosos	60	Rollar	2910	Setúbal
SCM Évora	12	UMP	7000	Évora
SCM Amadora	30	UMP	2610	Lisboa
SCM Arruda dos Vinhos	15	UMP	2630	Lisboa
SCM Coruche	15	UMP	2100	Santarém
SCM Alhos Vedros	45	UMP	2860	Setúbal

The areas that present changes in competition are the same 50 postcode areas that have their access scores changed, and which form the catchment area's span of the new supply point (illustrated in Figure 24).

Appendix C presents the HHI changes, and identifies the highest market shares and dominance thresholds in these 50 postcode areas, after the introduction of the new supply point. As shown, competition increases in 49 of the 50 postcode areas, as indicated by the HHI decrease in these areas. Only in the postcode area 7080, where the new supply point is introduced, there is an increase in the HHI. Thus, the market in that area becomes more concentrated, which indicates a weakening of the competition between the firms that own the long-term care supply points that serve that area.

The most significant change occurs in the dominated postcode area 2100, in the south of district of Santarém (see Figure 24), with a sharp decrease of more than 2500 points in the HHI value, from 9368 to 6811 (see Appendix C). Another substantial change, which alters the market concentration level from high to medium, occurs in postcode area 2985, located in the northeast of the Setúbal district (see Figure 24), with a HHI decrease from 2428 to 1974 points (see Appendix C). However, these two substantial changes in postcode areas 2100 and 2985 do not imply any change in the identification of the target regions: postcode area 2100 keeps being highly concentrated and dominated, and postcode 2985 is not dominated before the introduction of the new unit and remains not dominated thereafter.

The major change with implication to the identification of the target regions occurs in postcode area 7080, with the market remaining highly concentrated, but with the highest market share – of the new firm – becoming higher than the dominance threshold. The HHI increases from 2253 to 3038 points, the highest market share increases from 31.7%, of an incumbent firm, prior to the new firm's entry, to 49.7%, of the new entrant firm, and the dominance threshold decreases from 41.7% to 36.4%.

5.1.5.3. Final results

As noted in the beginning of this validation exercise, the expected changes, due to the introduction of a new entrant's supply point, were mainly: (i) an increase in the access of the residents of the region where the new firm will be active, i.e., the affected region; and (ii) a strengthening in competition in most areas of the affected region.

As seen in our validation exercise, the affected region is the region covered by the catchment area of the new supply point, because people outside the catchment reside too far. Their access does not change, because they will not be willing to travel to the new supply point, due to its excessive distance. Competition outside the catchment area also does not change, because the firms that compete for the patients that reside outside the catchment are not influenced by any competitive pressure of the new firm in those markets. In fact, the new entrant will compete only for the patients that reside inside the affected region. Thus, this first finding confirms the expected result that changes occur only in the affected region: the 50 postcode areas that constitute this region.

Regarding the changes to the patients' access to long-term care, we were able to confirm with our validation exercise that, with the introduction of a new supply point, our proposed EKD4SFCA method captures increases in all the 50 postcode areas of the affected region, as expected.

Finally, regarding the changes to the competition, we were also able to identify changes as expected: in most areas of the affected region, competition was strengthened by the new entry. In fact, we identified that 49 of the 50 areas presented a decrease in the market concentration, with some substantial changes. The major change occurred, though, in the postcode area in which the new supply point was introduced, where competition weakened. This occurred also as expected, because the market was competitive and there was no supply point in that area, or nearby, prior to the new firm's entry. Accordingly, the results indicated that the new firm faces weak competition after entry.

To conclude, we present the new target regions in Figure 27.

5. Validating the EKD4SFCA method

5.1. Introducing a new market entrant with a new supply point

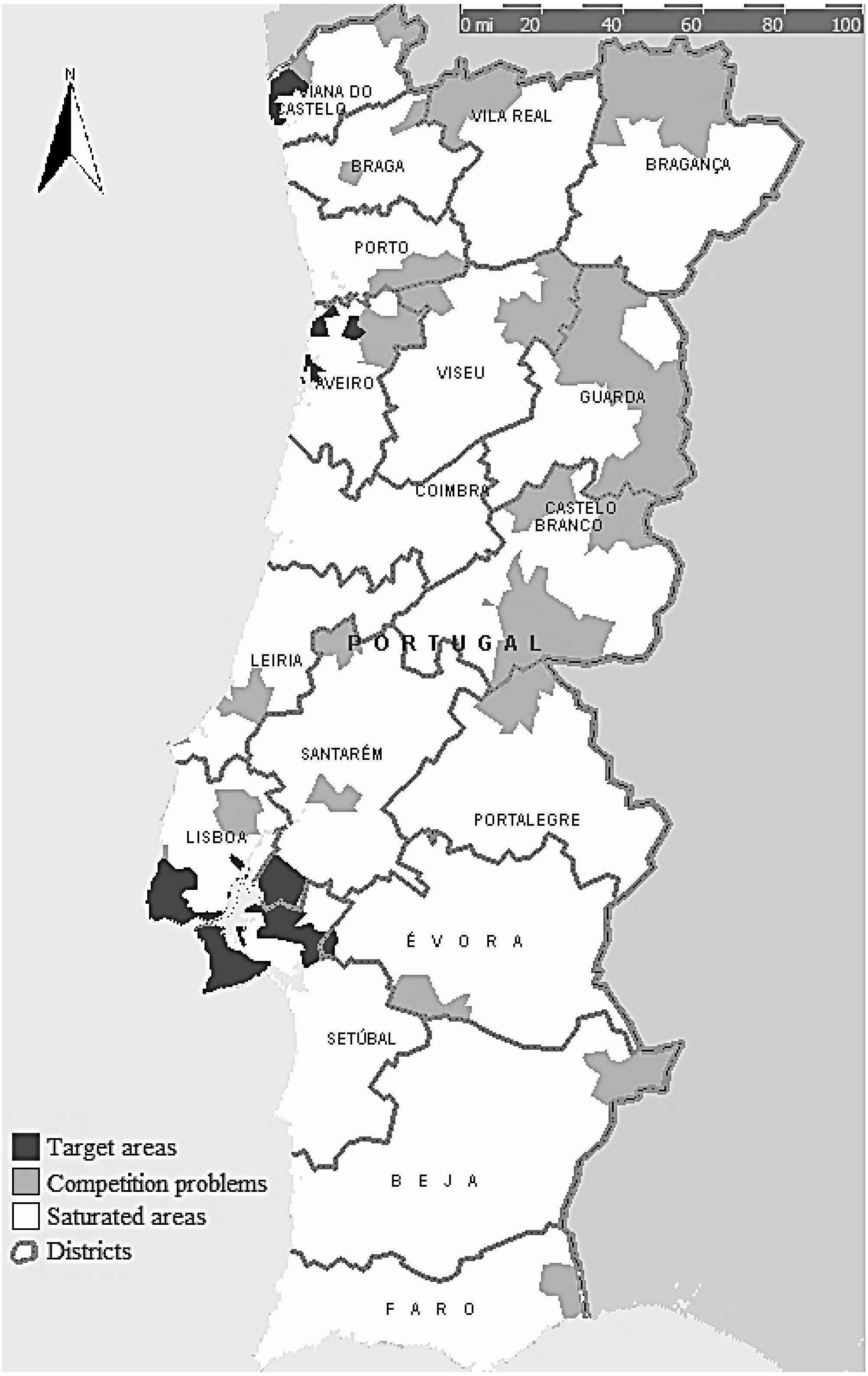

Figure 27 – New target regions after the introduction of the new firm's supply point

Comparing Figure 27 with Figure 23 (the situation *ex ante*, i.e., before the introduction of the supply point), we can see that small changes occur to the target regions in the districts of Évora and Setúbal, more specifically in the east of the district of Évora and in the north of the district of Setúbal. As can be seen, the target regions become smaller, because six target postcode areas lose their status as targets and, accordingly, change from dark shaded to white: the same six areas presented in Table 11. These changes are due to the changes in their access level, from low to medium, while in one postcode area, where the new supply point is introduced, the change is also due to the fact that it becomes dominated and highly concentrated, i.e., with potential competition problems.

Thus, summing up the changes, besides confirming the expected results with this validation exercise, we also conclude that the proposed method is useful to promote access and competition, since the net results of the introduction of a new supply point in a target area can be regarded as positive. The introduction of the new inpatient unit in the target postcode area 7080 improves access of the populations residing in the 50 postcode areas that are covered by the catchment area, producing six access level changes, and it strengthens competition in 49 areas. Even though the competitive situation deteriorates in the postcode area where the new firm enters the market, that area benefits from a significant increase in access: it is one of the areas where access changes from low to medium. This means that the population that resides in that area stops being a disadvantaged population, which shall compensate the risk of occurrence of competition problems in that small area.

5.2. Creating access barriers in the south of Portugal

For our second exercise to validate the application of the EKD4SFCA method and its results, we apply the method again with the data set that was used in chapter 4, from the Portuguese hospital health care sector. The change that we introduce to analyze the method’s sensitivity to change is the creation of access barriers, isolating the southwestern region of Portugal to access from patients that reside in the southeastern region, and vice versa. These access barriers could be caused, in real life, by a roads blockage, for example, as a result of a natural catastrophe, like an earthquake, fire or floods. They could also occur due to the destruction of the roads that separate west from east in the south of Portugal by rebels, protesters or terrorists in a civil war, or the introduction of extremely high tolls in these roads, which would make it far too costly for people to travel between the two regions.

5.2.1. The creation of the access barriers

With the change that we carry out, the three supply points located on the west side of the Faro district become inaccessible to patients that reside in the east side, and the two supply points located on the east side of Faro become inaccessible to patients that reside in the west side. Table 13 presents the identification of the five hospitals: their designation, the number of physicians in each hospital, the firms that own them, and their location.

Table 13 – Hospitals in the Faro district

Hospital	Number of physicians	Firm	Postcode area	District
Hospital de Santa Maria de Faro	67	HPP Saúde/CGD	8000	Faro (east)
HPA Unidade de Faro	8	Hospital Particular do Algarve	8005	Faro (east)
Hospital de São Camilo	12	Hospital Particular do Algarve	8500	Faro (west)
HPA Unidade de Alvor	56	Hospital Particular do Algarve	8500	Faro (west)
Hospital de São Gonçalo de Lagos	39	HPP Saúde/CGD	8600	Faro (west)

Considering the theoretical reasoning behind access and competition assessments, we can expect the following changes in the affected region:

(i) Access can increase or decrease, depending on the net result of two opposing effects;

(ii) Competition can strengthen or weaken, such that it also depends on the net result of two opposing effects.

Regarding access, we can expect that it will increase, due to the fact that fewer patients will compete for the supply in each side of the region. The creation of access barriers in our exercise is equivalent to the delineation of an imaginary border between the two sides of the southern region of Portugal, with no trespassing allowed, such that the supply in each side becomes available only to patients from the same side. However, after the creation of the barriers, it can also be expected that access will decrease, because fewer hospitals and less supply capacity become available to the patients that reside in each side: they will not be able to travel to the hospitals of the other side anymore.

As regards competition, the creation of the access barriers impedes competition between the firms with supply points located on one side of the southern region of Portugal for patients of the other side. Accordingly, we can expect that competition will become weaker. However, we can expect that competition will strengthen, if the affected region is highly concentrated and dominated by a firm with its supply located on one side. In that case, the access barriers can reduce the dominant span, isolating it to just one side.

5.2.1.1. Access changes

Considering how our EKD4SFCA method was constructed, the access barriers shorten the catchment areas of the five hospitals identified in Table 13, since, as mentioned above, the access barriers are meant to be equivalent to an imaginary border with no trespassing. Figure 28 shows this imaginary border, which obstructs access that would be otherwise possible until the 90 minutes maximum travel times:

- From the western populations of postcode areas 7665, 7670, 8300, 8365, 8375, 8400, 8500, 8550, 8600, 8650 and 8670 to the hospitals of the east side of the Faro district; and
- From the eastern populations of postcode areas 7700, 7780, 8000, 8005, 8100, 8125, 8135, 8150, 8200, 8700, 8800, 8900 and 8950 to the hospitals of the west side of the Faro district.

