U.PORTO

FEP FACULDADE DE ECONOMIA
UNIVERSIDADE DO PORTO

Space impact on retailers' performance – A study of a Portuguese grocery retailer

Diogo Cardoso Campos

Dissertation
Master in Management

Supervised by
Professor Paulo Teles
Professor Francisco Vitorino Martins

2018

Acknowledgments

I would like to deeply thank Professor Paulo Teles for the high degree of professionalism and dedication shown in this mentorship, being available all the time to help in the making of this dissertation. I am also grateful to Professor Vitorino for the valuable contribution and knowledge shared.

To Mário I would like to thank for all the support provided and for the patience every time I have bothered you asking for feedback. To André and Patricia I am grateful for the knowledge shared throughout this first year in Sonae, for the flexibility granted and for the trust you have placed in me.

To my parents and sister, I thank for the unconditional love given to me in every single stage of my life. Thank you for supporting me in every decision taken and thank you for making my life easier and happier.

To my older friends – Meleiro, Marta, Tadeu, Sofia, Forte, Luisa – thank you for understanding my absence in many situations, thank you for your patience and love. Thank you for being present when I needed the most. Regardless of the distance and time we know that we always can count on each other and it is comforting to have such good friends.

To my newer friends – M, Mafalda, Rita, Dani, Maçaira, Carlos – thank you for everything. I never thought that I would make such good friends in my first job experience. We have experienced a lot together in the last year and I hope this friendship holds in the future. Thank you for all the support, advices and for the patience and love shown even in these last darker months.

Abstract

Purpose – This work aims at understanding and measuring the space impact on retailers' performance, i.e., space-elasticity, as well as the influence of other attributes related to trade area characteristics on the sales-space relationship. Furthermore, it aims at fitting appropriate models to estimate retailers' sales revenues at the store and business unit level. These objectives have the purpose to support retailers concerning space allocation and site location decisions and store and business unit performance assessment.

Methodology – Regression models in panel datasets spanning monthly information in a year of analysis were applied. Most information was provided by a Portuguese leading grocery retailer.

Findings – We found that space has a strong positive correlation with stores' and business units' sales revenues. This relationship presents a high degree of stability as is not influenced by trade area characteristics. We applied multiple regression models with the purpose of estimating store and business unit sales revenues based on the attributes that define stores' sales potential.

Research limitations – The fitted models are static and consider sales revenues as the only performance measure. Moreover, business units' specific explanatory variables were not considered.

Originality and value – By using a Geographical Information System (GIS), we have overcome a limitation identified in prior studies related to trade area measurement which is required to calculate store sales potential accurately. Therefore, this is the first approach fitting regression models to estimate sales revenues considering variables related to store sales potential. Moreover, it is the first study estimating the sales-space relationship at the store level using panel data which enables the identification of store and period specific effects, important elements to help retailers in decision-making. Finally, this is the first study to discuss the sales-space relationship at the business unit level.

Keywords: Business Unit, Performance Assessment, Retail Location Decision, Regression Models, Space Allocation Management, Space-Elasticity, Store, Trade Area.

JEL Codes: C23, C33, M31

Resumo

Objetivo – Esta investigação visa perceber e medir o impacto do espaço na performance dos retalhistas, isto é, a elasticidade-espaço, bem como medir o impacto que diversos fatores que caracterizam a área de influência da loja exercem nesta relação entre espaço e vendas. Além disso, pretendemos aplicar modelos de regressão para estimar as vendas de lojas e unidades de negócio tendo em conta as características que definem o potencial de vendas das lojas. Estes objetivos visam apoiar os retalhistas em decisões relativas à alocação de espaço e à escolha da localização de novas lojas, bem como no processo de avaliação de lojas e unidades de negócio.

Metodologia – Aplicaram-se modelos de regressão em bases de dados em painel que contêm informação mensal respeitante a um ano de análise. A maioria destes dados foram fornecidos por um retalhista alimentar português.

Conclusões – Foi provada a existência de uma correlação positiva significativa entre o espaço e as vendas das lojas e unidades de negócio. Concluiu-se que esta relação apresenta um elevado grau de estabilidade, uma vez que não é influenciada pelas caraterísticas da área de influência das lojas. Os modelos aplicados permitiram estimar as vendas de lojas e unidades de negócio com base nas caraterísticas que definem o potencial de vendas de cada loja.

Limitações – Os modelos estimados são estáticos. Foram apenas consideradas vendas como medida de performance. Não foram considerados atributos específicos das unidades de negócio.

Originalidade e Valor – A utilização de sistemas de informação geográfica permite a mensuração das áreas de influência das lojas. Esta abordagem é a primeira a aplicar modelos de regressão com o intuito de estimar as vendas das lojas com base nas caraterísticas que definem o seu potencial de vendas. Para além disso, é também o primeiro estudo a usar dados em painel na estimação da elasticidade-espaço para as lojas, bem como a estudar esta variável ao nível das unidades de negócio.

Palavras-Chave: Área de Influência, Avaliação de Performance, Decisões de Escolha de Localização, Elasticidade-Espaço, Gestão de Espaço, Loja, Unidade de Negócio.

Contents

List of Tables

Chapter 1

Introduction

This study aims at contributing to a research field that has focused on the space impact on retailers' performance.

Concerning this issue, our research will address the space allocation problem that grocery retailers are currently facing through an empirical analysis that, supported on a dataset provided by a leading Portuguese grocery retailer, aims at measuring the space impact on stores' and business units' performance.

Before getting into the empirical analysis it is relevant to explain in detail:

- the context that is shaping the Portuguese grocery retail and its contribution to the intensification of the space allocation problem;

- space allocation processes and their respective problems;

- the objectives of the dissertation;

- the relevant contributions of previous academic research;

The first three issues are addressed in the introductory chapter, while the last is discussed in chapter 2.

1.1. An overview of the Portuguese Grocery Retail Sector

The Portuguese grocery retail sector is characterized by an increasing number of large grocery retailers that, through ambitious expansion plans and aggressive promotions policies, are struggling to increase market share in order to achieve a dominant position in the market.

These efforts aim to guarantee future sustainability in a highly competitive environment that is leading to reduced levels of profitability.

For this reason, retailers are actively engaged in achieving efficiency gains to develop long-term competitive advantages that will enable them to strengthen their position in the market. To this end, they are implementing a set of activities aiming to build up new processes and to develop new systems to manage their resources in a more efficient way.

One of the most important resources owned by retailers is the stores' space where the assortment is displayed to consumers. Space is limited to allocate all the available products offered for sale. Since these sales have a significant impact on retailers' performance it is important to manage space more efficiently, making efforts to improve space allocation decision-making processes.

Therefore, under this competitive context, it is increasingly important to acknowledge space as a resource to be managed more efficiently. If so, retailers can improve consumers' satisfaction and stores' performances.

1.2. Space Allocation Management – Processes and Problems

Space allocation decisions must be taken from the moment a location is identified as a future potential store, to the daily space allocation of products on the shelves of existing stores.

The decision of determining the space to allocate to a store is usually supported by an analysis of performance indicator ratios complemented by an analogue approach. The analogue approach identifies stores belonging to the same store format, with similar characteristics and operating under similar environments. It is important that this decision must be supported by a competent analysis, since it cannot be reversed easily and since changing the space allocated to a store requires a significant investment. Hence, the analysis must guarantee that the space allocated to a store is fit with its sales potential. If this is not the case, a retailer may incur in a substantial opportunity cost caused by the space cost such as rents or general expenses, if a store is over spaced, or by lost sales, if a store is under spaced.

On the other hand, from time to time, stores need to be revamped, either because they are becoming obsolete and the infrastructures and equipment need to be replaced or to implement a new store atmosphere and image to address new customers' needs. Since store revamps require a serious investment, it is a moment when an opportunity arises to reassess stores' space and its fit to sales potential and to adjust it if the investment is expected to be profitable for the retailer.

Space is a very expensive resource since it entails real estate costs, general expenses, investment in equipment, and other operational costs. Therefore, if a store sells 3,000 € per square meter (sqm), while the average space productivity of the retailers' stores is 10,000 € per sqm, it is intuitive to consider that the store is over spaced, and could sell the same with less space. In turn, this reduction would not only lead to a better productivity, but also to decreasing costs that would improve store profitability.

The importance of these decisions has led retailers to improve their processes of collecting and processing the relevant data supporting this type of analysis, since their output has an effective impact on retailers' performance.

In addition to the decision of defining store space allocation, there are other important decisions regarding the space allocation processes that retailers must carry out. In fact, it is important to determine the space to allocate to business units, categories of products, products and brands. These decisions are more flexible than the one described above. However, they also imply costs for the firm and cannot be reversed as easily. Indeed, retailers cannot adjust regularly the space allocated to product categories since these decisions involve layout modifications, replacement of equipment where products are displayed, labor utilization. Furthermore, it would create confusion in clients' minds, since they are accustomed to the product location in the stores layout.