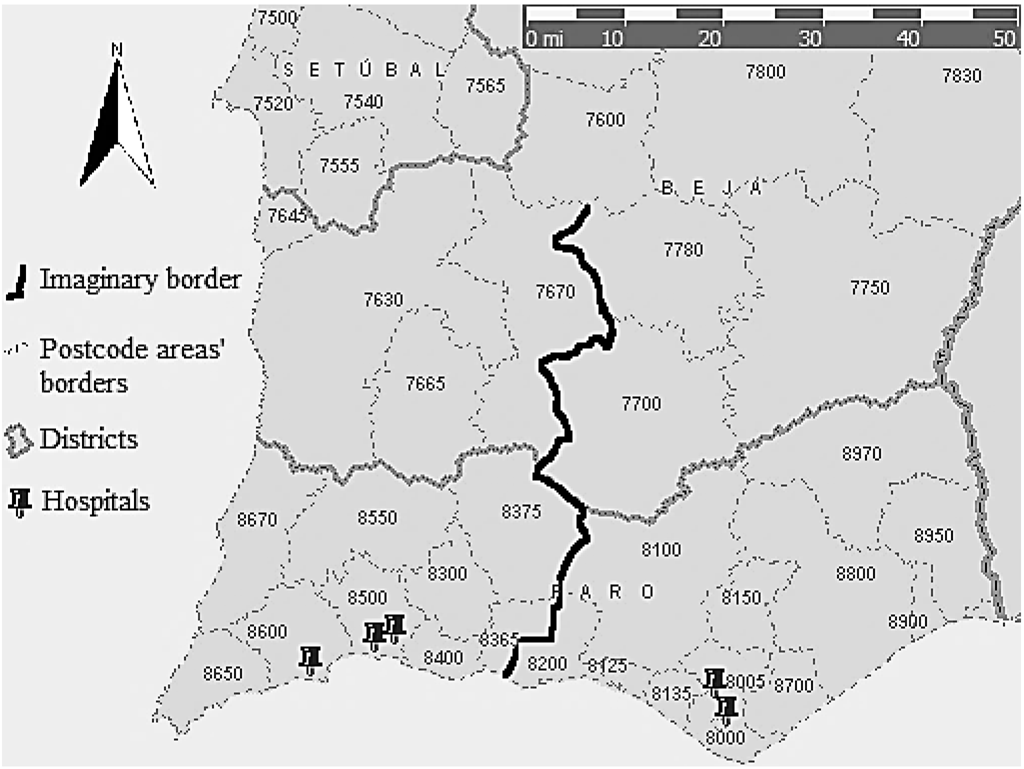

Figure 28 – The imaginary border that imposes access barriers

The shortening of catchment areas imposed by the imaginary border produces the two opposing effects in access described before, which are duly captured by our method's step 1 and 2. As a matter of fact, the increasing effect on access is captured primarily by the step 1's supply-to-demand ratio (see equation (14), in subsection 3.1.1.2), because the population size in the denominator of each ratio that is calculated is diminished with the shortening of the catchment areas. In turn, with this increasing effect, the access scores computed with step 2 also increase (see equation (15), in subsection 3.1.1.3).

Regarding the decreasing effect on access, this is captured only in the computation of the access scores, with step 2. With the shortening of the catchment areas, a smaller number of step 1's ratios is included in the step 2's sum, and this reflects appropriately the fact that fewer hospitals and less supply capacity becomes available to the affected populations.

In total, we are able to identify with the first two steps of our EKD4SFCA method that access changes occur in 29 postcode areas, in the districts of Faro, Beja and Setúbal, namely 17 access increases and 12 access decreases. Even though we identify more postcode areas with access increases, three of the areas where access decreases are the only ones that suffer changes in the access levels, from medium to low.

Table 14 lists all the 29 postcode areas with access scores changes, including the three ones with access levels changes, from medium to low (8125, 8135 and 8200) (the two bounds of the three access levels defined by the k-means clustering algorithm are 32.0 and 87.9 physicians per 100,000 inhabitants). In the table, it is possible to identify the postcode areas, the district of each postcode area, the access scores – before and after the creation of the access barriers –, the sign of each change (positive or negative), and each access level after the access barriers. These 29 postcode areas are illustrated in Figure 28.

The access level change from medium to low of the three postcode areas can be seen in the comparison between Figure 29, which presents the Faro district's access levels before the creation of the access barriers, and Figure 30, which shows Faro's access levels thereafter. The three postcode areas form a southern coastal region in the middle of the Faro district, which changes from light shaded to white, and these are the only visible access changes.

Table 14 – Medium to low access level changes

Postcode areas	District	Access scores		Change	Access level
		Before	After		
7600	Beja	0.12	0.16	+	Low
7630	Beja	5.23	8.67	+	Low
7645	Beja	1.90	3.16	+	Low
7665	Beja	21.20	29.70	+	Low
7670	Beja	10.83	10.69	-	Low
7700	Beja	13.95	7.38	-	Low
7780	Beja	5.89	3.25	-	Low
8000	Faro (east)	46.34	38.97	-	Medium
8005	Faro (east)	53.91	39.76	-	Medium
8100	Faro (east)	22.95	16.60	-	Low
8125	Faro (east)	43.09	27.08	-	Low
8135	Faro (east)	42.51	29.77	-	Low
8150	Faro (east)	37.37	33.28	-	Medium
8200	Faro (east)	51.31	24.59	-	Low
8700	Faro (east)	50.28	39.63	-	Medium
8800	Faro (east)	17.60	19.87	+	Low
8900	Faro (east)	17.94	19.44	+	Low
8950	Faro (east)	10.90	14.30	+	Low
8970	Faro (east)	3.16	4.16	+	Low
8300	Faro (west)	44.41	58.13	+	Medium
8365	Faro (west)	61.44	68.39	+	Medium
8375	Faro (west)	42.13	43.60	+	Medium
8400	Faro (west)	58.52	71.32	+	Medium
8500	Faro (west)	55.71	74.88	+	Medium
8550	Faro (west)	43.67	64.48	+	Medium
8600	Faro (west)	52.03	69.52	+	Medium
8650	Faro (west)	38.14	58.68	+	Medium
8670	Faro (west)	38.90	59.39	+	Medium
7565	Setúbal	1.56	1.49	-	Low

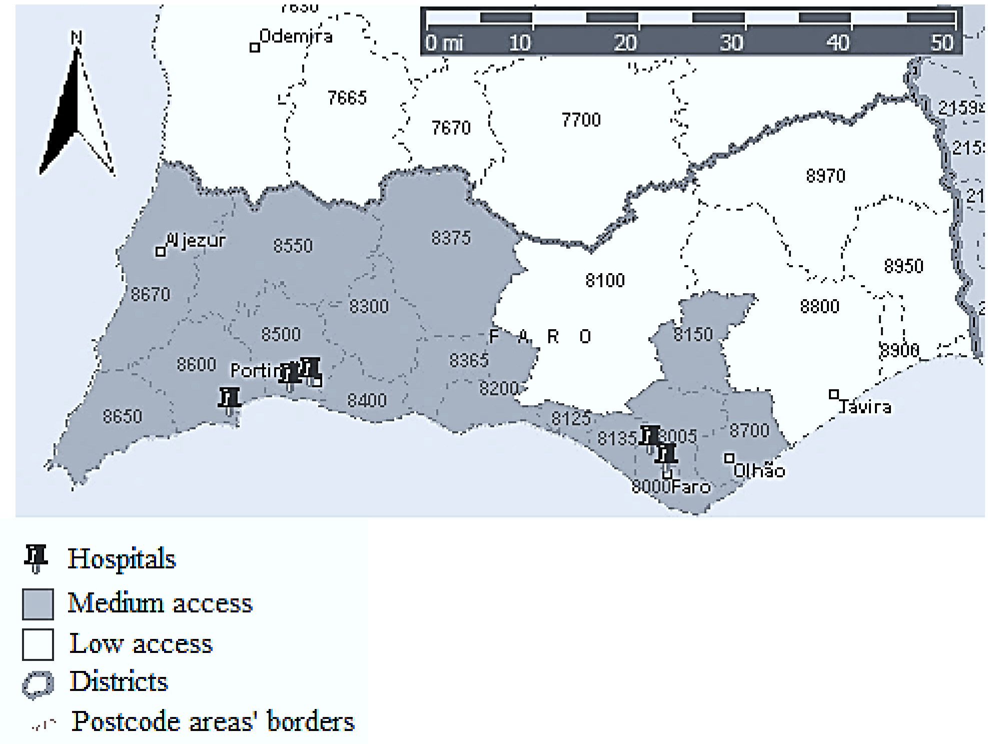

Figure 29 – Access levels before the access barriers

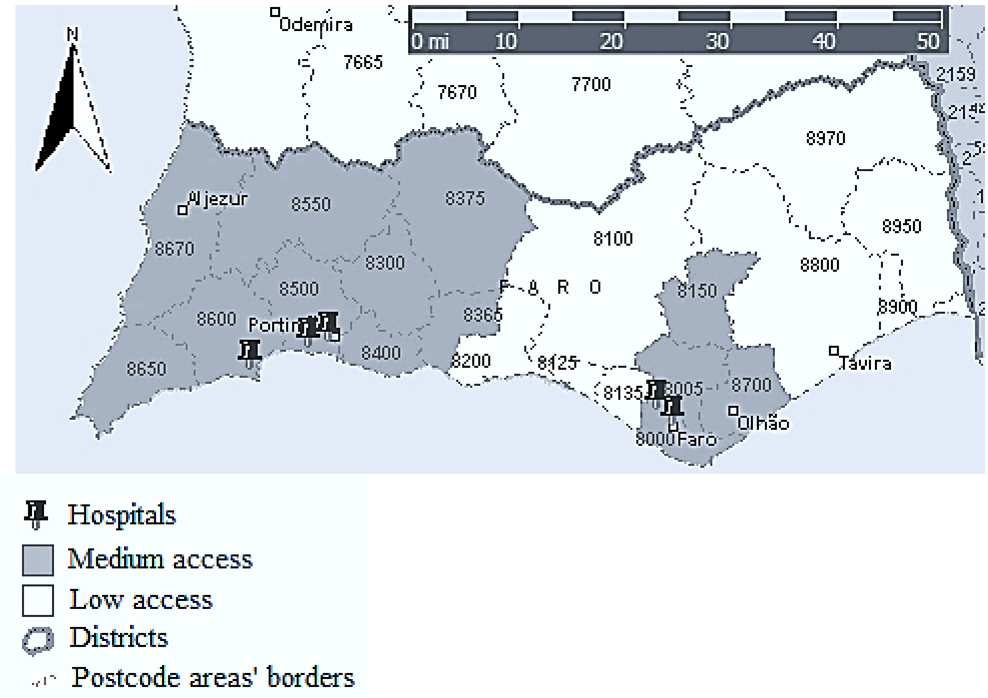

Figure 30 – Access levels after the access barriers

Two kinds of changes are now examined in more details, in order to demonstrate that the changes captured by our method are captured appropriately:

(i) The most relevant access scores' decreases, i.e., the access level changes from medium to low access;

(ii) The most significant access scores' increases identified in Table 14.

(i) The most relevant access scores' decreases

As indicated above, three postcode areas have their access level changed from medium to low after the creation of access barriers: 8125, 8135 and 8200. They are located in the eastern side of the Faro district, i.e., to the east of the imaginary border. In fact, most of the negative access changes occur in the eastern side, because there is more supply capacity available in the west, both in terms of number of supply points and total number of physicians. Besides, the population in the east is greater than the population in the west, such that the relation between supply and demand decreases in the east, after the creation of the access barriers, with much less supply becoming available to the relatively greater potential demand. Table 15 shows the information on the supply and on the population of each side of the Faro district.

Table 15 – Differences between west and east of the Faro district

Supply and population	**Faro district**	
	Faro (west)	**Faro (east)**
Number of supply points	3	2
Number of physicians	107	75
Population	163,951	287,055

As we could verify, the decrease in the supply-to-demand relation occurs especially in the postcode areas that are closer to the imaginary border, which include postcode areas 8125, 8135 and 8200. These are the areas that are more affected by the access barriers, because the populations that reside there are the ones that are closer to the western hospitals, which concentrate the largest supply capacity. As such, they are the ones from the east side that could benefit more from access to these hospitals, prior to the creation of the access barriers.

Still, we note that, regarding the eastern side of Faro, even though the majority of the eastern Faro postcode areas have their scores decreased (eight out of 12), there are a few postcode areas in the far east of Faro that present an access improvement (8800, 8900, 8950 and 8970). The populations residing at the farer east of Faro gain small increases in their access scores, because the effect from the western hospitals on their access before the barriers is very small, due to the long distances and the distance decay. This means that the negative effect on access that is caused by the access barriers to these populations is very small. In fact, it is more than compensated by the positive effect of the shortened catchment areas of the eastern hospitals, blocking access from the western populations, such that the net result is an access increase.

(ii) The most significant access scores' increases

As we could verify, the highest increases in the access scores computed with steps 1 and 2 our EKD4SFCA method are identified in the far west of Faro. There are increases of over 15 physicians per 100,000 inhabitants in the access scores of postcode areas 8500, 8550, 8600, 8650 and 8670. These are the western populations who depend more exclusively on the closest available, western supply, prior to the creation of the access barriers, and who benefit more from the barriers imposed to the eastern population. Thus, these high access increases in the far west of Faro occur, because, with the creation of the access barriers, their residents have a higher availability of supply, since the eastern patients stop competing with them for that supply. Besides, these changes are amplified, because, as seen in Table 15, the western supply capacity is higher than the eastern supply capacity, and the eastern population is larger than the western population.

Finally, we present in Table 16 the positive effects on access created by the shortening of the hospitals' catchment areas and the reduction of competition between the populations for the available supply. These effects are captured with the step 1 supply-to-demand ratio of our method and they explain all positive changes to the access scores. When only these effects occur, or when they more than compensate the negative effect of the shortening of the catchment areas in the calculation of the step 2 scores, then the final access change is positive.

Table 16 – Step 1 ratios (physicians per 100,000 inhabitants): before and after the creation of access barriers

Hospital	Step 1 ratios	
	Before	After
Hospital de Santa Maria de Faro	31.88	41.54
Hospital de São Camilo	6.42	10.70
Hospital de São Gonçalo de Lagos	20.86	34.56
HPA Unidade de Alvor	29.97	49.92
HPA Unidade de Faro	3.23	4.44

5.2.1.2. Competition changes

As far as the competition changes are concerned, we recall that the creation of the access barriers can produce also opposing effects in terms of competition. However, before we present the sensitivity to change analysis, we need to first clarify that, from the 29 postcode areas that have their access scores changed due to the creation of the access barriers, four of them are affected only because of the effect of the step 1 ratios on the computation of the access scores. Thus, the residents of these four postcode areas are not directly affected by the imaginary border, such that they can travel until the maximum travel time of 90 minutes to reach the available hospitals without any restriction. Figure 31 illustrates these postcode areas: 7600, 7630, 7645 and 8970.

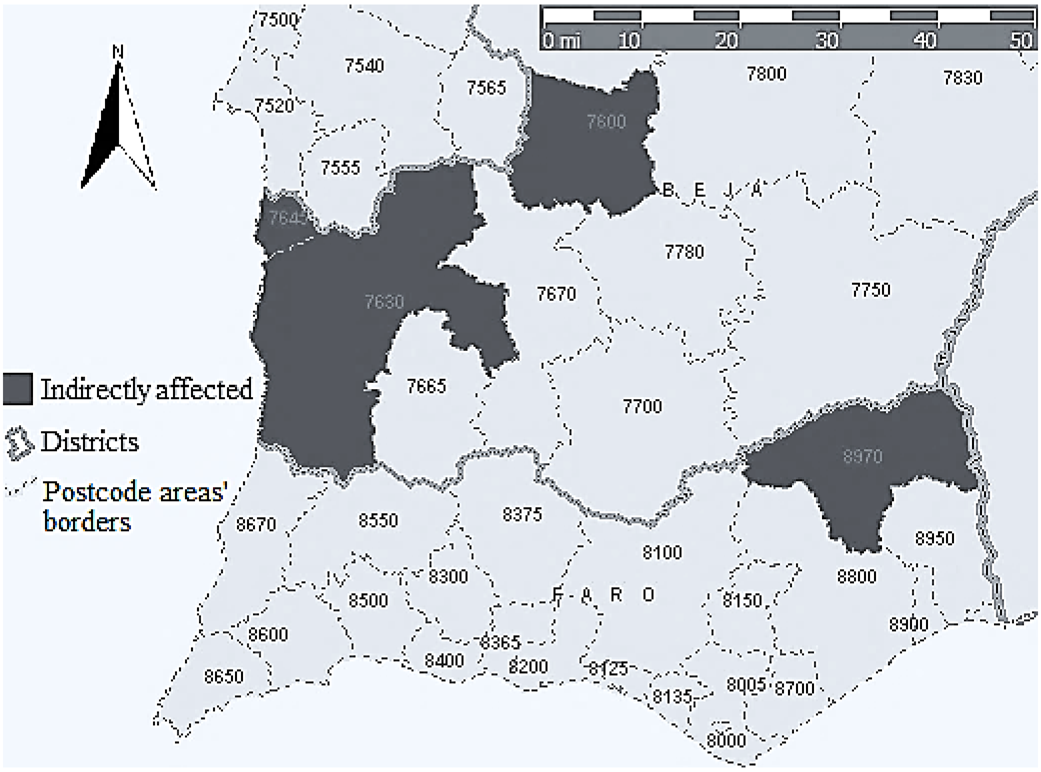

Figure 31 – Areas indirectly affected by the creation of the access barriers

Consequently, the changes in the competitive situation are verified only in 25 postcode areas, which have their catchment areas limited, such that the results obtained with the steps 3 and 4 of our EKD4SFCA method are affected (see equations (18) and (19), in subsections 3.2.2 and 3.2.3, respectively). These are the only geographic markets where firms suffer changes in their market shares. Accordingly, these are also the only markets where the market concentration and the dominance threshold change, as calculated by our extended HHI and the extended dominance identification method, respectively.

The changes that we identified with step 3 are presented in Table 17. Market concentration increases in the whole region, with all 25 highly concentrated postcode areas presenting HHI increases.

Table 17 – The 25 HHI changes caused by the access barriers

Postcode area	**District**	**HHI**		**Change**
		Before	**After**	
7665	Beja	5544	6406	862
7670		5080	5389	309
7700		5088	7185	2097
7780		5057	6627	1570
8000	Faro (east)	5963	8094	2131
8005		5641	8094	2453
8100		5485	7650	2165
8125		5332	7896	2564
8135		5521	8021	2500
8150		6052	7899	1847
8200		5095	7884	2789
8700		5781	8064	2284
8800		6865	7714	849
8900		6699	7744	1045
8950		7508	7510	2
8300	Faro (west)	5010	5380	370
8365		5046	5371	325
8375		5098	5375	276
8400		5001	5420	419
8500		5072	5608	536
8550		5294	5689	396
8600		5000	5184	184
8650		5018	5072	54
8670		5016	5075	60
7565	Setúbal	5421	5488	67

Regarding the dominance identification, all 25 postcode areas were already dominated before the creation of the access barriers, and they keep being dominated thereafter. However, the dominant firms change in postcode areas

7670, 8365, 8375, 8400 and 8600, which means that these geographic markets keep being dominated, but by another firm. Table 18 summarizes the results of the sensitivity to change analysis to the market dominance identification with step 4 of our proposed EKD4SFCA method (CGD stands for HPP Saúde/CGD, HPA is Hospital Particular do Algarve and ESS is Espírito Santo Saúde).

Table 18 – Changes in the market dominance identification

Postcode area	**District**	**Dominant firm**		**Highest share**		**Dom. threshold**	
		Before	**After**	**Before**	**After**	**Before**	**After**
7665	Beja	HPA	HPA	66.5%	76.5%	33.5%	23.5%
7670		CGD	HPA	56.3%	64.0%	43.7%	36.0%
7700		CGD	CGD	56.6%	83.1%	43.4%	16.9%
7780		CGD	CGD	55.3%	78.5%	44.7%	21.5%
8000	Faro (east)	CGD	CGD	71.9%	89.3%	28.1%	10.7%
8005		CGD	CGD	67.9%	89.3%	32.1%	10.7%
8100		CGD	CGD	65.6%	86.4%	34.4%	13.6%
8125		CGD	CGD	62.9%	88.1%	37.1%	11.9%
8135		CGD	CGD	66.1%	88.9%	33.9%	11.1%
8150		CGD	CGD	72.9%	88.1%	27.1%	11.9%
8200		CGD	CGD	56.9%	88.0%	43.1%	12.0%
8700		CGD	CGD	69.8%	89.1%	30.2%	10.9%
8800		CGD	CGD	80.5%	86.8%	19.5%	13.2%
8900		CGD	CGD	79.1%	87.0%	20.9%	13.0%
8950		CGD	CGD	85.4%	85.4%	14.6%	14.6%
8300	Faro (west)	HPA	HPA	52.2%	63.8%	47.8%	36.2%
8365		CGD	HPA	54.8%	63.6%	45.2%	36.4%
8375		CGD	HPA	57.0%	63.7%	43.0%	36.3%
8400		CGD	HPA	50.6%	64.5%	49.4%	35.5%
8500		HPA	HPA	56.0%	67.4%	44.0%	32.6%
8550		HPA	HPA	62.1%	68.6%	37.9%	31.4%
8600		CGD	HPA	50.3%	59.6%	49.7%	40.4%
8650		HPA	HPA	50.3%	56.0%	49.7%	44.0%
8670		HPA	HPA	52.8%	56.1%	47.2%	43.9%
7565	Setúbal	ESS	ESS	70.5%	70.9%	27.2%	27.0%

The highest market share augments in all 25 postcode areas, indicating that market dominance becomes stronger (we note that the increase in postcode area 8950 cannot be identified in Table 18, because it is a very small increase, of around 0.012 percentage points).

To sum up the results obtained with steps 3 and 4 of our method, we conclude that competition weakens in the whole affected region formed by the 25 postcode areas that have their catchment areas shortened. This occurs because the access barriers block competition between the western and eastern hospitals in Faro, which are owned by two firms in both sides: CGD and HPA.

Regarding competition between these two firms in the affected region, they are the sole competing firms in Faro and Beja. While CGD dominates the entire eastern region, including the eastern postcode areas of Beja (7700 and 7780), and remains dominant, after the access barriers, HPA dominates a great part of the western region, including the western postcode areas of Beja (7665 and 7670), and becomes the only dominant firm. Thus, what our methods captures after the creation of the access barriers is a weaker competitive pressure of HPA in the east of Faro and Beja, and a weaker competitive pressure of CGD in the west of Faro and Beja. The largest CGD hospital in the east of Faro stops competing with the largest HPA hospital in the west of Faro, and vice versa, such that we see more uneven competitive pressures after the barriers: in the east, the smallest HPA hospital becomes the only competitor of CGD; in the west, the smallest CGD hospital becomes the only competitor of HPA.

In postcode area 7565, in the district of Setúbal, other hospitals compete with each other, located in different regions, including the large hospitals of Lisbon owned by the dominant firm ESS. Nevertheless, the change that our method captures is due to the fact that the competition from HPA of the eastern side of Faro is blocked with the access barriers. We note that the change is very small, due to the long distance from the eastern side of Faro until postcode area 7565 and the distance decay. In fact, the HHI increases only 67 points and the highest market share, of ESS, increases only 0.4%, as seen in Table 17 and Table 19, respectively.

5.2.1.3. Final results

As seen, the access and competition changes that are produced by the creation of the access barriers, as captured by our method, are logical, expected changes. This indicates that the method behaves as it should, in accordance with how it was constructed and what it proposes to assess, producing further evidence that the method attains construct validity.

We finally note that no changes in the target markets occur, because, even though there are some access level changes, from medium to low, in Faro, these areas remain highly concentrated and dominated after the competition changes. Thus, the target markets remain the same as the ones identified in Figure 21 (see section 4.3).

5.3. Conclusions

This chapter is dedicated to the validation of the proposed EKD4SFCA method. We determined that our method attains content validity and construct validity. First, we established that the EKD4SFCA method attains content validity, considering the literature and expert reviews that were carried out, which aided in the development of the method and of its application. Then, we presented two validation exercises, based on sensitivity to change analysis, to conclude on the construct validity of the method. The two validation exercises, presented in sections 5.1 and 5.2, established that the method reacts as expected to the changes that we introduced to the data sets that were used in these exercises. The changes captured by our method were consistent with the theoretical reasoning and the expectations regarding the changes that access and competition assessments should capture.

6. Concluding remarks and future research

This thesis proposes a new method to identify fine-resolution geographic target markets with low access and no potential competition problems for firms and public administrations to promote new supply. The method corrects limitations of previous methods, enhancing the assessments of access and of competition. Furthermore, it innovatively integrates the two kinds of assessment. This integration makes it possible to process a comprehensive set of data, in order to, on one hand, identify geographic areas for a comparatively higher chance of firms' survival and expansion post-market entry. On the other hand, this method can be used to support public policies that aim at reducing access inequalities and increasing health care supply to disadvantaged populations.

Based on an enhanced influential method to assess patients' access to health care, the KD2SFCA method, the development of our proposed EKD4SFCA method improves it by:

(i) Incorporating non-spatial components that characterize the health needs of the populations;

(ii) Emulating Stouffer's concept to properly represent the populations' mobility when traveling to health care supply points; and

(iii) Introducing an adequate distance decay function to fit best the weighing of proximity in the health care context.

It also improves the competition assessment methods that are integrated into the model, namely the HHI and the dominance identification method of Melnik et al. (2008). More specifically, it corrects the following limitations of these methods:

(i) The disregard of border crossing of consumers;

(ii) The lack of variability in market concentration within the relatively large areas;

(iii) The lack of measurement of distance decay or impedance; and

(iv) The vulnerability to the MAUP.

In the presentation of the four constituent parts of the proposed EKD4SFCA method, we explain the limitations of the existing methods and the problems that are addressed by our adaptations and extensions. Presenting case studies, we establish that our innovations introduce enhancements that improve the assessments of access and of competition. In particular, we show that the adaptions and extensions enable an improved identification of low access regions, where disadvantaged populations reside, and of areas with potential competition problems, i.e., highly concentrated and dominated markets.