These space allocation decisions usually result from a set of interactions between several stakeholders that actively participate in the process:

- Operational and store managers take decisions regarding the space to allocate to products within stores, based on their experience, operational concerns, and with the support of performance indicators analysis;

- Specialized teams of analysts perform detailed analysis to stores' performance indicators ratios. Based on computing norms and internal benchmarks, calculated through comparisons among stores and business units/categories/products, business units/categories/products productivity and profitability are measured and their space fit with the potential of sales is reassessed. The results of this analysis are space adjustment recommendations;

- Specialized teams of space management experts, with the support of the aforementioned analysis, design stores' layout and shelves planograms, considering, layout, infrastructure, equipment and operational concerns.

These space allocation processes are carried out with the purpose of improving retailers' performance. However, there are obstacles preventing retailers to achieve better results:

- Retailers cannot quantify accurately the impact of changing the space in a store, business unit or product category. Considering a store that sells 10 million € annually and that has 2,000 sqm, the space productivity ratio is 5,000 € / sqm. However, retailers cannot measure precisely the impact on store sales revenues resulting from the store space adjustment, since ratios assume constant returns and it would not be reasonable to consider that this store would sell 20 million € if the space increased to 4,000 sqm.

- Retailers cannot aggregate all store attributes influencing performance into relevant store clusters. In the knowledge that store performance variations are a result of their attributes differences, retailers aggregate stores with similar attributes, i.e., sales area, location and operating under similar environments into store clusters. The objective is to create groups of comparable stores, to perform comparative analysis to performance indicator ratios between similar stores excluding the effects of the store attributes in their sales. However, the criteria used to determine these clusters is too simplistic. On the one hand, simplicity is required to allow the creation of relevant groups of comparable stores. However, on the other hand, disregarding some relevant attributes affecting stores' performance could lead to biased conclusions and space misallocations.

Despite the increasing importance that this theme represents for retailers' sustainability, there are few studies in the literature that have been focusing in assessing the impact of space on stores' performance.

Therefore, considering:

- the context of the Portuguese grocery sector that is leading retailers to recognize the increasing importance of managing the space allocation processes more efficiently;
- the obstacles that retailers are facing to improve the effectiveness of their space allocation processes;
- the opportunity to add contributions to this research field.

This study aims at providing a greater understanding of the space allocation process and of the relationship between space and stores and business units' performance.

1.3. Main Objectives

This work aims at providing support to space allocation decision-making processes:

- concerning the definition of the store and business unit space allocation;
- concerning the assessment of store and business units' space performance
- concerning the estimation of store and business unit sales revenues;

Bearing these objectives in mind, the expected outcome is:

- to develop econometric models capable of measuring the relationship between stores' and business units' space and sales revenues;
- to understand and measure the impact of store relevant attributes on sales revenues and on the sales-space relationship;
- to identify opportunities of space adjustments at the store and business unit level.

1.4. Structure

This work is structured into five chapters. In Chapter 1, the research theme was presented, a brief description of the environment that is shaping the Portuguese grocery retail sector was made and the obstacles that companies are facing to improve retailers' performance through the space allocation processes were described.

In Chapter 2, the literature review is discussed on the relevant studies focusing on the space allocation processes and on the impact of space and other store attributes on stores' performance. A greater emphasis is placed on studies covering the sales-space relationship at the store and the business unit/category level.

In Chapter 3, the empirical analysis of this work is carried out. This chapter is divided into three different sections. Firstly, the methodology that guides the analysis is presented. Secondly, the datasets supporting the analysis are described, together with their collection and processing methods. Finally, the methodology is applied to the datasets and the results of the study and their managerial implications are discussed.

In chapter 4, the study main conclusions and their central managerial implications are summed up and the limitations of the work are addressed.

Chapter 2

Literature Review

In this chapter, the literature review is carried out and is divided into five different but complementary sections.

In section 2.1., the importance of assessing store performance for retailers is discussed and the factors found to affect such performance are presented. In section 2.2., the space allocation management decisions retailers must take are presented and those most relevant for this study are identified. In section 2.3., the importance of retail site location and store space allocation decisions are addressed, as well as the methods used by retailers to support such decisions. Furthermore, the importance of accurately estimating stores sales revenues is addressed and the essential role of the process of measuring such trade area in the estimation process is discussed. In section 2.4., the importance of space in order to estimate store sales revenue potential for new and existing stores is explained. Then, the most common methods used by retailers to define the store space allocation are presented as well as their limitations in assessing sales-space relationship. Finally, in section 2.5., the most relevant studies are presented addressing the relationship between retailers' space and sales revenues developing statistical models to understand and measure this relationship.

2.1. Store Performance Assessment

This work is part of the research field that has been focusing on studying the factors influencing store performance.

Store performance evaluation has been a concern for retailers for a long time. This concern has been rising due to the highly competitive conditions shaping the retail sector. In the last decades, both practitioners and academics have been developing new methods to understand and measure the impact of several factors on store performance, generating important contributions to support managerial decision-making.

The academic field has studied these factors globally to act upon them locally (Kamakura and Kang, 2007). Each store has different attributes and operates under different environments and these specific characteristics must be taken into consideration in performance assessment. In order to assess accurately store performance there is a need to collect relevant information on its intrinsic attributes. These attributes can be divided into two different groups: internal and external (Silva and Cardoso, 2005).

Internal attributes are related to store characteristics generally controlled and defined by retailers (Nilsson *et al.*, 2015). They are associated with the physical store space characteristics and with aspects of the marketing-mix (product, price, place and promotion) determined by retailers to create a value proposition capable of attracting consumers.

Regarding the physical store characteristics, store sales area influences performance (Castro, 2011) since its size is linked with assortment' width and depth and the service level offered to costumers (Sinaglia, 1997). Furthermore, a larger sales area has a positive influence on clients' visual perception, leading consumers to buy products with greater exhibition space more frequently (Philipps and Bradshaw, 1993). On the other hand, the larger the space allocated to a product the less stock-out is observed, positively influencing its sales performance (Borin *et al.*, 1994). Other attributes such as the number of employees', number of checking counters, parking places, store's layout design (Fernie *et al.*, 2015) or the existence of self-scanning devices (Cho and Fiorito, 2010) were also found to affect store performance. However, the causality effect of these attributes is not clear (Davies, 1977), i.e., most of these factors are defined using an estimation of the store sales potential. Moreover, their majority can be adjusted over time based on the store actual performance. In fact, the number of employees is adjusted by taking into account the expected monthly sales. In Christmas holidays, part-time employees are hired to reinforce operational teams since a sales increase during this period is expected every year. Therefore, it may be considered that these factors are influenced by store performance and not the opposite.

Moreover, the literature also highlighted the importance of store internal factors such as visibility or accessibility as having an impact on stores' performance (Silva and Cardoso, 2005). Nilsson *et al.* (2015) observed that accessibility by car was the most important store attribute when consumers had to choose a store to shop. Even though such attributes do not have a high degree of controllability, since they often depend on decisions taken by external entities, they must be taken into consideration in the moment of choosing a location to open a store.

Concerning the aspects of the marketing-mix, while some studies have proved that product quality is the most important attribute in consumers' store choice decision (Reutterer and Teller, 2009; Wong and Dean, 2009), others have emphasized the greater importance of product assortment (Briesh *et al.*, 2008), product supply (Carpenter and Moore, 2006) or price level (Mitchell and Harris, 2005). Other marketing-mix aspects as stores' atmosphere (Hoffman and Turley, 2002), loyalty programs (Bridson *et al.*, 2008), store promotion policies (Dhar and Hoch, 1997), service quality (Reutterer and Teller, 2009) or employees' friendliness (Woodside and Trappey, 1992) were also found to affect stores' performance, and must be taken into consideration in the marketing-mix strategy.

On the other hand, external attributes are related with the trade area characteristics of the store location and their degree of controllability is very low (Nilsson, 2015).

Stores are installed in locations where they must compete with other stores – either belonging to the same retailer or to competitors – to attract potential clients. Stores' attractiveness to potential clients is geographically limited. Hence, it is important to measure stores' trade area that defines the geographical limits of the store influence in the market where it is located. There are two important factors related to store's trade area: competition and market potential.

Concerning competition, there are contradictory results on its impact on store performance. On the one hand, it is intuitive to consider that as the number and the dimension of competitors increases in the trade area of a store, its performance is negatively affected. Clarke *et al.* (2003) remarked that stores performed poorly in over-shopped areas. On the other hand, since retailers usually locate their stores in areas with greater market potential, i.e., in highly-populated areas where consumers have higher purchasing power, it is observed that best performing stores are located in areas with stronger competition (Silva and Cardoso, 2005). However, is intuitive to consider that a store operating in a specific trade area is negatively affected by the introduction of new competitors if the market potential is kept constant.

Market potential is defined by the number of residents, their socio-economic and demographic characteristics and by the potential clients' preferences. These factors are essential to assess stores' sales potential which is important to support the retail site location decision (Silva and Cardoso, 2005). In fact, this assessment prevents retailers from opening new stores in locations with little market potential which would result in significant losses. Besides, store sales potential also depends on the characteristics of the clients living in its trade area. Clients' characteristics were found to be significant to determine stores' performance, since clients choose a store offering a value proposition that meets their specific needs and preferences (Jones, 1999). A number of studies have proved the influence of the population socio-demographic and economic characteristics such as age, gender or income on clients' store choice decision (Carpenter and Moore, 2006; Prasad and Aryasri, 2011).