Regarding the validation of the EKD4SFCA, we first determine that our method attains content validity, based on the literature review that was carried out and the feedback of experts gathered in conferences, in reviews of submitted papers and at the Portuguese Health Regulatory Agency. We also present two construct validity exercises, in which changes to the data are introduced and the reaction of the method to these changes is assessed. With these exercises, we analyze the responsiveness of our method to confirm that the method captures changes in the data as expected, considering how it is constructed, the information that it uses and its objectives.

Finally, we can conclude that the main objective of this thesis, which, as defined in chapter 1, was to create a new method to assess access to health care and competition between health care providers in an integrated way, was achieved. To the best of our knowledge, this is the first time that these two kinds of methods are combined, making it possible to assess access and competition in an integrated way to identify fine-resolution target markets. The usefulness of the results of such an assessment is clear for firms' market entry strategies and for public policies directed to reduce access disparities in an efficient way.

The supplementary objectives were also accomplished. As a matter of fact, we introduced improvements in the access and competition assessment methods that are integrated into our EKD4SFCA method, and we contributed to the literature on access to health care and competition among health care providers.

The specific expected results regarding the theoretical background, the access assessment, the competition assessment and our proposed method were obtained. In fact, we were able to:

1. Identify the main methods in the context of our research focus, define enhancements to these methods and establish that a new, integrated method was needed to enable the identification of target geographic markets for the promotion of new supply;
2. Propose the EKD2SFCA method to assess access to health care supply points by producing access scores, using a simple GIS, easily obtainable information on supply, and other data collected systematically by national statistics institutes;
3. Propose innovative competition assessment methods, namely the extended HHI and the extended dominance identification method; and
4. Integrate the assessments of access and of competition with the proposed EKD4SFCA method, incorporating the extended HHI and the extended dominance identification method into the EKD2SFCA framework.

Nevertheless, this thesis clearly does not exhaust the research that we followed, since some limitations of the method, its development and its applications remain, leaving space for future research. In fact, the main limitation was already mentioned in chapter 4, when we remarked that the method does not substitute the preparation of a thorough business plan that is needed to support entry decisions (see section 4.3).

Although the method enables an integrated assessment of access and competition that can be useful to support firms' decisions and public policies, a more complete analysis of access and competition must take into consideration additional information and assessments. This means that other lines of work can complement our work, but also that our method and its application can still be improved, aiming to achieve a more complete analysis.

Therefore, we conclude this thesis with a list of paths for future research that we identified, which shall be explored to improve our work:

(i) Location-allocation: as we noted in section 3.4, the EKD4SFCA method could play a role as a location-allocation model. Further application of the method could explore this role, for the identification of the best locations for entrant firms.

(ii) EKD2SFCA method: we compared our method to assess access with the KD2SFCA method. Since a multitude of new 2SFCA-based methods have been published in the scientific literature, and still much more are expected to be developed and published, future work could also compare the EKD2SFCA method with these newer 2SFCA-based versions, in order to identify additional improvement opportunities.

(iii) Health needs and commuting indices: even though we indicated reasonable justifications in subsection 3.1.1 for constructing the health needs and commuting indices with the first component of PCA, we also indicated that the limitation of using only the first component is that it uses less information than it would be the case if more components were used. Future work could explore the utilization of more components and corresponding weights.

(iv) Utilization data: if the researcher is able to collect complete and updated utilization data, these data could be used to try to define in a more precise way: the distance decay function that we used in the four steps of our method; the catchment areas sizes; the variables to be used in the health needs index; and the calculation of the market shares.

(v) Commuting: future research could explore the comparison of the results of our method, with the results of a modified method with the capping function of McGrail and Humphreys (2009a, 2009b, 2009c), instead of the commuting index.

(vi) Dominance identification parameter: future work could explore the estimation of the parameter of the dominance identification method. As we noted in subsection 2.2.3, that parameter represents industry-specific entry barriers.

(vii) Industry: future research could explore how our method should be adapted to be applied to other industries. Depending on the industry, price or the financial status of the populations should be considered in the model, as a representation of the financial barriers that will be affecting access more intensely. In that case, the health needs index could be substituted by another index that would reflect the financial status of the populations.

Also, regarding the application of our method, the consideration of telemedicine could also be explored. Although it is still marginal, it shall imply modifications to our work, because long-distance medical consultations and treatments could affect the catchment areas sizes.

References

ACCC (2008). Merger Guidelines from the Australian Competition & Consumer Commission. Retrieved from http://www.accc.gov.au/publications/merger-guidelines.

Aday, L. A. and Andersen, R. (1974). A Framework for the Study of Access to Medical Care. *Health Services Research,* 9 (3), 208-220.

Alcaraz, K. I., Kreuter, M. W. and Bryan, R. P. (2009). Use of GIS to identify optimal settings for cancer prevention and control in African American communities. *Preventive Medicine,* 49 (1), 54-57.

Alexiadis, P., Dillon, D., Cole, M. and Pita Barros, P. (2003). Methodologies for market definition and market analysis: study for ICP-ANACOM. Retrieved from http://www.anacom.pt/streaming/WIK.pdf?contentId=128770&field=ATTACHED_FILE.

Allin, S., Masseria, C., Sorenson, C., Papanicolas, I. and Mossialos, E. (2007). Measuring inequalities in access to health care: a review of the indices: research note for the European Commission. Retrieved from http://eprints.lse.ac.uk/29837/.

Andersen, R. (1995). Revisiting the behavioral model and access to medical care: does it matter? *Journal of Health and Social Behavior,* 36 (1), 1-10.

Andersen, R., McCutcheon, A., Aday, L. A., Chiu, G. Y. and Bell, R. (1983). Exploring dimensions of access to medical care. *Health Services Research,* 18 (1), 49-74.

Andersen, R. and Newman, J. F. (1973). Societal and individual determinants of medical care utilization in the United States. *Milbank Memorial Fund Quarterly: Health and Society,* 51 (1), 95-124.

Andrulis, D. P. (1998). Access to Care Is the Centerpiece in the Elimination of Socioeconomic Disparities in Health. *Annals of Internal Medicine,* 129 (5), 412-416.

Bagheri, N., Benwell, G. L. and Holt, A. (2008). Modelling Accessibility to Primary Health Care Using a Spatial Accessibility Index and a Need Index. *Hawai'i Journal of Public Health,* 1 (1), 14-27.

Bei, L.-T. and Cheng, T.-C. (2012). Brand power index – using principal component analysis. *Applied Economics,* 45 (20), 2954-2960.

Bernstein, A. B. and Gauthier, A. K. (1998). Defining competition in markets: why and how? (Examining the Role of Regulation in an Evolving Healthcare Marketplace). *Health Services Research,* 33 (5), 1421-1438.

Birkin, M., Clarke, G. and Gibson, P. (2005). Reevaluating funding mechanisms for UK Primary Care Trusts: a new model of secondary care usage incorporating lifestyle data. *Environment and Planing C: Government and Policy,* 23 (3), 317-336.

Borruso, G. (2005). Network Density Estimation: Analysis of Point Patterns over a Network *Computational Science and Its Applications – ICCSA 2005* (Vol. 3482/2005, pp. 126-132).

Bullen, N., Moon, G. and Jones, K. (1996). Defining localities for health planning: A GIS approach. *Social Science & Medicine,* 42 (6), 801-816.

Burkey, M. L. (2012). Decomposing geographic accessibility into component parts: methods and an application to hospitals. *The Annals of Regional Science,* 48 (3), 783-800.

Burlingame, G. M., Seaman, S., Johnson, J. E., Whipple, J., Richardson, E., Rees, F., Earnshaw, D., Spencer, R., Payne, M. and O'Neil, B. (2006). Sensitivity to change of the Brief Psychiatric Rating Scale-Extended (BPRS-E): An item and subscale analysis. *Psychological Services,* 3 (2), 77-87.

CC (2013). Guidelines for market investigations. Retrieved from https://www.gov.uk/government/uploads/system/uploads/attachment_data/file/284390/cc3_revised.pdf.

CC&OFT (2010). Merger assessment guidelines. Retrieved from https://www.gov.uk/government/uploads/system/uploads/attachment_data/file/284449/OFT1254.pdf.

Cervigni, F., Suzuki, Y., Ishii, T. and Hata, A. (2008). Spatial Accessibility to Pediatric Services. *Journal of Community Health,* 33 (6), 444-448.

CGAP (2000). Assessing the relative poverty of microfinance clients: A CGAP operational tool. Retrieved from http://www.ifpri.org/publication/assessing-relative-poverty-microfinance-clients?print.

Chaix, B., Merlo, J. and Chauvin, P. (2005). Comparison of a spatial approach with the multilevel approach for investigating place effects on health: the example of healthcare utilisation in France. *Journal of Epidemiology and Community Health,* 59 (6), 517-526.

Chua, C. L., Palangkaraya, A. and Yong, J. (2011). Hospital Competition, Technical Efficiency and Quality. *The Economic Record,* 87 (227), 252-268.

Cinnamon, J., Schuurman, N. and Crooks, V. A. (2008). A method to determine spatial access to specialized palliative care services using GIS. *BMC Health Services Research,* 8 (140), 1-11.

Clarke, G., Eyre, H. and Guy, C. (2002). Deriving Indicators of Access to Food Retail Provision in British Cities: Studies of Cardiff, Leeds and Bradford. *Urban Studies,* 39 (11), 2041-2060.

Cooper, H., Bossak, B., Tempalski, B., Des Jarlais, D. and Friedman, S. (2009a). Geographic approaches to quantifying the risk environment: Drug-related law enforcement and access to syringe exchange programmes. *International Journal of Drug Policy,* 20 (3), 217-226.

Cooper, H., Bossak, B., Tempalski, B., Des Jarlais, D. and Friedman, S. (2009b). Temporal Trends in Spatial Access to Pharmacies that Sell Over-the-Counter Syringes in New York City Health Districts: Relationship to Local Racial/Ethnic Composition and Need. *Journal of Urban Health,* 86 (6), 929-945.

Corcoran, J., Higgs, G. and Anderson, T. (2013). Examining the use of a geodemographic classification in an exploratory analysis of variations in fire incidence in South Wales, UK. *Fire Safety Journal,* 62, Part A (11), 37-48.

Correia, I. and Veiga, P. (2010). Geographic distribution of physicians in Portugal. *The European Journal of Health Economics,* 11 (4), 383-393.

Cromley, E. K. and Albertsen, P. C. (1993). Multiple-Site Physician Practices and Their Effect on Service Distribution. *Health Services Research,* 28 (4), 503-522.

Cromley, E. K. and Shannon, G. W. (1986). Locating Ambulatory Medical Care Facilities for the Elderly. *Health Services Research,* 21 (4), 499-514.

Dai, D. and Wang, F. (2011). Geographic disparities in accessibility to food stores in southwest Mississippi. *Environment and Planning B: Planning and Design,* 38 (4), 659-677.

Dalgic, T. and Leeuw, M. (1994). Niche Marketing Revisited: Concept, Applications and Some European Cases. *European Journal of Marketing,* 28 (4), 39-55.

Darby, M. R. and Karni, E. (1973). Free Competition and the Optimal Amount of Fraud. *Journal of Law and Economics,* 16 (1), 67-88.

Delmelle, E., Delmelle, E., Casas, I. and Barto, T. (2010). H.E.L.P: A GIS-based Health Exploratory AnaLysis Tool for Practitioners. *Applied Spatial Analysis and Policy,* 4 (2), 113-137.

Dewulf, B., Neutens, T., De Weerdt, Y. and Van de Weghe, N. (2013). Accessibility to primary health care in Belgium: an evaluation of policies awarding financial assistance in shortage areas. *BMC Family Practice,* 14 (1), 1-13.

Donabedian, A. (1972). Models for Organizing the Delivery of Personal Health Services and Criteria for Evaluating Them. *The Milbank Memorial Fund Quarterly,* 50 (4), 103-154.

EC (1997). Commission Notice on the definition of relevant market for the purposes of Community competition law. *Official Journal of the European Communities,* C 372 (3), 5-13.

EC (2004). Guidelines on the assessment of horizontal mergers under the Council Regulation on the control of concentrations between undertakings. *Official Journal of the European Union,* C 31 (3), 5-18.

EC (2007). Concentration operation, Case N. COMP/M.4367. Retrieved from http://ec.europa.eu/competition/mergers/cases/decisions/m4367_20070316_20212_en.pdf.

ERS (2011). Acesso dos Utentes aos Cuidados Continuados de Saúde. Retrieved from http://www.ers.pt/uploads/writer_file/document/144/Relatorio_Cuidados_Continuados.pdf.

ERS (2013). Avaliação do Acesso dos Utentes aos Cuidados Continuados de Saúde. Retrieved from http://www.ers.pt/pages/73?news_id=620.

ERS (2014a). Avaliação Concorrencial dos Mercados de Meios Complementares de Diagnóstico e Terapêutica de Gastrenterologia. Entidade Reguladora da Saúde. Unpublished.

ERS (2014b). Avaliação Concorrencial dos Mercados de Serviços de Hemodiálise. Entidade Reguladora da Saúde. Unpublished.

ERS (2014c). Parecer da ERS sobre a Operação de Concentração com a Referência Ccent 23/2014 – José de Mello Saúde/Espírito Santo Saúde. Retrieved from http://www.ers.pt/pages/73?news_id=1062.

EU (2011). Directive 2011/24/EU of the European Parliament and of the Council of 9 March 2011 on the application of patients' rights in cross-border healthcare. *Official Journal of the European Union,* L 88 (54), 45-65.

EU (2012). Consolidated Version of the Treaty on the Functioning of the European Union. *Official Journal of the European Union,* C 326 (55), 47-390.

Field, K. (2000). Measuring the need for primary health care: an index of relative disadvantage. *Applied Geography,* 20 (4), 305-332.

Field, K. and Briggs, D. J. (2001). Socio-economic and locational determinants of accessibility and utilization of primary health-care. *Health & Social Care in the Community,* 9 (5), 294-308.

Filmer, D. and Pritchett, L. H. (2001). Estimating Wealth Effects Without Expenditure Data - Or Tears: An Application To Educational Enrollments In States Of India. *Demography,* 38 (1), 115-132.

Fortney, J., Rost, K. and Warren, J. (2000). Comparing Alternative Methods of Measuring Geographic Access to Health Services. *Health Services and Outcomes Research Methodology,* 1 (2), 173-184.

Fotheringham, A. S. and Wong, D. W. S. (1991). The modifiable areal unit problem in multivariate statistical analysis. *Environment and Planning A,* 23 (7), 1025-1044.

Gatrell, A. C., Bailey, T. C., Diggle, P. J. and Rowlingson, B. S. (1996). Spatial point pattern analysis and its application in geographical epidemiology. *Transactions of the Institute of British Geographers,* 21, 256-274.

Gatrell, A. C. and Wood, D. J. (2012). Variation in geographic access to specialist inpatient hospices in England and Wales. *Health & Place,* 18 (4), 832-840.

Gaynor, M. (2006). *What Do We Know About Competition and Quality in Health Care Markets*. NBER Working Paper 12301.

Gaynor, M. and Haas-Wilson, D. (1999). Change, consolidation, and competition in health care markets. *Journal of Economic Perspectives,* 13 (1), 141-164.

Gaynor, M., Ho, K. and Town, R. (2014). *The Industrial Organization of Health Care Markets*. NBER Working Paper 19800.

Gaynor, M., Moreno-Serra, R. and Propper, C. (2012a). Can competition improve outcomes in UK health care? Lessons from the past two decades. *Journal of Health Services Research & Policy,* 17 (1), 49-54.

Gaynor, M., Moreno-Serra, R. and Propper, C. (2012b). *Death by Market Power: Reform, Competition and Patient Outcomes in the National Health Service*. NBER Working Paper 16164.

Gesler, W. M. and Meade, M. S. (1988). Locational and population factors in health care-seeking behavior in Savannah, Georgia. *Health Services Research,* 23 (3), 443-462.

Geurs, K. T. and van Wee, B. (2004). Accessibility evaluation of land-use and transport strategies: review and research directions. *Journal of Transport Geography,* 12 (2), 127-140.

Gold, M. (1998). Beyond coverage and supply: measuring access to healthcare in today's market. *Health Services Research,* 33 (3), 625-684.

Goldenberg, M. J. (2012). Defining "quality of care" persuasively. *Theoretical medicine and bioethics,* 33 (4), 243-261.

Goodchild, M. F. and Haining, R. P. (2004). GIS and spatial data analysis: Converging perspectives. *Papers in Regional Science,* 83 (1), 363-385.

Goody, B. (1993). Defining rural hospital markets. *Health Services Research,* 28 (2), 183-200.

Graves, A. (2009). A Model for Assessment of Potential Geographical Accessibility: A Case for GIS. *Online Journal of Rural Nursing & Health Care,* 9 (1), 46-55.

Guagliardo, M. (2004). Spatial accessibility of primary care: concepts, methods and challenges. *International Journal of Health Geographics,* 3 (1), 1-13.

Guagliardo, M. and Ronzio, C. R. (2005). Is Region of Country a Useful Variable for Child Health Studies? *Pediatrics,* 116 (6), 1542-1545.

Guagliardo, M., Ronzio, C. R., Cheung, I., Chacko, E. and Joseph, J. G. (2004). Physician accessibility: an urban case study of pediatric providers. *Health & Place,* 10 (3), 273-283.

Gulzar, L. (1999). Access to health care. *Journal of Nursing Scholarship,* 31 (1), 13-19.

Halbersma, R. S., Mikkers, M. C., Motchenkova, E. and Seinen, I. (2011). Market structure and hospital-insurer bargaining in the Netherlands. *The European Journal of Health Economics,* 12 (6), 589-603.

Hansen, W. G. (1959). How Accessibility Shapes Land Use. *Journal of the American Institute of Planners,* 25 (2), 73 - 76.

Harrington, D. W., Wilson, K., Bell, S., Muhajarine, N. and Ruthart, J. (2012). Realizing neighbourhood potential? The role of the availability of health care services on contact with a primary care physician. *Health & Place,* 18 (4), 814-823.

Havard, S., Deguen, S., Bodin, J., Louis, K., Laurent, O. and Bard, D. (2008). A small-area index of socioeconomic deprivation to capture health inequalities in France. *Social Science & Medicine,* 67 (12), 2007-2016.

Hay, A. M. (1995). Concepts of equity, fairness and justice in geographical studies. *Transactions of the Institute of British Geographers,* 20 (4), 500-508.

Haynes, S. N., Richard, D. C. S. and Kubany, E. S. (1995). Content validity in psychological assessment: A functional approach to concepts and methods. *Psychological Assessment,* 7 (3), 238-247.

Hays, R. D. and Hadorn, D. (1992). Responsiveness to change: an aspect of validity, not a separate dimension. *Quality of Life Research,* 1 (1), 73-75.

Hellmer, S. and Wårell, L. (2009). On the evaluation of market power and market dominance - The Nordic electricity market. *Energy Policy,* 37 (8), 3235-3241.

Henry, C., Sharma, M., Lapenu, C. and Zeller, M (2003). Microfinance Poverty Assessment Tool - CGAP Technical Tool series. Retrieved from http://documents.worldbank.org/curated/en/2003/10/2857591/microfinance-poverty-assessment-tool.

Higgs, G. (2004). A Literature Review of the Use of GIS-Based Measures of Access to Health Care Services. *Health Services and Outcomes Research Methodology,* 5 (2), 119-139.

Higgs, G. (2009). The role of GIS for health utilization studies: literature review. *Health Services and Outcomes Research Methodology,* 9 (2), 84-99.

Higgs, G., Langford, M. and Fry, R. (2013). Investigating variations in the provision of digital services in public libraries using network-based GIS models. *Library & Information Science Research,* 35 (1), 24-32.

Hirschman, A. O. (1964). The Paternity of an Index. *The American Economic Review,* 54 (5), 761.

Howe, L. D., Hargreaves, J. R. and Huttly, S. R. (2008). Issues in the construction of wealth indices for the measurement of socio-economic position in low-income countries. *Emerging Themes in Epidemiology,* 5 (3).

Hughes, J. G., Budetti, P. P., Chapman, D. D., Cramblett, H. G., Mathies, A. W., Meyer, B. P., Seidel, H., Slungaard, R. K., Connelly, J. P. and Bornstein, S. L. (1981). Critique of the Final Report of the Graduate Medical Education National Advisory Committee. *Pediatrics,* 67 (5), 585-596.

Huiban, J. P. (2011). The spatial demography of new plants: urban creation and rural survival. *Small Business Economics,* 37 (1), 73-86.

Humphreys, J. S. (1998). Delimiting 'Rural': Implications of an Agreed 'Rurality' Index for Healthcare Planning and Resource Allocation. *Australian Journal of Rural Health,* 6 (4), 212-216.

Ikkersheim, D. and Koolman, X. (2012). Dutch healthcare reform: did it result in better patient experiences in hospitals? A comparison of the consumer quality index over time. *BMC Health Services Research,* 12 (35), 76.

Jain, A. K. (2010). Data clustering: 50 years beyond K-means. *Pattern Recognition Letters,* 31 (8), 651-666.

Jarvis, E. (1866). Influence of distance from and nearness to an insane hospital on its use by the people. *American Journal of Insanity,* 22, 361-433.

Jolliffe, I. T. (2002). *Principal Component Analysis, Second Edition*. New York, NY: Springer.

Joseph, A. E. and Bantock, P. R. (1982). Measuring potential physical accessibility to general practitioners in rural areas: A method and case study. *Social Science & Medicine,* 16 (1), 85-90.

Joseph, A. E. and Phillips, D. R. (1984). *Accessibility and Utilization: Geographical perspectives on health care delivery*. New York, NY: Harper & Row.

Kaiser, H. F. (1974). An Index of Factorial Simplicity. *Psychometrika,* 39 (1), 31-36.

Kane, M. (2012). Validating score interpretations and uses. *Language Testing,* 29 (1), 3-17.

Kaplan, R. M., Bush, J. W. and Berry, C. C. (1976). Health status: types of validity and the index of well-being. *Health Services Research,* 11 (4), 478-507.

Kayri, M. and Zırhlıoğlu, G. (2009). Kernel Smoothing Function and Choosing Bandwidth for Non-Parametric Regression Methods. *Ozean Journal of Applied Sciences,* 2 (1), 49-54.

Knoche, F. and Thöni, M. (2011). Sozial- und wettbewerbsrechtliche Konflikte bei der Fusionskontrolle in der Gesetzlichen Krankenversicherung in Deutschland. *Zeitschrift für die gesamte Versicherungswissenschaft,* 100 (4), 539-560.

Kwan, M. (2004). GIS Methods in Time-Geographic Research: Geocomputation and Geovisualization of Human Activity Patterns. *Geografiska Annaler: Series B, Human Geography,* 86 (4), 267-280.

Kwan, M. and Weber, J. (2003). Individual Accessibility Revisited: Implications for Geographical Analysis in the Twenty-first Century. *Geographical Analysis,* 35 (4), 341-353.

Langford, M., Higgs, G. and Fry, R. (2012). Using floating catchment analysis (FCA) techniques to examine intra-urban variations in accessibility to public transport opportunities: the example of Cardiff, Wales. *Journal of Transport Geography,* 25 (11), 1-14.

Lee, G. and Hong, I. (2013). Measuring spatial accessibility in the context of spatial disparity between demand and supply of urban park service. *Landscape and Urban Planning,* 119, 85-90.

Lewis, D. J. and Longley, P. A. (2012). Patterns of Patient Registration with Primary Health Care in the UK National Health Service. *Annals of the Association of American Geographers,* 102 (5), 1135-1145.

Lian, M., Struthers, J. and Schootman, M. (2012). Comparing GIS-Based Measures in Access to Mammography and Their Validity in Predicting Neighborhood Risk of Late-Stage Breast Cancer. *PLoS ONE,* 7 (8), e43000.

Lin, S.-J., Crawford, S. Y. and Warren Salmon, J. (2005). Potential access and revealed access to pain management medications. *Social Science & Medicine,* 60 (8), 1881-1891.

Love, D. and Lindquist, P. (1995). The geographical accessibility of hospitals to the aged: a geographic information systems analysis within Illinois. *Health Services Research,* 29 (6), 629-651.

Luo, W. (2004). Using a GIS-based floating catchment method to assess areas with shortage of physicians. *Health & Place,* 10 (1), 1-11.

Luo, W. and Qi, Y. (2009). An enhanced two-step floating catchment area (E2SFCA) method for measuring spatial accessibility to primary care physicians. *Health & Place,* 15 (4), 1100-1107.

Luo, W. and Wang, F. (2003). Measures of spatial accessibility to health care in a GIS environment: synthesis and a case study in the Chicago region. *Environment and Planning B: Planning and Design,* 30 (6), 865-884.

Makuc, D. M., Haglund, B., Ingram, D. D., Kleinman, J. C. and Feldman, J. J. (1991). The Use of Health Service Areas for Measuring Provider Availability. *The Journal of Rural Health,* 7 (4), 347-356.

Manyong, V. M., Okike, I. and Williams, T. O. (2006). Effective Dimensionality and Factors Affecting Crop-Livestock Integration in West African Savannas: A Combination of Principal Component Analysis and Tobit Approaches. *Agricultural Economics,* 35 (2), 145-155.