2.2. Space Management Decisions

In the previous section, multiple factors that help to explain stores' performance were presented. It is necessary to identify and understand these factors to develop innovative processes and systems able to help managers in their decision-making process. There are multiple decisions made by several managers responsible for steering different departments aiming to improve retailers' performance. These decisions are related with marketing, logistics, product development, etc. and their overall impact is expected to enhance retailers' performance. However, depending on the processes being addressed, it is important to focus in the most relevant factors affecting them. This study aims at supporting space management decisions that are part of the broad marketing field of research. More specifically, it aims at providing important inputs for the store space allocation definition and the business unit space allocation decision.

According to Davidson *et al.* (1984) there are three different types of space management decisions: strategic, tactic and operational. Strategic decisions relate to the choice of a store location – the retail site location decision – and to the definition of store design – definition of store sales area and other operational characteristics. Tactic decisions are related to the definition of store layout and the necessary equipment, of the assortment to display in stores and the space allocation to business units / product categories.

Operational decisions relate to the micro-space and the space allocation to products and brands.

Therefore, this work focuses on space management strategic decisions – store space allocation definition – and tactic decisions – space allocation to business units.

2.3. Retail Site Location and Trade Area Measurement

Retail site location decisions are very important for retailers and play a key role in determining the success of a store (García-Palomares, *et al.* 2012). Choosing a new store location involves serious financial and corporate risks to the company (Alarcón, 2011). Therefore, it is very important to perform a solid sales revenues estimation to ensure that a new store has the expected performance in its trade area (Hernández & Bennison, 2000).

The most common method that has been used over time by retailers and consulting firms to support this decision is the analogue approach, developed in 1932 by William Applebaum (Davies, 1977). Using this approach, retailers estimate the sales potential of new stores, comparing them to other stores presenting similar internal attributes – sales area, visibility, etc. – and operating in similar environments in what concerns their location and competition (Davies 1977; Rogers and Green, 1979).

Retailers' goal is to estimate store sales potential through a comparison with similar stores operating in similar trade environments (Clarke *et al.* 2003). However, this complex decision is usually oversimplified. The first step is to estimate store market penetration of existing stores in their trade areas. Then the relationship between store performance, known market factors and store characteristics – the analogues – is analyzed. These analogues are then extrapolated to forecast sales revenues of potential new sites (Rogers and Green, 1979).

This approach presents some limitations. The first is related to the reliance on the market analyst expertise, since he must be able to assess and select the appropriate analogues (Rogers and Green, 1979). Second, this analysis assumes the use of "rules-of-thumb" to measure stores' trade area and some important factors.

On the other hand, statistical techniques, such as multiple regression models, have been used to support the analogue approach (Davies, 1977). They can define and measure correlations between store sales revenues and variables within the catchment area that influence performance. These techniques can be quite useful when retailers own a great number of stores, since they can provide a summary of the strength of the factors influencing stores' performance.

However, like the analogue approach, such statistical techniques are not capable of measuring accurately stores' trade area, which is crucial in order to assess stores' market potential. (Davies, 1977; Wood and Tasker, 2008).

The trade area can be defined "as the geographic area in which a retailer attracts customers and generates sales during a specific period" (Roig-Tierno, *et al.*, 2013), and is generally defined through simple "rules-of-thumb" as the three-mile radius or an average ten-minute drive. Silva and Cardoso (2005) developed a regression tree to predict store sales performance using three different rings – 5 minutes, 10 minutes and 15 minutes driving – as potential store trade area.

The difficulty in accurately measuring store trade area might undermine a fitting site-location decision, since the spatial dispersion of both consumers and vendors is important to estimate correctly the market potential of the trade area in which a store is located (Baviera-Puig *et al.* 2011).

The market potential of a trade area depends on two main factors: geodemand and geocompetition (Roig-Tierno, *et al.* 2013). Geodemand can be defined "as the location of customers who purchase a product or a service in a specific market". Geocompetition can be defined as "the location of competitors of a business and delineation of their trade areas in particular markets." Therefore, trade areas establish the boundaries of the stores' influence near potential customers and of the stores' direct competitors (Baviera-Puig, 2012). A trade area precise estimation enables retailers to accurately calculate store sales potential which is important either to take decisions of "go"/" don't go" in potential stores openings or to select a new store among a list of possible locations. (Wood and Reynolds, 2012).

In recent years, the emergence of Geographical Information Systems (GIS) has contributed to find a more accurate estimation of stores' trade area. Hence, these systems are playing an important role for retail site-location decisions (Church, 2002). They have the capacity to generate spatial representations of geodemographic and retail data (Hernández, 2007). Moreover, they can deal with large quantities of information (Roig-Tierno, *et al.* 2013) being able to support spatial interaction models that can calculate stores' trade area by measuring the relationship between store attractiveness and distance from consumers (Wood and Tasker, 2008). These new techniques are contributing to find a solution to the trade area definition problem and are complementary to the methods described above.

Despite of the important contributions made by these techniques for the retail site location decisions, managers' intuition, experience and "rules of thumb" are still important to be considered alongside quantitative methods to enhance the results of the decision-making process (Wood and Tasker, 2008). In such complex decisions, there are relevant factors influencing store performance that cannot be measured accurately by retailers. In these cases, managers' expertise provides valuable inputs that, together with the results of the quantitative methods, will enable an improvement in retailers' performance.

2.4. The Importance of Space Allocation in Sales Performance

The definition of trade area allows to obtain information on a store's market potential and competition. However, the size of the trade area depends on the store attractiveness, which in turn is linked with the store size. Therefore, in order to estimate a store's sales potential, it is important to consider its space since it influences the size of the trade area and its sales potential. In this section, the importance of space for retailers' performance is described.

Space is a limited resource to allocate the increasing number of products available to offer to costumers (Nogales and Suarez; 2005) and, simultaneously, is one of the most expensive resources owned by retailers (Ramaseshan *et al.*, 2009). Space cost includes real estate expenses – rents and condominium – general expenses as electricity, cleaning, maintenance or investment in equipment in which the products will be displayed to consumer, among others. As space increases, these costs also increase, affecting negatively store profitability. Therefore, space is one of the most important resources managed by retailers and its allocation processes must be properly addressed, either in strategic, tactic or operational decisions (Davidson *et al.*, 1984).

Concerning the definition of store space allocation, either for new or existing stores, it is important to guarantee that the store size is fitted with the market potential to guarantee a high productivity and assure store profitability. Retailers generally define store space through an analogue approach, described in the previous section, and through performance indicator ratio analysis. When a store is decided to open, similar stores, operating under similar environments, are identified. Based on the assessment of performance indicator ratios calculated for these comparable stores, the space to allocate to a new store is defined. This assessment also allows to identify the fit between the space of existing stores and their performance which could lead to sales area adjustments in future revamping.

Since retailers define store sales area with the support of this type of analyses, space can be interpreted as being both a cause and an effect of store sales area. This causality issue remains central in the literature. Desmet and Renaudin (1998) stated that the impact of space in retailers' performance could only be proved through experimentation. However, it is very costly and time-consuming to perform space experimentation at the store or business unit level. In fact, it is unreasonable to think that retailers can change the space allocated to a store or business units several times in order to observe the impact in stores' or business units' performance. For this reason, it is a challenge for retailers to measure the sales-space relationship. Ratio analyses that are performed to support these space allocation decisions assume constant returns which, as explained in chapter 1, is inaccurate to measure this relationship. From this research field, several authors have developed statistical models to measure sales-space relationship – space-elasticity – using cross-sectional data from a high number of stores. Since there is not enough data of the space variability over time of a store, such variability is calculated through the differences observed among stores.

The concept of space-elasticity is crucial to understand and measure the extent of the impact of space on sales performance – the central objective of this study. In the next section the most relevant studies approaching this concept are presented.

2.5. Space-Elasticity – Similar Studies and Methods

Space-elasticity was defined by Yang and Chen (1999) as the ratio of the percentage variation of sales revenues and the percentage variation of sales area and measures the relationship between space area and sales revenues.

Sales-space relationship was found to take the form of a S-Curve. A number of studies have proved that space-elasticity has decreasing marginal returns (Jallais *et al.*, 1993; Desmet and Renaudin, 1998). That is to say that, starting from a certain sales area, the store incremental sales caused by sales area increases are less than proportional. As space increases, the variable costs also increase and may not be covered by incremental sales, yielding losses for the firm.

Therefore, the existence of decreasing marginal returns stresses the importance of measuring space-elasticity to support space allocation decisions. Ratio analyses assume constant returns and is unrealistic to consider that a store selling 10,000 € / sqm would present the same productivity regardless of its sales area.

Space-elasticity enables finding a better forecast of the impact of space on retailers' sales. However, as mentioned in the previous section, there are not enough space changes over time at the store and business unit level to assess the impact of space on performance. Therefore, the variations among stores are considered to explain the differences between store sales revenues.

However, as presented in section 3.1. there are many factors affecting store performance (Silva and Cardoso, 2005) and consequently the sales-space relationship. Consequently, identifying and understanding the impact of the most relevant factors in store performance is required to estimate space-elasticity more accurately.