Matsumoto, M., Ogawa, T., Kashima, S. and Takeuchi, K. (2012). The impact of rural hospital closures on equity of commuting time for haemodialysis patients: simulation analysis using the capacity-distance model. *International Journal of Health Geographics,* 11 (28).

McDowell, I. (2006). *Measuring Health: A Guide to Rating Scales, and Questionnaires* (Third ed.). New York, New York: Oxford University Press, Inc.

McGrail, M. (2012). Spatial accessibility of primary health care utilising the two step floating catchment area method: an assessment of recent improvements. *International Journal of Health Geographics,* 11 (1), 50.

McGrail, M. and Humphreys, J. S. (2009a). The index of rural access: an innovative integrated approach for measuring primary care access. *BMC Health Services Research,* 9 (1), 124.

McGrail, M. and Humphreys, J. S. (2009b). Measuring spatial accessibility to primary care in rural areas: Improving the effectiveness of the two-step floating catchment area method. *Applied Geography,* 29 (4), 533-541.

McGrail, M. and Humphreys, J. S. (2009c). A new index of access to primary care services in rural areas. *Australian and New Zealand Journal of Public Health,* 33 (5), 418-423.

McIntosh, C. and Hellmer, S. (2012). Necessary and sufficient conditions in merger control: the use of HHI and threshold value. *Applied Economics,* 44 (7), 867-878.

McLafferty, S. L. (2003). GIS and Health Care. *Annual Review of Public Health,* 24 (1), 25-42.

McLafferty, S. L. and Grady, S. (2004). Prenatal Care Need and Access: A GIS Analysis. *Journal of Medical Systems,* 28 (3), 321-333.

McLaughlin, C. G. and Wyszewianski, L. (2002). Access to care: remembering old lessons. *Health Services Research,* 37 (6), 1441-1443.

Melnik, A., Shy, O. and Stenbacka, R. (2008). Assessing market dominance. *Journal of Economic Behavior & Organization,* 68 (1), 63-72.

Messer, L. C., Laraia, B. A., Kaufman, J. S., Eyster, J., Holzman, C., Culhane, J., Elo, I., Burke, J. G. and O'Campo, P. (2006). The Development of a Standardized Neighborhood Deprivation Index. *Journal of Urban Health,* 83 (6), 1041-1062.

Messick, S. (1990). Validity of Test Interpretation and Use. Retrieved from http://eric.ed.gov/?id=ED395031.

Mooi, E. A. and Sarstedt, M. (2011). *A Concise Guide to Market Research. The Process, Data and Methods using IBM SPSS Statistics*. Berlin: Springer.

Motta, M. (2004). *Competition Policy: Theory and Practice*. New York: Cambridge University Press.

Muchara, B., Ortmann, G. and Mudhara, M. (2014). Collective action and participation in irrigation water management: A case study of Mooi River Irrigation Scheme in KwaZulu-Natal Province, South Africa. *Water SA,* 40 (4), 699-708.

Nakamba, P., Hanson, K. and McPake, B. (2002). Markets for hospital services in Zambia. *International Journal of Health Planning and Management,* 17 (3), 229-247.

Neupane, M., Basnyat, B., Fischer, R., Froeschl, G., Wolbers, M. and Rehfuess, E. A. (2015). Sustained use of biogas fuel and blood pressure among women in rural Nepal. *Environmental Research,* 136 (1), 343-351.

Neutens, T., Versichele, M. and Schwanen, T. (2010). Arranging place and time: A GIS toolkit to assess person-based accessibility of urban opportunities. *Applied Geography,* 30 (4), 561-575.

Ngui, A. N. and Vanasse, A. (2012). Assessing spatial accessibility to mental health facilities in an urban environment. *Spatial and Spatio-temporal Epidemiology,* 3 (3), 195-203.

Norris, T. L. and Aiken, M. (2006). Personal Access to Health Care: A Concept Analysis. *Public Health Nursing,* 23 (1), 59-66.

OECD (2006a). Health Care Quality Indicators Project: Initial Indicators Report. Retrieved from http://EconPapers.repec.org/RePEc:oec:elsaad:22-en.

OECD (2006b). Policy Roundtables - Evidentiary Issues in Proving Dominance. Retrieved from http://www.oecd.org/competition/abuse/41651328.pdf.

OECD (2011). Competition Assessment Toolkit - Volume I: Principles. Retrieved from http://www.oecd.org/daf/competition/46193173.pdf.

OECD (2012a). Policy Roundtables - Competition in Hospital Services. Retrieved from http://www.oecd.org/daf/competition/50527122.pdf.

OECD (2012b). Policy Roundtables - Market Definition. Retrieved from http://www.oecd.org/daf/competition/Marketdefinition2012.pdf.

OFT (2004a). Empirical indicators for market investigations - Part 1. Retrieved from http://products.ihs.com/Ohsis-SEO/OFFICEOFFAIRTRADING.html.

OFT (2004b). Market definition - Understanding competition law. Retrieved from http://www.gov.uk/government/uploads/system/uploads/attachment_data/file/284423/oft403.pdf.

OFT (2007). Completing competition assessments in impact assessments. Retrieved from http://webarchive.nationalarchives.gov.uk/20140402142426/http://www.oft.gov.uk/shared_oft/reports/comp_policy/oft876.pdf.

Olafsson, A. and Skov-Petersen, H. (2013). The use of GIS-based support of recreational trail planning by local governments. *Applied Spatial Analysis and Policy,* 7 (2), 149-168.

Oliveira, M. D. (2010). Evolução da oferta. In J. Simões (Ed.), *30 anos do Serviço Nacional de Saúde: Um Percurso Comentado* (pp. 271-295). Coimbra: Almedina.

Oliveira, M. D. and Bevan, G. (2003). Measuring geographic inequities in the Portuguese health care system: an estimation of hospital care needs. *Health Policy,* 66 (3), 277-293.

Oliver, A. and Mossialos, E. (2004). Equity of access to health care: outlining the foundations for action. *Journal of Epidemiology and Community Health,* 58 (8), 655-658.

Openshaw, S. (1984). *The Modifiable Areal Unit Problem*. Norwich, UK: Geo Books.

Parzen, E. (1962). On Estimation of a Probability Density Function and Mode. *Annals of Mathematical Statistics,* 33 (3), 1065-1076.

Pavlyuk, D. (2009). Spatial Competition Pressure as a Factor of European Airports' Efficiency. *Transport and Telecommunication,* 10 (4), 8-17.

Penchansky, R. and Thomas, J. W. (1981). The Concept of Access: Definition and Relationship to Consumer Satisfaction. *Medical Care,* 19 (2), 127-140.

Peng, Z.-R. (1997). The Jobs-Housing Balance and Urban Commuting. *Urban Studies,* 34 (8), 1215-1235.

Pepall, L. M. and Richards, D. J. (2002). The Simple Economics of Brand Stretching. *Journal of Business,* 75 (3), 535-552.

Polit, D. F. and Beck, C. T. (2006). The Content Validity Index: Are You Sure You Know What's Being Reported? Critique and Recommendations. *Research in Nursing & Health,* 29 (5), 489-497.

Polo, G., Acosta, C. M. and Dias, R. A. (2013). Spatial accessibility to vaccination sites in a campaign against rabies in São Paulo city, Brazil. *Preventive Veterinary Medicine,* 111 (1–2), 10-16.

Polzin, P., Borges, J. and Coelho, A. (2013). Identifying Target Geographic Areas for New Market Entrants. *Advances in Business-Related Scientific Research Conference, Conference Proceedings,* ISBN 978-961-6347-52-5.

Polzin, P., Borges, J. and Coelho, A. (2014a). An extended kernel density two-step floating catchment area method to analyze access to health care. *Environment and Planing B: Planing and Design,* 41 (4), 717-735.

Polzin, P., Borges, J. and Coelho, A. (2014b). Validating the EKD4SFCA method using data of the long-term care sector. *Advances in Business-Related Scientific Research Journal,* 5 (2), 135-152.

Polzin, P., Borges, J. and Coelho, A. (2015). A decision support method to identify target geographic markets for health care providers. *Papers in Regional Science,* forthcoming.

Porta, S., Latora, V., Wang, F., Strano, E., Cardillo, A., Scellato, S., Iacoviello, V. and Messora, R. (2009). Street centrality and densities of retail and services in Bologna, Italy. *Environment and Planning B: Planning and Design,* 36 (3), 450-465.

Propper, C. (2012). Competition, Incentives and the English NHS. *Health Economics,* 21 (1), 33-40.

Pulaj, E. and Kume, V. (2013). Measuring Market Concentration of Construction Industry. Vlora Region Evidence. *European Scientific Journal,* 9 (32), 121-136.

Radke, J. and Mu, L. (2000). Spatial Decompositions, Modeling and Mapping Service Regions to Predict Access to Social Programs. *Geographic Information Sciences,* 6 (2), 105-112.

Rhoades, S. A. (1993). The Herfindahl-Hirschman Index. *Federal Reserve Bulletin,* 79 (3), 188-189.

Rice, N. and Smith, P. C. (2001). Ethics and geographical equity in health care. *Journal of Medical Ethics,* 27 (4), 256-261.

Ricketts, T. C. and Goldsmith, L. J. (2005). Access in health services research: The battle of the frameworks. *Nursing Outlook,* 53 (6), 274-280.

Salmond, C., Crampton, P., King, P. and Waldegrave, C. (2006). NZiDep: A New Zealand index of socioeconomic deprivation for individuals. *Social Science & Medicine,* 62 (6), 1474-1485.

Schmid, A. and Ulrich, V. (2013). Consolidation and concentration in the German hospital market: The two sides of the coin. *Health Policy,* 109 (3), 301-310.

Schmidtlein, M. C., Deutsch, R. C., Piegorsch, W. W. and Cutter, S. L. (2008). A Sensitivity Analysis of the Social Vulnerability Index. *Risk Analysis,* 28 (4), 1099-1114.

Schönfelder, S. and Axhausen, K. W. (2003). Activity spaces: measures of social exclusion? *Transport Policy,* 10 (4), 273-286.

Schuurman, N., Bérubé, M. and Crooks, V. A. (2010). Measuring potential spatial access to primary health care physicians using a modified gravity model. *Canadian Geographer / Le Géographe Canadien,* 54 (1), 29-45.

Schuurman, N., Crooks, V. A. and Amram, O. (2010). A protocol for determining differences in consistency and depth of palliative care service provision across community sites. *Health and Social Care in the Community,* 18 (5), 537-548.

Seaman, D. E. and Powell, R. A. (1996). An Evaluation of the Accuracy of Kernel Density Estimators for Home Range Analysis. *Ecology,* 77 (7), 2075-2085.

Shaw, H. (2012). Access to healthy food in Nantes, France. *British Food Journal,* 114 (2), 224-238.

Siegfried, J. J. and Evans, L. B. (1994). Empirical studies of entry and exit: A survey of the evidence. *Review of Industrial Organization,* 9 (2), 121-155.

Sloas, L. B., Steele, P. D. and Hare, T. S. (2011). Geographical access to treatment for sex offenders under community supervision in Kentucky. *Journal of Sexual Aggression,* 18 (3), 294-310.

Sohler, K. B. and Clapis, J. A. (1972). Jarvis' law and the planning of mental health services. *HSMHA Health Reports,* 87 (1), 75-80.

Song, P., Zhu, Y., Mao, X., Li, Q. and An, L. (2013). Assessing Spatial Accessibility to Maternity Units in Shenzhen, China. *PLoS ONE,* 8 (7), e70227.

Spencer, J. and Angeles, G. (2007). Kernel density estimation as a technique for assessing availability of health services in Nicaragua. *Health Services and Outcomes Research Methodology,* 7 (3), 145-157.

Stouffer, S. A. (1940). Intervening Opportunities: A Theory Relating Mobility and Distance. *American Sociological Review,* 5 (6), 845-867.

Strong, N., Bollard, A. and Pickford, M. (2000). Defining Market Dominance: A Study of Antitrust Decisions on Business Acquisitions in New Zealand. *Review of Industrial Organization,* 17 (2), 209-227.

Tanser, F. (2006). Methodology for optimising location of new primary health care facilities in rural communities: a case study in KwaZulu-Natal, South Africa. *Journal of Epidemiology and Community Health,* 60 (10), 846-850.

Tareque, M. I., Begum, S. and Saito, Y. (2014). Inequality in Disability in Bangladesh. *PLoS ONE,* 9 (7), e103681.

Tay, A. (2003). Assessing competition in hospital care markets: the importance of accounting for quality differentiation. *RAND Journal of Economics,* 34 (4), 786-814.

Thomas, J. W. and Penchansky, R. (1984). Relating Satisfaction with Access to Utilization of Services. *Medical Care,* 22 (6), 553-568.

Tobler, W. R. (1970). A Computer Movie Simulating Urban Growth in the Detroit Region. *Economic Geography,* 46, 234-240.

Tremblay, B. and Merritt, D. (1995). The Frequency Function of Elliptical Galaxy Intrinsic Shapes. *The Astronomical Journal,* 110, 1039-1060.

UKDH (2007). Radiotherapy: developing a world class service for England - Report to Ministers from National Radiotherapy Advisory Group. Retrieved from http://webarchive.nationalarchives.gov.uk/20130107105354/http://www.dh.gov.uk/prod_consum_dh/groups/dh_digitalassets/@dh/@en/documents/digitalasset/dh_074576.pdf.

UN (1948). The Universal Declaration of Human Rights. Retrieved from http://www.un.org/en/documents/udhr/index.shtml.

van der Merwe, J. H. and van Niekerk, A. (2013). Application of geospatial technology for gap analysis in tourism planning for the Western Cape. *South African Journal of Science,* 109 (3-4), 1-10.

van der Zee, J. and Kroneman, M. W. (2007). Bismarck or Beveridge: a beauty contest between dinosaurs. *BMC Health Services Research,* 7 (94), 1-11.

Vermeersch, D. A., Lambert, M. J. and Burlingame, G. M. (2000). Outcome Questionnaire: Item Sensitivity to Change. *Journal of Personality Assessment,* 74 (2), 242-261.

Vyas, S. and Kumaranayake, L. (2006). Constructing socio-economic status indices: how to use principal components analysis. *Health Policy and Planning,* 21 (6), 459-468.

Walker, M. (2006). Competition Law, Anti-Competitive Behavior, and Merger Analysis: Economic Foundations. Office of the General Counsel. Asian Development Bank.

Wang, F. (2000). Modeling Commuting Patterns in Chicago in a GIS Environment: A Job Accessibility Perspective. *The Professional Geographer,* 52 (1), 120 - 133.

Wang, F. (2012). Why police and policing need GIS: an overview. *Annals of GIS,* 18 (3), 159-171.

Wang, F. and Luo, W. (2005). Assessing spatial and nonspatial factors for healthcare access: towards an integrated approach to defining health professional shortage areas. *Health & Place,* 11 (2), 131-146.

Wang, F. and Tang, Q. (2013). Planning toward equal accessibility to services: a quadratic programming approach. *Environment and Planning B: Planning and Design,* 40 (2), 195-212.

Wang, L. (2007). Immigration, ethnicity, and accessibility to culturally diverse family physicians. *Health & Place,* 13 (3), 656-671.

Wang, L. (2011). Analysing spatial accessibility to health care: a case study of access by different immigrant groups to primary care physicians in Toronto. *Annals of GIS,* 17 (4), 237-251.

Weibull, J. W. (1976). An axiomatic approach to the measurement of accessibility. *Regional Science and Urban Economics,* 6 (4), 357-379.

Whitehead, M., Evandrou, M., Haglund, B. and Diderichsen, F. (1997). As the health divide widens in Sweden and Britain, what's happening to access to care? *BMJ,* 315 (7114), 1006-1009.

WHO (1978). Report of the International Conference on Primary Health Care, Alma Ata. Retrieved from http://www.who.int/hpr/NPH/docs/declaration_almaata.pdf.

WHO (2006). Quality of Care - A process for making strategic choices in health systems. Retrieved from http://www.who.int/management/quality/assurance/QualityCare_B.Def.pdf.

Wing, P. and Reynolds, C. (1988). The availability of physician services: a geographic analysis. *Health Services Research,* 23 (5), 649-667.

Wong, H. S., Zhan, C. and Mutter, R. (2005). Do Different Measures of Hospital Competition Matter in Empirical Investigations of Hospital Behavior. *Review of Industrial Organization,* 26 (1), 27-60.

Wynd, C. A., Schmidt, B. and Schaefer, M. A. (2003). Two Quantitative Approaches for Estimating Content Validity. *Western Journal of Nursing Research,* 25 (5), 508-518.

Xu, K. and Cui, W. (2012). A GIS-Based Assessment of Spatial Accessibility to County Hospitals: A Case Study of Dancheng County, China. In X. Qu and Y. Yang (Eds.), *Information and Business Intelligence* (Vol. 268, pp. 454-460): Springer Berlin Heidelberg.

Yaghmaie, F. (2003). Content validity and its estimation. *Journal of Medical Education,* 3 (1), 25-27.

Yang, D.-H., Goerge, R. and Mullner, R. (2006). Comparing GIS-Based Methods of Measuring Spatial Accessibility to Health Services. *Journal of Medical Systems,* 30 (1), 23-32.

Yang, H.-C., Hsieh, H.-Y. and Fann, C. S. J. (2008). Kernel-Based Association Test. *Genetics,* 179 (2), 1057-1068.

Yao, J., Murray, A. T. and Agadjanian, V. (2013). A geographical perspective on access to sexual and reproductive health care for women in rural Africa. *Social Science & Medicine,* 96 (C), 60-68.

APPENDICES

APPENDIX A

Visual Basic for Applications code to calculate travel times from each of the 460 postcode areas until each of the 459 other postcode areas

```
Private Sub CommandButton1_Click()

 Set oApp = CreateObject("MapPoint.Application")
 oApp.Visible = False
 Set objMap = oApp.NewMap
 Set ObjRoute = objMap.ActiveRoute

 For i = 2 To 211141
 szZip1 = Worksheets("Sheet1").Cells(i, 1)
 szZip2 = Worksheets("Sheet1").Cells(i, 2)

 ObjRoute.Clear
 ObjRoute.Waypoints.Add objMap.FindAddressResults(, , , , szZip1).Item(1)
 ObjRoute.Waypoints.Add objMap.FindAddressResults(, , , , szZip2).Item(1)
 ObjRoute.Calculate

 Worksheets("Sheet1").Cells(i, 3) = ObjRoute.DrivingTime

 Next i

End Sub
```

APPENDIX B

The 168 hospitals of continental Portugal in December 2011

Hospital	Group	Private/Public
British Hospital Lisbon XXI	Galilei Saúde	Private
Casa de Saúde da Associação de Socorros Mútuos dos Empregados no Comércio e Indústria	-	Private
Casa de Saúde da Boavista	-	Private
Casa de Saúde de Amares, Lda.	-	Private
Casa de Saúde de Guimarães	AMI	Private
Casa de Saúde de São Lázaro, Lda.	-	Private
Casa de Saúde São Mateus - Hospital Particular	-	Private
Cento Clínico da Amadora (Hospital da Luz)	Espírito Santo Saúde	Private
Centro de Ambulatório de Amarante (Clipóvoa - Clínica de Amarante)	Espírito Santo Saúde	Private
Centro de Ambulatório de Cerveira (Clipóvoa - Clínica de Cerveira)	Espírito Santo Saúde	Private
Centro de Ambulatório do Porto (Clipóvoa - Clínica do Porto)	Espírito Santo Saúde	Private
Clínica CUF Alvalade	José de Mello Saúde	Private
Clínica CUF Belém	José de Mello Saúde	Private
Clínica CUF Cascais	José de Mello Saúde	Private
Clínica CUF Torres Vedras	José de Mello Saúde	Private
Clínica de Montes Claros, Lda.	-	Private
Clínica de Paços de Ferreira	Trofa Saúde	Private
Clínica de Santo António - CLISA - Amadora	-	Private
Clínica de São Cristóvão - Associação de Socorros Mútuos de Empregados no Comércio de Lisboa	-	Private
Clínica de Todos-os-Santos	-	Private
Clínica do Campo Grande	Galilei Saúde	Private
Clínica Europa, Serviços Médicos, S.A.	-	Private
Clínica Médico-Cirúrgica de Santa Tecla, Lda.	-	Private
Clínica Parque dos Poetas (Hospital da Luz)	Espírito Santo Saúde	Private
Clínica Particular de Barcelos, Lda.	CESPU	Private
Clínica São João de Deus - CTD - Centro de Tratamento de Doentes, Lda.	-	Private
Clinigrande - Clínica da Marinha Grande, Lda.	-	Private
Clipóvoa - Hospital Privado (Clipóvoa)	Espírito Santo Saúde	Private

Hospital	**Group**	**Private/Public**
Cliria - Centro Médico de Águeda	Espírito Santo Saúde	Private
Cliria - Clínica de Oiã	Espírito Santo Saúde	Private
Cliria - Hospital Privado de Aveiro	Espírito Santo Saúde	Private
Fundação Nossa Senhora da Guia	-	Private (not for profit)
Galilei Saúde - Unidade Alcobaça	Galilei Saúde	Private
Galilei Saúde - Unidade de Leiria	Galilei Saúde	Private
Galilei Saúde - Unidade de Pombal	Galilei Saúde	Private
Hospitais da Universidade de Coimbra	Centro Hospitalar e Universitário de Coimbra, EPE	Public
Hospital Amato Lusitano - Castelo Branco	Unidade Local de Saúde de Castelo Branco, EPE	Public
Hospital Arcebispo João Crisóstomo - Cantanhede	-	Public
Hospital Bernardino Lopes de Oliveira - Alcobaça	Centro Hospitalar do Oeste Norte	Public
Hospital Conde de Bertiandos - Ponte de Lima (Centro Hospitalar do Alto Minho)	Unidade Local de Saúde do Alto Minho, EPE	Public
Hospital Conde de São Bento - Santo Tirso	Centro Hospitalar do Médio Ave, EPE	Public
Hospital CUF Descobertas	José de Mello Saúde	Private
Hospital CUF Infante Santo	José de Mello Saúde	Private
Hospital CUF Porto	José de Mello Saúde	Private
Hospital Curry Cabral, EPE - Lisboa	-	Public
Hospital da Arrábida	Espírito Santo Saúde	Private
Hospital da Boavista	HPP Saúde/CGD	Private
Hospital da Confraria de Nossa Senhora da Nazaré	-	Private (not for profit)
Hospital da Covilhã	Centro Hospitalar Cova da Beira, EPE	Public
Hospital da Cruz Vermelha Portuguesa - Lisboa	-	Private (not for profit)
Hospital da Fundação Aurélio Amaro Diniz - Oliveira do Hospital	-	Private (not for profit)
Hospital da Luz	Espírito Santo Saúde	Private
Hospital da Misericórdia de Évora	Espírito Santo Saúde	Private
Hospital da Misericórdia de Paredes	CESPU	Private
Hospital da Misericórdia de Sangalhos	HPP Saúde/CGD	Private
Hospital da Senhora da Oliveira - Guimarães	Centro Hospitalar do Alto Ave, EPE	Public
Hospital da Venerável Irmandade de Nossa Senhora da Lapa - Porto	-	Private (not for profit)

Hospital	**Group**	**Private/Public**
Hospital da Venerável Irmandade de Nossa Senhora do Terço e Caridade - Porto	-	Private (not for profit)
Hospital da Venerável Ordem Terceira de Nossa Senhora do Carmo - Porto	-	Private (not for profit)
Hospital da Venerável Ordem Terceira de São Francisco - Porto	-	Private (not for profit)
Hospital da Venerável Ordem Terceira de São Francisco da Cidade - Lisboa	-	Private (not for profit)
Hospital de Braga	-	Public
Hospital de Cândido Figueiredo - Tondela	Centro Hospitalar Tondela-Viseu, EPE	Public
Hospital de Cascais	HPP Saúde/CGD	Private
Hospital de Chaves	Centro Hospitalar de Trás-os-Montes e Alto Douro, EPE	Public
Hospital de Dia da Maia	Trofa Saúde	Private
Hospital de Dia de Vila Nova de Famalicão	Trofa Saúde	Private
Hospital de Eduardo Santos Silva (Unidade I) - Vila Nova de Gaia	Centro Hospitalar de Vila Nova de Gaia/Espinho, EPE	Public
Hospital de Egas Moniz - Lisboa	Centro Hospitalar de Lisboa Ocidental, EPE	Public
Hospital de Faro, EPE	-	Public
Hospital de Jesus - Venerável Ordem Terceira da Penitência de S. Francisco, a Jesus	-	Private (not for profit)
Hospital de Lamego	Centro Hospitalar de Trás-os-Montes e Alto Douro, EPE	Public
Hospital de Nossa Senhora do Rosário - Barreiro	Centro Hospitalar Barreiro Montijo, EPE	Public
Hospital de Saint Louis	Galilei Saúde	Private
Hospital de Santa Luzia - Viana do Castelo	Unidade Local de Saúde do Alto Minho, EPE	Public
Hospital de Santa Maria - Lisboa	Centro Hospitalar Lisboa Norte, EPE	Public
Hospital de Santa Maria de Faro	HPP Saúde/CGD	Private
Hospital de Santa Maria	Galilei Saúde	Private
Hospital de Santa Marta - Lisboa	Centro Hospitalar de Lisboa Central, EPE	Public
Hospital de Santarém, EPE	-	Public
Hospital de Santiago	Espírito Santo Saúde	Private