In section 2.2., the types of space allocation decisions retailers need to take were identified – strategic, tactic and operational.

Most studies addressing the space-elasticity issue are related to operational space allocation decisions, i.e., product and brand space allocation within shelves. On the one hand, it is easier to assess space-elasticity at product level since is not expensive or time-consuming to modify the space allocated to products within shelves. Therefore, the impact of space on products is assessed through experimentations within-store. On the other hand, the development of category management has led retailers and academics to work on space allocation problems at the product level, rather than at aggregate levels (Desmet and Renaudin, 1998). In this respect, many authors (Cox, 1970; Curhan, 1972; Corstjens and Doyle 1981; Abbot and Palekar, 2008) have conducted studies on the impact of space on product performance through experimentation in retailer stores. They managed to calculate shelf-space elasticity, i.e., the impact on sales resulting from changing the space of the product categories, products or brands displayed on shelves. Curhan (1972) found an average value of 0.212 at the product category level while Corstjens and Doyle (1981) calculated a lower average value of 0.086 at the product level. On average, the results of these studies reached a space-elasticity for product categories/products/brands ranging from 0.08 to 0.2 which means that an increase of 100% in category/product/brand space results in an increase in sales revenues between 8% to 20%.

On the other hand, there are not many studies approaching the space allocation issue at an aggregate level. At the store level, Davies (1977) developed a set of simple regression models to quantify the correlation between sales and internal store attributes such as sales area, number of employees, number of cash tills and annual rents paid by the retailer, obtaining an average value of 0.748 for the store space-elasticity. Thurik (1988) performed a study using cross-sectional data that found an average space-elasticity of 0.51 for hypermarkets, considering 68 observations, and 0.68 to supermarkets, considering 121 observations.

Desmet and Renaudin (1998), developed a model to measure the relationship between space and sales revenues at the category level. Space was measured in linear meters. The authors supported the study in a panel dataset spanning monthly data over one year for more than 200 stores belonging to the same French town-center variety store chain. These stores are divided into three different store formats – essential stores, plus stores and standard stores. The use of this type of data allows to consider space variability within the same store over time – period effects – and among different stores – store effects. An econometric model based on a demand function linking the share of sales to the share of space allocated to a product category was applied. The authors mentioned the need to consider other stores – location and competition – and category attributes – width of assortment, price and promotional policies – in addition to the category sales area to fully understand sales revenues variability and to measure the impact of space to this indicator more precisely. However, they failed to incorporate these variables in the model because it was impossible to collect this type of information. Among all store and category attributes, they only have identified the impulse-buying categories to observe the differences of the space impact between this specific group of categories and the other. Results showed that the average value of space-elasticity for all the categories of products was 0.205, meaning that on average the sales revenues of a category from the French retailer increases (decreases) 0.205% when the respective space increases (decreases) 1% However, the results obtained for each category were significantly different since were obtained values for space-elasticity ranging from -0.44 to 0.80. Furthermore, it was proved that impulse-buying categories exhibit higher space-elasticity which supports the hypothesis that space has an impact on sales revenues. Finally, differences between category space-elasticity in each store format were found to be nonsignificant. In this study, cross-elasticities were not considered, i.e., the impact of space changes of a category on the sales revenues of all the others (Corstjens and Doyle, 1981). This decision is justified by the fact that cross-elasticities were found to be significantly weaker than direct space-elasticities.

Castro (2011), has carried out a study aiming at estimating the sales-space relationship at the store and category level. The study was supported on cross-sectional data for one year spanning information on 106 stores divided into three store formats – hypermarkets, supermarkets and convenience stores. Space was measured in square meters. Firstly, the author developed a simple regression model using the least squares method to calculate space-elasticity at the store level and reported an average space-elasticity of 1.21, meaning that, on average, a 1% increase (decrease) of store sales area leads to a 1.21% increase (decrease) in sales revenues. Moreover, differences among store formats were found – an average value of space-elasticity of 0.76 for hypermarkets, 1.16 for supermarkets and 0.53 for convenience stores. In this study, the relevance in considering other store and category attributes was also mentioned. The author considered the price index of the company against its main competitor; the degree of consumer satisfaction; the number of competitors; demographic – population and population density – and economic characteristics of the population in which a store is located. However, this study failed to prove the impact of competition and population variables (number of residents, population density) on store sales revenues which may be explained by the inability to measure store trade area accurately. Adding the variables that were proven to be significant in explaining store sales revenues, a multiple regression model was applied considering four independent variables – sales area, retailer price index, store consumer satisfaction index and purchasing power of the city in which a store is located. The results of this regression showed an average space-elasticity of 0.90 which is lower than that estimated in the first model. This conclusion was also observed for each store format, since space-elasticities of 0.71, 1.12 and 0.40 were found for hypermarkets, supermarkets and convenience stores respectively. This model also measured the impact of the remaining variables introduced in the model. Store consumer satisfaction was found to have an elasticity of 0.59, the price index against the main competitor of 1.77 and the purchasing power of the population living in the city in which a store is located of 0.18, meaning that a 1% increase (decrease) in each index value with all the other variables constant) is expected to increase (decrease) store sales in 0.59%, 1.77% or 0.18% respectively. Later in the study, the impact of space and of the other independent variables mentioned above on sales revenues of product categories was also measured. The results show that space-elasticities vary greatly among categories ranging from -0.24 to 1.84.

After having presented the most relevant contributions that have been made over the last decades regarding the space allocation process, an empirical analysis concerning a Portuguese retailer is conducted in the next chapter.

Chapter 3

Application to a Portuguese Leading Grocery Retailer

Recall that this dissertation aims at understanding and measuring the impact of space on grocery retailer stores and business unit sales revenues. These dimensions define the scope of the analysis and are related with two specific space management activities which are store space allocation and space allocation to business units. The definition of the space to be allocated to a store happens in two different moments. The first and most important is in the retail site location decision, when retailers choose a location to open a new store. The second happens later when a store is set to be revamped. On the contrary, the process of allocating store space to business units is more flexible and can occur more often. In this decision, the broader scope of the stores' assortment that will be displayed to clients is defined. In the following empirical analysis, we propose econometric models with the purpose of measuring the magnitude of the space impact on store and business unit sales revenues. Moreover, they also intend to analyze the impact that other store attributes have on store and business unit performance and on the relationship between space and sales which is central to the analysis.

For these purposes, the analysis is applied to Sonae MC, a Portuguese leading retailer. This chapter is divided into three sections. In section 3.1., the methodology to be used in the analysis is described. In section 3.2., the characteristics of the datasets supporting the analysis are described. In section 3.3., the methodology is applied and the results obtained are presented.

3.1. Methodology

The following empirical analysis is organized into three stages.

In section 3.3.1., an analysis on the relationship between space and sales revenues is performed at the store level. Firstly, a simple regression model (model 1) is developed considering store sales area as the only independent variable explaining store sales revenues. In this stage, all the stores comprised in the dataset are considered regardless of their store format. Then, a regression model (model 2) is also developed taking into account the differences among the three store formats operated by the Portuguese grocery retailer. In this section, the magnitude of the impact of space on store sales revenues is observed and other variables are not considered.

In section 3.3.2., other explanatory factors of store sales revenues besides space are analyzed. These factors are essentially related to the store trade area. Firstly, the reasons are presented for considering such variables. Then, simple regression models for each of these variables are fitted in order to test their influence on store sales revenues. Finally, a multiple regression model (model 3) is fitted comprising sales area and the other independent variables. This model is replicated to consider the differences among store formats (model 4). In this stage, models able to estimate new or existing stores' sales revenues based on their sales area and trade area characteristics are developed. Moreover, the impact that these other variables have on store space-elasticity is observed, comparing the results of this stage with those reported in the first stage.

In section 3.3.3., the scope of the analysis is extended to the business unit level and a similar approach is used. Firstly, simple regression models are estimated for all the aforementioned variables in order to analyze their importance to explain business unit sales revenues, with a particular focus on the sales-space relationship (model 5). Then, a multiple regression model is developed capable of estimating business unit performance based on the values of those variables (model 6). The main objective is again to measure the impact of space on business unit sales revenues either in a simple regression model or jointly with other independent variables.

To implement this methodology, two datasets were collected and processed to support the two dimensions of the analysis – store and business unit level. Most of the information was provided by the mentioned Portuguese grocery retailer. It is important to fully understand the characteristics of these datasets since they affect the methods to be implemented and consequently their results. Therefore, in section 3.2. the datasets, their collection and processing methods and the characteristics of the variables being studied are described.

3.2. Datasets

This section is divided into two subsections. In subsection 3.2.1., the Portuguese leading grocery retailer firm that provided the information supporting this work is introduced. In subsection 3.2.2., the characteristics of the datasets are described.

3.2.1. The Company: Sonae MC

Sonae – Sociedade Nacional de Estratificados was founded in 1959 and is a multinational company managing a diversified portfolio of businesses in retail, financial services, technology, shopping centers and telecommunications.

In its early days, the firm operated exclusively in the wood processing sector. However, in 1983, it formed a joint venture to renew the fragmented distribution and retail business in Portugal in which many small operators were prevailing. In 1985, opened the first Continente hypermarket and that was the moment Sonae Distribuição began.

In the present, Sonae Distribuição became Sonae MC which holds the hypermarket and supermarket chains owned by Sonae Group. Sonae MC owns 260 grocery stores and is the market leader in the Portuguese grocery sector.

Sonae MC operates three different store formats under the umbrella of three different brands:

- Continente (41 stores) is the banner of the hypermarket format comprising stores with larger sales areas. These stores are generally located in the suburbs of highly populated metropolitan areas near large commercial sites which are highly attractive to consumers. They are characterized for displaying a wider assortment both at the category and the product level.

- Continente Modelo (123 stores) is the banner of the supermarket format comprising stores with an average sales area of 2,000 sqm. These stores are either located in the suburbs of large metropolitan areas, in commercial areas, or are located in rural areas. Compared with Continente stores, they offer a less extensive range of categories and products to consumers.

- Continente Bom Dia (96 stores) is the banner of the convenience format comprising stores with smaller sales area. They are located in the center or in residential areas of large cities and aim at offering convenience to consumers. They offer a less extensive range of products placing a greater focus on fresh products that are purchased more often by consumers.

Apart from the grocery sector, Sonae MC also invested in other small retail-related businesses such as bakeries, pharmacies, pet shops and stationery shops. More recently expanded its activity scope into the dentist and esthetic services market.

3.2.2. Dataset Description

To apply the methodology, two datasets were collected and processed, the first at the store level and the second at the business unit level. Their characterization is a requirement to fully understand the empirical analysis and its results.

The first dataset spans monthly information of 192 stores over 12 months, from September 2015 to August 2016, on store sales revenues, sales area and on several store trade area variables. The second dataset contains monthly information of 23 business units, over 12 months, from September 2015 to August 2016, on business unit sales revenues, sales area and the same trade area characteristics as the first dataset. Therefore, these datasets differ in the cross-sectional variable to be studied, i.e. the first dataset supports the analysis at the store level, and the second at the business unit level. Next, the differences between these two cross-sectional variables and the characteristics of all the variables part of both datasets are described.

3.2.2.1. Cross-sectional Dimensions

Stores

Sonae MC divides its store chain into three different store formats named with distinct banners: Continente, Continente Modelo, Continente Bom Dia. This division aims at aggregating stores into groups with similar characteristics to ease the management processes.

The first dataset comprises 192 of the total 260 grocery retail stores held by Sonae MC at the end of 2017

- 40 stores operating under the Continente banner;
- 116 stores operating under the Continente Modelo banner;
- 36 stores operating under the Continente Bom Dia banner.

In this work, stores opened after December 2013 were not considered. This decision ensures the consistency of this dataset since, according to managerial expertise, it takes at least a year for a store to reach its maturity, i.e., its expected sales potential. During the maturation process, a store must standardize operational processes, attract new customers and build trust with new clients. For this reason, at the beginning of their activity, stores have lower sales revenues when compared to similar older stores, making the comparison of their performances unreasonable. Consequently, it has been decided to withdraw the stores mentioned above from the dataset because their inclusion could lead to biased conclusions.

Business Units

Retailers organize their products according to a market structure designed to replicate the way consumers mentally organize the different groups of products displayed in stores. As an example, considering a bottom-up approach, several products such as apple juice, orange juice or lemon juice are organized into a category of products, i.e. juices. This category is then aggregated with other related categories, such as beer and water, into business units, i.e., beverages. These business units include all the products displayed to clients in retailer stores.

Table 1 below shows the 23 business units' part of the commercial structure of the mentioned Portuguese leading retailer.

Business Units	BU description	Business Units	BU description
BU01	Savory	**BU18**	Take Away
BU02	Sweet Savory	**BU19**	Cafeteria
BU03	Drinks	**BU30**	Leisure
BU05	Hygiene and Beauty	**BU31**	Home
BU06	Home Cleaning	**BU33**	Culture
BU07	Frozen	**BU34**	Brico & Auto
BU08	Dairy	**BU35**	Pet and Care
BU11	Butchery	**BU41**	Baby Apparel
BU12	Fishery	**BU42**	Children Apparel
BU13	Cheese and Cold Meats	**BU43**	Women Apparel
BU15	Fruits and Vegetables	**BU44**	Men Apparel
BU16	Bakery		

Table 1: Business units' commercial structure

3.2.2.2. Variable Description

According to the outlined methodology, regardless of the stage or the cross-sectional dimension being studied, there are two main variables in this study: sales revenues is the dependent variable and space, measured by sales area, is the central independent variable. In fact, the sales-space relationship plays a central role in this work.

The sales-space relationship is also analyzed at the store level for each store format operated by the retailer. Thus, the banner of each store format is also a *dummy* variable to be considered.

In the second and third stages of the methodology, other variables are introduced that are expected to influence the dependent variable, i.e., store or business unit sales revenues. The choice of these variables is explained by their importance in assessing store sales revenues potential. As it was already mentioned, this work aims at supporting the space allocation process at the store level where it is essential to estimate the store sales revenues potential, either to new or existing stores, and to guarantee the fit of the store sales area to the expected sales revenues of a new or existing store. Store space definition occurs in two different moments in time, when a store is set to be open or revamped. An accurate store sales revenues estimation requires the definition of the store trade area in order to establish geographical limits for the store attractiveness. A trade area is characterized by its market potential and competition. On the one hand, it is important to estimate its market potential that depends on the number of residents living in the area and their socio-economic characteristics. On the other hand, the store potential sales in the market are limited by the existence of competitors. Therefore, in this study variables characterizing store trade area are considered which is essential to estimate sales revenues which in turn are important to define the space to allocate to a store.

There is a large set of other factors influencing store performance that are not considered in this work. On the one hand, it would be impossible to include all these variables. On the other hand, according to the literature and to Sonae MC directors' managerial expertise, they have a minor importance in explaining store sales revenues variability. Therefore, four variables are considered besides space since together they can help to characterize a store's trade area:

- **Population**: Number of residents living in stores' trade area.

- **Population Density:** Number of residents per square meter living in stores' trade area.

- **Purchasing Power Index:** Reflects the population economic strength which is important to assess people's willingness to spend, helping to support the estimation of the demand side in the trade area.

- **Competition:** Space of the competitors that are already established offering a similar assortment. Provides an insight of the supply side of a store trade area.

These variables are considered together with sales area and, applying a regression analysis, the store sales potential can be estimated for new or existing stores.

The table below displays the main characteristics of the variables to be studied. These characteristics are described in detail next.

Store Attributes		Variable	Description	Type	Unit Measure	Source
Internal Attributes		Store Sales	Monthly reported net sales (sales deducted by VAT and discounts given to clients) per store	numerical - continuous	number / unit	Internal
		Business Unit Sales	Monthly reported net sales (sales deducted by VAT and discounts given to clients) per business unit	numerical - continuous	number / unit	Internal
		Store Sales Area	Monthly sales area per store	numerical - continuous	sqm	Internal
		Business Unit Sales Area	Monthy sales area of a business unit	numerical - continuous	number / unit	Internal
		Store Format	Store banners' name, identifying its store format - "Continente", "Continente Modelo", "Continente Bom Dia"	categorical - nominal	n.a	Internal
External Attributes	City	PPI	Purchasing Power Index per capita in the city where a store is located	numerical - continuous	number / index	National Inst. of Statistics
	Trade Area (TA)	Population	Number of Residents in stores' trade area	numerical - discrete	number / unit	National Inst. of Statistics
		Population Density	Number of Residents per square meter in stores' trade area	numerical - discrete	number / sqm	National Inst. of Statistics
		Competitors' sales area	Competitors' monthly sales area in stores' trade area	numerical - continuous	sqm	Nielsen

Table 2: Variables description

Sales Revenues

In this work, both store and business unit performance are measured by their respective sales revenues which is the central dependent variable of the econometric models to be fitted. Sales revenues account for the value in Euros that customers pay for the products purchased in stores, deducted by the Value-Added Tax to be delivered to the Portuguese state and by the discounts given to costumers and deposited in their loyalty cards for future purchases. In short, sales revenues account for the money that was earned by the retailer.

Sales Area

Space is the central independent variable of this empirical analysis and is measured by sales area. Store sales area accounts not only for the shelf space where the products are displayed but also includes checking counters, service counters and circulation areas. It includes all the space visible to customers required to provide them a satisfactory shopping experience. Therefore, it does not consider other areas such as storage, loading bays or parking places. Business unit sales area accounts not only for the space where the products of the respective business units are displayed but also for the surrounding circulation area. Store areas as the checking or service counters are allocated to a residual business unit to ensure that the sum of the space allocated to every business unit equals the store total sales area.

Store Format

In section 3.2.1., the three grocery retail formats operated by Sonae MC were introduced. They are named with three distinctive banners, i.e., Continente, Continente Modelo and Continente Bom Dia. The dataset includes 40 Continente, 116 Continente Modelo and 36 Continente Bom Dia stores.