Hospital	Group	Private/Public
Hospital de Santo André - Leiria	Centro Hospitalar de Leiria-Pombal, EPE	Public
Hospital de Santo António dos Capuchos - Lisboa	Centro Hospitalar de Lisboa Central, EPE	Public
Hospital de São Bernardo - Setúbal	Centro Hospitalar de Setúbal, EPE	Public
Hospital de São Camilo	Hospital Particular do Algarve	Private
Hospital de São Francisco Xavier - Oeiras	Centro Hospitalar de Lisboa Ocidental, EPE	Public
Hospital de São Gonçalo de Lagos	HPP Saúde/CGD	Private
Hospital de São João - Porto	Centro Hospitalar de São João, EPE	Public
Hospital de São José - Fafe	Centro Hospitalar do Alto Ave, EPE	Public
Hospital de São José - Lisboa	Centro Hospitalar de Lisboa Central, EPE	Public
Hospital de São Pedro Gonçalves Telmo - Peniche	Centro Hospitalar do Oeste Norte	Public
Hospital de São Sebastião - Santa Maria da Feira	Centro Hospitalar de Entre o Douro e Vouga, EPE	Public
Hospital de Sousa Martins - Guarda	Unidade Local de Saúde da Guarda, EPE	Public
Hospital de Vila Franca de Xira	-	Public
Hospital de Vila Real	Centro Hospitalar de Trás-os-Montes e Alto Douro, EPE	Public
Hospital Distrital Caldas da Rainha (Centro Hospitalar das Caldas da Rainha)	Centro Hospitalar do Oeste Norte	Public
Hospital Distrital da Figueira da Foz, EPE	-	Public
Hospital Distrital da Póvoa de Varzim	Centro Hospitalar da Póvoa de Varzim-Vila do Conde, EPE	Public
Hospital Distrital de Abrantes	Centro Hospitalar do Médio Tejo, EPE	Public
Hospital Distrital de Águeda	Centro Hospitalar do Baixo Vouga, EPE	Public
Hospital Distrital de Bragança	Unidade Local de Saúde do Nordeste, EPE	Public

Hospital	Group	Private/Public
Hospital Distrital de Lagos	Centro Hospitalar do Barlavento Algarvio, EPE	Public
Hospital Distrital de Macedo de Cavaleiros	Unidade Local de Saúde do Nordeste, EPE	Public
Hospital Distrital de Mirandela	Unidade Local de Saúde do Nordeste, EPE	Public
Hospital Distrital de Pombal	Centro Hospitalar de Leiria-Pombal, EPE	Public
Hospital Distrital de Tomar	Centro Hospitalar do Médio Tejo, EPE	Public
Hospital Distrital de Torres Novas	Centro Hospitalar do Médio Tejo, EPE	Public
Hospital Distrital de Vila do Conde	Centro Hospitalar da Póvoa de Varzim-Vila do Conde, EPE	Public
Hospital Distrital de Vila Nova de Gaia (Unidade II)	Centro Hospitalar de Vila Nova de Gaia/Espinho, EPE	Public
Hospital Distrital São João da Madeira	Centro Hospitalar de Entre o Douro e Vouga, EPE	Public
Hospital Distrital Torres Vedras	Centro Hospitalar de Torres Vedras	Public
Hospital do Barlavento Algarvio - Portimão	Centro Hospitalar do Barlavento Algarvio, EPE	Public
Hospital do Espírito Santo de Évora, EPE	-	Public
Hospital do Fundão	Centro Hospitalar Cova da Beira, EPE	Public
Hospital do Litoral Alentejano, EPE - Santiago do Cacém	-	Public
Hospital do Montijo	Centro Hospitalar Barreiro Montijo, EPE	Public
Hospital Dom Luiz I - Peso da Régua (Centro Hospitalar de Vila Real/Peso da Régua)	Centro Hospitalar de Trás-os-Montes e Alto Douro, EPE	Public
Hospital dos Lusíadas	HPP Saúde/CGD	Private
Hospital Dr. Francisco Zagalo - Ovar	-	Public
Hospital Dr. José de Almeida - Cascais	-	Public

Hospital	Group	Private/Public
Hospital Dr. José Maria Grande - Portalegre	Unidade Local de Saúde do Norte Alentejano, EPE	Public
Hospital Garcia de Orta, EPE - Almada	-	Public
Hospital Geral da Colónia Portuguesa do Brasil (Hospital dos Covões) - Coimbra	Centro Hospitalar e Universitário de Coimbra, EPE	Public
Hospital Geral de Santo António - Porto	Centro Hospitalar do Porto, EPE	Public
Hospital Infante D. Pedro - Aveiro	Centro Hospitalar do Baixo Vouga, EPE	Public
Hospital José Joaquim Fernandes - Beja	Unidade Local de Saúde do Baixo Alentejo, EPE	Public
Hospital José Luciano de Castro - Anadia	-	Public
Hospital Nossa Senhora da Ajuda (Unidade III) - Espinho	Centro Hospitalar de Vila Nova de Gaia/Espinho, EPE	Public
Hospital Nossa Senhora da Assunção - Seia	Unidade Local de Saúde da Guarda, EPE	Public
Hospital Nossa Senhora da Conceição - Valongo	Centro Hospitalar de São João, EPE	Public
Hospital Padre Américo - Vale de Sousa - Penafiel	Centro Hospitalar do Tâmega e Sousa, EPE	Public
Hospital Particular de Lisboa	Trofa Saúde	Private
Hospital Particular de Viana do Castelo, Lda.	-	Private
Hospital Pedro Hispano - Matosinhos	Unidade Local de Saúde de Matosinhos, EPE	Public
Hospital Privado da Boa Nova	Trofa Saúde	Private
Hospital Privado da Trofa	Trofa Saúde	Private
Hospital Privado de Braga	Trofa Saúde	Private
Hospital Privado de Guimarães	AMI	Private
Hospital Professor Doutor Fernando Fonseca, EPE - Amadora	-	Public
Hospital Pulido Valente - Lisboa	Centro Hospitalar Lisboa Norte, EPE	Public
Hospital Santa Cruz - Oeiras	Centro Hospitalar de Lisboa Ocidental, EPE	Public
Hospital Santa Luzia de Elvas	Unidade Local de Saúde do Norte Alentejano, EPE	Public
Hospital Santa Maria Maior, EPE - Barcelos	-	Public

Hospital	Group	Private/Public
Hospital São Gonçalo - Amarante	Centro Hospitalar do Tâmega e Sousa, EPE	Public
Hospital São João de Deus - Vila Nova de Famalicão	Centro Hospitalar do Médio Ave, EPE	Public
Hospital São Miguel - Oliveira de Azeméis	Centro Hospitalar de Entre o Douro e Vouga, EPE	Public
Hospital São Teotónio - Viseu	Centro Hospitalar Tondela-Viseu, EPE	Public
Hospital Visconde Salreu - Estarreja	Centro Hospitalar do Baixo Vouga, EPE	Public
HPA Unidade de Alvor	Hospital Particular do Algarve	Private
HPA Unidade de Faro	Hospital Particular do Algarve	Private
Instituto Cuf Diagnóstico e Tratamento	José de Mello Saúde	Private
Intercir - Centro Cirúrgico de Coimbra, S.A.	-	Private
Montepio Rainha Dona Leonor - Associação Mutualista	-	Private (not for profit)
Sanfil - Casa de Saúde de Santa Filomena, S.A.	-	Private
Santa Casa da Misericórdia da Mealhada	-	Private (not for profit)
Santa Casa da Misericórdia da Póvoa de Lanhoso	-	Private (not for profit)
Santa Casa da Misericórdia de Alvaiázere	-	Private (not for profit)
Santa Casa da Misericórdia de Benavente	-	Private (not for profit)
Santa Casa da Misericórdia de Esposende	-	Private (not for profit)
Santa Casa da Misericórdia de Fão	-	Private (not for profit)
Santa Casa da Misericórdia de Felgueiras	-	Private (not for profit)
Santa Casa da Misericórdia de Lousada	-	Private (not for profit)
Santa Casa da Misericórdia de Marco de Canaveses	-	Private (not for profit)
Santa Casa da Misericórdia de Riba de Ave	-	Private (not for profit)

Hospital	Group	Private/Public
Santa Casa da Misericórdia de Vila do Conde	-	Private (not for profit)
Santa Casa da Misericórdia de Vila Verde	-	Private (not for profit)
Santa Casa da Misericórdia do Entroncamento	-	Private (not for profit)
Santa Casa da Misericórdia do Porto - Hospital da Prelada	-	Private (not for profit)

APPENDIX C

Competition changes after the introduction of an artificial supply point

Postcode area	District	HHI		HHI change	Highest market share	Dominance threshold
		Before	After			
2100	Santarém	9367.78	6810.56	-2557.22	80.9%	16.0%
7080	Évora	2242.53	3038.01	795.49	49.7%	36.4%
2985	Setúbal	2427.63	1973.64	-453.99	29.7%	42.3%
7040	Évora	4656.47	4235.07	-421.40	50.8%	29.0%
2120	Santarém	3512.50	3262.72	-249.77	53.4%	34.4%
2965	Setúbal	1970.89	1774.97	-195.92	30.8%	42.9%
2130	Santarém	2695.49	2507.34	-188.16	43.8%	38.4%
7000	Évora	5040.20	4887.44	-152.76	59.1%	25.6%
7580	Setúbal	1569.42	1420.89	-148.53	21.0%	46.3%
2950	Setúbal	1629.58	1538.03	-91.56	25.5%	44.3%
2900	Setúbal	1694.41	1615.79	-78.63	24.9%	44.7%
2910	Setúbal	1675.63	1597.03	-78.59	22.9%	44.8%
2955	Setúbal	1828.71	1777.65	-51.06	30.2%	42.5%
2975	Setúbal	1552.05	1515.49	-36.57	25.1%	43.9%
2830	Setúbal	1530.47	1494.94	-35.53	24.8%	43.9%
2840	Setúbal	1351.12	1322.44	-28.68	23.3%	45.0%
2860	Setúbal	1544.23	1517.76	-26.47	25.2%	43.7%
2870	Setúbal	1641.68	1618.96	-22.72	29.6%	43.1%
2835	Setúbal	1804.50	1784.14	-20.36	29.3%	42.3%
2890	Setúbal	1551.99	1536.60	-15.39	26.6%	43.6%
2805	Setúbal	1188.00	1173.41	-14.59	21.9%	46.2%
2845	Setúbal	1359.54	1345.27	-14.27	23.7%	45.0%
2865	Setúbal	1405.85	1392.88	-12.97	24.6%	44.6%
2925	Setúbal	1644.23	1632.54	-11.69	26.1%	43.5%
2135	Santarém	2186.54	2178.22	-8.32	32.0%	40.0%
2825	Setúbal	1209.84	1202.22	-7.63	22.8%	45.9%
2800	Setúbal	1198.67	1191.22	-7.45	22.6%	46.0%
1990	Lisboa	1294.85	1288.03	-6.82	22.3%	46.0%
2815	Setúbal	1208.40	1203.34	-5.05	22.8%	45.9%
2600	Lisboa	1504.90	1499.90	-5.00	29.3%	44.6%
2810	Setúbal	1193.16	1190.77	-2.39	22.6%	46.0%
1500	Lisboa	1209.06	1207.08	-1.99	21.2%	46.6%
1885	Lisboa	1305.60	1304.21	-1.39	22.6%	45.9%
2620	Lisboa	1327.55	1326.20	-1.35	21.8%	45.6%
1750	Lisboa	1296.38	1295.17	-1.21	22.1%	46.0%
7570	Setúbal	3909.17	3908.24	-0.93	46.6%	31.5%
2690	Lisboa	1304.36	1303.59	-0.77	22.2%	46.0%

Postcode area	District	HHI		HHI change	Highest market share	Dominance threshold
		Before	After			
2675	Lisboa	1305.39	1304.84	-0.55	21.7%	45.8%
2685	Lisboa	1294.77	1294.18	-0.59	22.4%	46.1%
1700	Lisboa	1277.70	1277.13	-0.58	22.4%	46.1%
1070	Lisboa	1194.96	1194.46	-0.50	21.6%	46.5%
1800	Lisboa	1282.53	1282.07	-0.47	22.4%	46.0%
2695	Lisboa	1291.69	1291.44	-0.25	22.0%	46.1%
2610	Lisboa	1213.30	1213.18	-0.12	21.4%	46.5%
2855	Setúbal	1374.75	1374.67	-0.08	25.1%	44.9%
2795	Lisboa	1208.92	1208.87	-0.05	21.8%	46.6%
2050	Lisboa	2257.67	2257.65	-0.02	43.5%	40.0%
2630	Lisboa	1588.79	1588.79	-0.01	30.1%	43.9%
1300	Lisboa	1190.99	1190.98	-0.01	23.0%	46.3%
1675	Lisboa	1299.21	1299.20	-0.01	21.6%	45.9%