Trade Area

A store’s trade area defines the boundaries of its geographical market where it must strive to attract consumers facing the competition of other grocery retailers’ stores. Trade area measurement is important for retailers’ decision-making, since it is necessary to estimate store sales potential. For this reason, it is an important process in the retail site-location decision and in the stores performance assessment. However, it is proven to be a huge challenge for retailers which have defined trade areas through simple “rules-of-thumb” based on managerial expertise and experience. The appearance of Geographical Information Systems (GIS) is helping retailers to overcome this issue since such systems are able to support spatial interaction models that establish the correlation between store attractiveness and distance to consumers which enables to establish more accurately the trade area limits and, consequently, to calculate several aspects of the trade area such as population or competition that are used in store sales revenues estimation. A geographical information system software (QGIS) has been used to measure stores’ trade area.

Trade area includes the region surrounding a store and is defined by a circle having a radius representing the distance of the residence of the last client accounting for an established percentage of the store sales. This information was provided by the Portuguese retailer.

Adopting this approach, the radius of the circles was estimated defining the trade area of each store. Next, these radiuses were aggregated into four different intervals: 2.5 km, 5 km, 7.5 km and 10 km, as displayed in the table below. This option was taken to standardize the process of collecting information on population and competition. In fact, this process was repeated for every retailer store and it would be unfeasible to personalize it according to the exact radius distance for every store. A maximum of 10 km has been defined for the radius of a store. This decision intends to limit the seasonality effects. There are stores located in tourist areas where most sales are made in the summer by people living in other cities of the country. Therefore, it is not reasonable to consider that most shoppers are part of such stores' trade area which, following this approach, would have an extension of hundreds of kilometers, would result in a huge market potential and in hundreds of competitors. Therefore, by establishing an upper limit of 10 km, it is guaranteed that these stores will not have an unreasonable and biased trade area.

	Layers	**Trade Area**
A	2.5 km	0 km ≤ TA < 3.75 km
B	5 km	3.75 km ≤ TA < 6.25 km
C	7.5 km	6.25 km ≤ TA < 8.75km
D	10 km	TA ≥ 8.75 km

Table 3: Store trade area layers

The analysis of each store's trade area allows to capture the surrounding areas of each city that are more relevant to stores' performance by means of assessing how many residents live nearby and how many competitors are also fighting to attract those consumers.

Population and Population Density in Stores' Trade Area

To determine the population within the trade area the open-source geographic information system software (QGIS) has been used.

For this analysis, it was required to gather and use information about the spatial location of the assessed stores, the major characteristics of the competitors (e.g. spatial location, sales area, brand) and demographic data.

The dataset includes detailed information concerning the number of residents at the parish level, collected from the National Institute of Statistics. The population existing in each parish was distributed per square meter based on the approach applied in the scientific research project PRISE – "Avaliação de perdas e risco sísmico dos edificios em Portugal" (Marques *et al.*., 2014). In this project, the population distribution per square meter was obtained through a combination of two sources:

- Population by parishes, from Census 2011 data, made available by the National Institute of Statistics;

- Detailed information, provided by the initiative Landscan that, through satellite images, predicts urban areas and rural areas, estimating the population density.

For the latter variable, weights were assigned to each square km based on the likeliness to exist residential population. Higher weights were considered in urban areas and zero values were assigned in zones such as rivers or dense forests. To update the information on the residential population in Portugal from 2011 to 2015, the variation of the population in each district between 2011 to 2015 was calculated. This variation was then proportionally distributed per square meter. Through the application of spatial algorithms available in QGIS, the circles around each store for the four layers – 2.5 km, 5 km, 7.5 km and 10 km – were created, defining in this way the corresponding theoretical trade areas.

Competitors' sales Area in Stores' Trade Area

Similarly, a dataset was collected for each month of analysis, of the total number of grocery retailer stores – both internal and external – as well as the respective sales area, measured in square meters. Information from internal competitors was provided by Sonae while information from external competitors was provided by Nielsen, a global information, data and measurement company that collects information from grocery retailers, organizes it and distributes it back to them. However, some retailers do not deliver this information.

In such cases, it is still possible to know exactly how many stores they own and where they are located. However, concerning their sales area it is usually estimated an average value of sales area for those stores.

To determine the average number of competitor stores from the same and different brands and the average sales area of those stores in the trade areas of each store, spatial queries were applied to the spatially distributed stores by using QGIS. Again, the coordinates of all grocery stores in the country were required in the model. This dataset was analyzed for each month of the time interval considered in the study in order to assess the stores inside the influence area of each Sonae MC store. It was then calculated the number of competitor stores and their total sales area for each interval – 2.5 km, 5 km, 7.5 km and 10 km – surrounding a Sonae store.

Purchasing Power Index of the City Population

Concerning the purchasing power index (PPI), an index that compares the residents' economic power among different areas of residence, and that can be considered as a *proxy* of the willingness of the population to spend, it was not possible to collect it for each store's trade area. For this reason, the PPI of the city in which a store is located was used. This information was made available by the National Institute of Statistics. This fact represents a limitation of this study since there are many stores operating in larger cities that will have the same purchasing power index.

3.3. The Sales-Space Relationship of a Portuguese Grocery Retailer

In this section, the methodology outlined in section 3.1. is applied, supported by the datasets described in section 3.2. In short, in the first stage, the simple sales-space relationship is studied at the store level, also considering the differences among store formats. In the second stage, other variables related to store trade area are added and their impact on store sales revenues and on the sales-space relationship is assessed, also considering differences among store formats. Finally, in the last stage the sales-space relationship is studied at the business-unit level.

In every stage of the analysis simple or multiple econometric models are fitted in order to estimate the correlations between store sales revenues and the explanatory variables.

The type of methods used for each step of the analysis depends on the variables to be considered. When considering all the stores independently of the store format, it is used the panel data least square method with cross-sectional and period fixed effects. Fixed effects were considered since each store and period intrinsic characteristics are important for this analysis. Moreover, a secondary goal of the analysis is to assess the performance of each considered store and month from the estimated regression. In fact, since there are factors influencing store sales revenues that are not considered in the analysis, when defining the space allocated to a store, other stores presenting similar attributes can be identified based on these factors left out and the differences of those stores' performance could be assessed from the fitted regression. Based on such differences, retailers could adjust the estimated sales regression acknowledging the existence of store effects. Therefore, fixed-effect analysis provides useful information for retailers that can improve decision-making.

On the other hand, considering the different store formats lead to the introduction of *dummy* variables in the models identifying the banner of each store format. In such cases, an ordinary least square method is used since the panel least square method does not allow the introduction of *dummy* variables. Therefore, only the period effects can be estimated in these models.

3.3.1. Store Level: Simple Regression Models

In the first stage of the empirical analysis, the objective is to measure the impact of space in stores' performance over time, i.e., the percentage change induced in store sales revenues by a 1% change of store sales area. This defines store space-elasticity.

To this purpose, a panel dataset that includes 2304 observations is used, i.e., 192 stores over 12 months, from September 2015 to August 2016. The dataset includes a significant number of heterogeneous stores, since monthly store sales area range from 474 square meters to 15,822 square meters, while the monthly store sales revenues range from 207,988 € to 9,255,335 €. This heterogeneity is important to capture the impact of space on store performance since if sales area was constant across stores, it would be impossible to estimate how space changes affect store sales.

Descriptive Statistics	Store Sales Revenues	Store Sales Area
Average	1,320,232	2,899
Median	908,726	2,050
Maximum	9,255,335	15,822
Minimum	207,988	474
St. Deviation	1,242,486	2,515
Skewness	2	2
Kurtosis	9	8
Sample Size	2,304	2,304

Table 4: Descriptive statistics – model 1

In order to estimate the impact of space on store sales revenues, a simple regression model is applied in which sales revenues are explained by sales area (all the variables are in logs).

Model 1

$$\log \text{sales\_store} = \beta_0 + \beta_1 \log \text{area\_store} + \varepsilon \quad (3.1)$$

sales_store: monthly store sales revenues (euros);

area_store: monthly store sales area (square meters);

β_0: regression intercept;

β_1: regression coefficient;

ε: error term.

The model was fitted by a panel least squares method. The results of this regression found that there is a positive correlation between store sales area and store sales revenues, yielding a space-elasticity of 0.67 which means that on average a 1% increase (decrease) in store sales area results in a 0.67% increase (decrease) in monthly sales revenues.

Model 1 Results		
Parameters	**Estimate**	**P-Value**
C	8.65	0
log (areas_store)	0.67	0.0048

Table 5: Results – model 1

This model shows an excellent fit since the R^2 is 99%, meaning that 99% of the log store sales variability is explained by sales area. Moreover, the regression is highly significant since the p-value of the F-statistic is approximately zero.

Model 1 Statistics	
Statistics	**Coefficient**
R^2	0.988
Adjusted R^2	0.986
S.E. of Regression	0.084
Prob. (F-statistic)	0.000

Table 6: Statistics – model 1

As previously explained, it is relevant to analyze the periods effects (12 months from September 2015 to August 2016) and the store effects (192 stores).

Period effects

Period effects are displayed for the fitted model in the following table. Bearing in mind that the estimated model yielded an average space-elasticity of 0.67, the period effect highlights the expected differences in store sales revenues in each month. Therefore, a store is expected to sell more 22%, 11% and 10% in December, August and July than the average monthly sales revenues estimated in the model.

These results agree with managers' expectations. In fact, owing to Christmas and New Year Eve holidays and to the Christmas allowance paid to employees, December is the period of the year with higher consumption levels. This effect is also expected in July and August due to the positive impact of foreigners' consumption and to the summer allowance paid to employees. On the other hand, February is the month with the most pronounced negative effect since it is the month of the year with fewer days. Moreover, in January and February there is a "hangover" effect after the consumption acceleration occurring in December. This reason helps to explain the negative effects observed for these months.

Period	sep/15	oct/15	nov/15	dec/15	jan/16	feb/16	mar/16	apr/16	may/16	jun/16	jul/16	aug/16
Effect	-3.5%	-4.1%	-3.1%	21.8%	-6.5%	-12.0%	-1.3%	-5.2%	-5.3%	-1.3%	11.0%	9.7%

Table 7: Period effects – model 1

Period effects allow the estimation of monthly sales revenues considering the month-specific mean differences. As an example, suppose that the Portuguese grocery retailer is planning to open a new store with 2,000 square meters of sales area. According to the fitted regression, this store will sell an average value of 912,452 € a month and, consequently, 10,949,420 € a year. It is relevant for the retailer to accurately estimate monthly sales to support managerial decision-making. Therefore, by applying the fitted period effects this store is estimated to sell 1,111,426 € in December and 802,756 € in February. Full results for the remaining months are displayed in the table below.

month	store sales
January	852,793
February	802,756
March	900,551
April	864,803
May	863,750
June	900,995
July	1,012,937
August	1,000,535
September	880,493
October	874,600
November	883,780
December	1,111,426
Total	**10,949,420**

Table 8: Example of period effects practical application – model 1

Store effects

Store effects allow a more accurate estimation of the average monthly sales revenues for an existing store considering the store-specific mean differences. In fact, the estimated regression represents an average value of the 192 stores considered. However, as mentioned before, stores exhibit differences in their internal attributes, in addition to sales area, and they operate under different environments. Since this multiplicity of factors has a significant impact on store sales revenues and they were not considered in the regression model, stores' actual sales revenues do not match their estimated values. The table below displays the store effects that stood out from the 192 stores considered in the analysis (all store effects are displayed in the Appendix 1).

Top 10		Bottom 10	
store	effect	store	effect
L0003	85.5%	L0471	-87.9%
L0463	85.1%	L0317	-84.1%
L0012	76.3%	L0867	-78.7%
L0464	72.2%	L0305	-71.7%
L0001	71.6%	L0315	-69.0%
L0007	70.0%	L0314	-61.7%
L0203	68.0%	L0318	-60.0%
L0002	66.2%	L1704	-59.0%
L0006	65.0%	L0330	-57.5%
L0004	62.1%	L0328	-54.8%

Table 9: Example of store effects practical application – model 1

Store effects must be taken into consideration when assessing the performance of existing stores. Moreover, they might also help in the estimation of new store sales revenues. Considering the above example, a new store with 2,000 sqm is estimated to sell 10,949,420 € annually. However, the retailer managers have concluded that this store will operate under a similar environment of the store L0203, facing a similar number of competitors, being located in a similar shopping mall, etc. Store L0203 was estimated to sell more 68% than the average value yielded by the regression. Therefore, this positive effect can be extrapolated to the new store in order to estimate more precisely its sales revenue. Applying the positive effect of 68%, the store is expected to sell 18,394,106 € in a year.

Period and store effects can be used together to get a more precise estimation of a store's monthly sales. In the same example, given both the store and period effects, the new store can be expected to sell 1,867,103 € in December and 1,348,562 € in February. The full results of this new store expected monthly sales are displayed in the next table.

Store Effect

L0203	68.0%
month	**store sales**
January	1,432,620
February	1,348,562
March	1,512,851
April	1,452,797
May	1,451,028
June	1,513,596
July	1,701,649
August	1,680,815
September	1,479,154
October	1,469,254
November	1,484,676
December	1,867,103
Total	**18,394,106**

Table 10: Example of period and store effects combined practical application – model 1

The second step of this stage aims at measuring the sales-space relationship for the different store formats operated by the Portuguese leading retailer, i.e., Continente, Continente Modelo and Continente Bom Dia. Since these store formats exhibit different characteristics and operational standards, it is relevant to understand their differences.

Bearing this objective in mind, *dummy* variables were added to the prior model in order to identify the store format banner of each store – 40 Continente, 116 Continente Modelo and 36 Continente Bom Dia. The most relevant statistics for each store format are displayed in the table below.

Descriptive Statistics	Continente		Continente Modelo		Continente Bom Dia	
	Store Sales	Store Sales Area	Store Sales	Store Sales Area	Store Sales	Store Sales Area
Average	3,176,247	7,017	944,455	2,069	468,830	996
Median	2,776,823	7,410	886,325	2,049	416,781	934
Maximum	9,255,335	15,822	3,398,974	3,860	1,198,333	1,619
Minimum	841,056	3,069	286,742	1,131	207,988	474
St. Deviation	1,595,194	2,778	330,842	328	205,812	329
Skewness	1	1	1	1	1	0
Kurtosis	0	0	4	8	1	-1
Sample Size	480	480	1,344	1,344	432	432

Table 11: Descriptive statistics – model 2

In this case, an ordinary least square method was applied since the inclusion of *dummy* variables does not allow to capture the store-specific heterogeneity required in the panel least square method. The regression model applied is as follows:

Model 2

$$\text{log sales\_store} = \beta_0 + \beta_1 \text{ continente modelo} + \beta_2 \text{ continente} + \beta_3 \text{ log areas\_store} + \beta_4 \text{ log areas\_store continente modelo} + \beta_5 \text{ log areas\_store continente} + \varepsilon \quad (3.2)$$

sales_store: monthly store sales (euros);

areas_store: monthly store sales area (square meters);

continente modelo: *dummy* variable identifying continente modelo stores

continente: *dummy* variable identifying continente stores;

β_0, β_1, β_2: regression intercept for Continente Bom Dia, Continente Modelo and Continente, respectively;

β_3, β_4, β_5: regression coefficients for Continente Bom Dia, Continente Modelo and Continente Bom Dia respectively;

ε: error term.

The results obtained show a higher space-elasticity for the total number of stores (0.98) when compared to the space-elasticity obtained in the prior model (0.67). This difference is explained by the use of different estimation methods, namely, ordinary least square and panel least square, respectively. The results obtained for this model show that space has a different impact on store format sales revenues. In fact, Continente Modelo stores exhibit the highest space-elasticity, 1.24, which means that, on average, a 1% increase (decrease) in Continente Modelo store sales area leads to a 1.24% increase (decrease) in store sales revenues. On the other hand, Continente Bom Dia stores exhibit the lowest space-elasticity, 0.77, whereas Continente stores were found to have a space-elasticity of 1.02.

Model 2 - Results			
Store Formats	**Parameters**	**Estimate**	**P-Value**
Continente	C	5.93	0.00
	log (areas_store)	1.02	0.00
Continente Modelo	C	4.26	0.00
	log (areas_store)	1.24	0.00
Continente Bom Dia	C	7.72	0.00
	log (areas_store)	0.77	0.00

Table 12: Results – model 2

This model shows an excellent fit since the R^2 is 85%, meaning that 85% of the log store sales variability is explained by sales area. Moreover, the regression is highly significant since the p-value of the F-statistic is approximately zero.

Model 2 - Statistics	
Statistics	Coefficient
R^2	0.850
Adjusted R^2	0.849
S.E. of Regression	0.280
Prob. (F-statistic)	0.000

Table 13: Statistics – model 2

3.3.2. Store Level: Multiple Regression Models

In the second stage of the methodology other explanatory variables of sales variability besides sales area are introduced. By adding these new independent variables to the store sales area, a multiple regression model to estimate store sales revenues is developed. These variables, whose choice was justified earlier in the beginning of the chapter, attempt to represent three main factors:

1. **Population and Trade Area**
   a) Number of residents in the stores' trade area
   b) Number of residents per square meter in the stores' trade area

2. **Economic characteristics of the population**
   a) Purchasing power index of the city in which the store is located.

3. **Competition in the trade area**
   a) Competitors' sales area (square meters) in the store's trade area

The descriptive statistics of the variables introduced in the multiple regression model are displayed in the table below. It is important to highlight the heterogeneity of the dataset. For instance, there is a store located in a trade area whose competitors own 343,295 square meters, while there is another store in which the competition space is only 800 square meters. This heterogeneity is important to measure the impact of these factors on stores' performance.

Descriptive Statistics	Store Sales	Store Sales Area	Population	Pop. Density	Competitors' Sales Area	PPI
Average	1,320,232	2,899	100,128	2,031	22,332	104
Median	908,726	2,050	61,504	712	13,125	99
Maximum	9,255,335	15,822	1,604,922	10,883	343,295	217
Minimum	207,988	474	9,564	15	800	53
St. Deviation	1,242,486	2,515	126,215	2,402	31,015	34
Skewness	2	2	4	1	4	2
Kurtosis	9	8	22	4	20	7
Sample Size	2,304	2,304	2,304	2,304	2,304	2,304

Table 14: Descriptive statistics – model 3

In this stage, the panel least squares method is used, considering 192 stores over 12 months of analysis, from September 2015 to August 2016 and fixed effects are considered to represent month and store-specific means, measuring the differences between the actual and the fitted values of monthly sales revenues for these specific stores and periods.

This method does not allow the introduction of the purchasing power index in the regression model since this variable is related to the city where a store is located. This consideration makes this variable equal to many stores located in the same city. Hence, the variable is not a unique store attribute which is a requirement for using the panel least square method.

Before applying the regression, it is important to assess the relevance of each variable to explain sales variability. Therefore, simple regressions models were fitted for each variable – population, population density and competition sales area – to measure their effect on store sales revenues. The results are displayed in Table 15 and show that every variable has a significant impact on store sales revenues.

Population was estimated to have a positive impact of 0.14, i.e. a 1% increase (decrease) in a store trade area population is expected to increase (decrease) its sales revenues by 0.14%. This result is expected since the larger the population, the larger the number of potential clients.

On the other hand, population density was estimated to have a negative impact on store sales revenues of 0.10, i.e., a 1% increase (decrease) in the trade area population density is estimated to decrease (increase) store sales in 0.10%. This result was unexpected since a higher population density means more potential clients per square meter in a store trade area. However, this result can be explained by two factors: the Portuguese clients' buying habits for grocery products and the trade area measurement process aforementioned. Stores with larger sales revenues are the bigger hypermarkets located in the suburbs of major cities. These stores are highly attractive for consumers, either for their size and the assortment width offered to consumers or because they are usually located near large commercial areas. For this reason, the Portuguese consumer still does most of grocery shopping in these larger stores, travelling by car. On the other hand, these consumers go to convenience stores located near their residential areas to purchase some missing grocery items. Since the attractiveness of larger hypermarkets is stronger their trade area is also larger, while convenience stores' trade area is lower. However, these convenience stores with lower sales revenues are usually located in highly populated areas. Therefore, they exhibit higher values of population density when compared to the larger hypermarkets. These reasons help to interpret the negative relationship estimated in the regression of sales revenues on population density. Next, it will be interesting to analyze the impact of this variable on each store format.

Finally, competitors' sales area was found to have a positive impact on store sales revenues of 0.09, i.e., a 1% increase (decrease) in competitors' sales area is expected to increase (decrease) store sales revenues by 0.09%. These results agree some contradictory conclusions obtained in other studies discussed in Chapter 2. Although competition is expected to decrease stores' performance, the best performing stores are usually located in areas with strong market potential, where many competitors are also settled. Therefore, it is important to assess the impact of competition jointly with population factors.

Simple Regression Model on Population		
Parameters	**Estimate**	**P-Value**
C	12.29	0.00
log (population)	0.14	0.00

Simple Regression Model on Population Density		
Parameters	**Estimate**	**P-Value**
C	14.49	0.00
log (population density)	-0.10	0.00

Simple Regression Model on Competitors Sales Area		
Parameters	**Estimate**	**P-Value**
C	12.95	0.00
log (comp_sales area)	0.09	0.00

Table 15: Results of the simple regression models with the variables in Model 3

All these simple regression models show an excellent fit since the R^2 is 99%, meaning that 99% of the log store sales variability is explained by these variables. Moreover, all regressions are highly significant since the p-value of the F-statistic is approximately zero.

Statistics - Simple Regression Models			
Variables	**Population**	**Population Density**	**Compet. Sales Area**
Statistics	**Coefficient**	**Coefficient**	**Coefficient**
R^2	0.988	0.988	0.988
Adjusted R^2	0.987	0.986	0.987
S.E. of Regression	0.083	0.084	0.083
Prob. (F-statistic)	0.000	0.000	0.000

Table 16: Statistics of the simple regression models with the variables in model 3

After concluding that the estimated parameters of all the variables are significant, a multiple regression model is applied considering these four independent variables together, i.e., store sales area, population, population density and competitors' sales area to explain store sales revenues. The model is presented below:

Model 3

$$\log \text{sales\_store} = \beta_0 + \beta_1 \log \text{areas\_store} + \beta_2 \log \text{population} + \beta_3 \log \text{population density} + \beta_4 \log \text{competitors' sales area} + \varepsilon \quad (3.3)$$

sales_store: monthly store sales revenues (euros);

areas_store: monthly store sales area (square meters);

population: number of residents living in a store trade area;

population density: number of residents per square meter living in a store trade area;

competitors' sales area: sales area of competitor grocery retailer stores in store trade area (square meters);

β_0: regression intercept;

β_1: regression coefficient of sales area;

β_2: regression coefficient of population;

β_3: regression coefficient of population density;

$\beta 4$: regression coefficient of competitors' sales area;

ε: error term.

The results of the regression are displayed below. All estimated parameters were found to be significant since their p-value is less than 5%. A space-elasticity of 0.66 was found, meaning that a 1% increase (decrease) of store sales area is expected to increase (decrease) store sales revenues by 0.66%. This result is almost coincident with that obtained in model 1 in the first stage of the analysis (0.67).

Concerning population density, a negative impact of 0.07 was estimated, meaning that a 1% increase (decrease) of the population density of a store's trade area is expected to decrease (increase) sales revenues by 0.07%. This result is similar to that obtained in the simple regression model (0.10).

The major differences that need to be analyzed were found for population and competition variables. The estimated impact of population was found to be much stronger in this model (0.21) than in the simple regression in which it was considered as the only independent variable (0.14). On the other hand, the competitors' sales area was estimated to negatively influence store sales (-0.08), contradicting the results found in the simple regression model, in which it was found to have a positive impact of 0.09.

These results highlight the importance of considering both variables. Previously, it was explained that the positive impact of competition sales area on store sales revenues obtained in the simple regression model between these two variables, although seemingly contradictory, was explained by the fact that most best performing stores are located in highly populated areas facing more competitors. By considering population and competition sales area in the same regression model, we conclude that the population positive impact is amplified while competition turns out to have the expected negative impact on store sales revenues.

Model 3 - Results		
Parameters	**Estimate**	**P-Value**
C	7.58	0.00
log (area_store)	0.66	0.00
log (population)	0.21	0.00
log (population density)	-0.07	0.02
log (comp_sales area)	-0.08	0.05

Table 17: Results – model 3

This model shows an excellent fit since the R^2 is 99%, meaning that 99% of the log store sales variability is explained by these variables. Moreover, these regressions are highly significant since the p-value of the F-statistic is approximately zero.

Model 3 - Statistics	
Statistics	Coefficient
R^2	0.988
Adjusted R^2	0.987
S.E. of Regression	0.083
Prob. (F-statistic)	0.000

Table 18: Statistics – model 3

The table below summarizes the space-elasticities estimated using the panel least square method, based on the simple regression model fitted in the first stage of the empirical analysis and on the multiple regression model fitted in this stage. It can be observed that the store space-elasticity (0.66) is almost coincident with that estimated in model 1 (0.67). These results also agree with those obtained in the literature at the store level. Davies (1977) found an average space-elasticity of 0.748, Thurik (1988) estimated a space-elasticity of 0.51 for hypermarkets and 0.68 for supermarkets. Finally, Castro (2011) found an average space-elasticity of 1.21, using an ordinary least square method. In section 3.2., a simple regression model is also estimated to measure space-elasticities among store formats and the value obtained for all stores was 0.98 which is more similar to that obtained in Castro (2011). Therefore, this work supports the results of the literature that have proven the positive impact of store space on store sales revenues.

On the other hand, it may be concluded that the variables characterizing stores' trade area have no impact on the grocery retailer sales-space relationship. Furthermore, according to the fitted model, a 1% increase (decrease) of the store sales area would induce an average increase (decrease) between 0.66% and 0.67% of the store sales revenues. This analysis must be combined with the costs derived from modifying store sales area to assess if that change is profitable for the retailer. Moreover, this analysis must be conducted for each store, considering store effects since the output of the model is an average value for monthly sales revenues.

Panel Least Square Method	Space-Elasticity
Simple Regression Model (model 1)	0.67
Multiple Regression Model (model 3)	0.66

Table 19: Summary of the estimated space-elasticities using the panel least square method – models 1 and 3

Although the estimated values for the space-elasticity in both models are similar, substantial differences are observed concerning store effects. In fact, there are some stores which in model 3 were found to have a negative store effect while in model 1 were found to have a positive effect. The opposite also occurs. As an example, the actual value of sales revenues of store L0389 is -18.6% lower than the fitted sales revenues obtained in model 1. Therefore, if store sales revenues were only influenced by store sales area, store L0389's sales revenues should be 18.6% larger than its actual value. In model 3, three new explanatory variables were added, i.e., population, population density and competition. For this model, store L0389's actual sales revenue was found to be 0.2% larger than the fitted value. In this case, the joint effect of population, population density and competition made the fitted value in the regression move towards the store's sales revenues actual value. On the other hand, the joint impact of population, population density and competition on the sales revenues fitted value was positive for some stores relatively to the same value estimated in model 1.

The table below displays the top 10 and bottom 10 store effect differences. The differences for every store included in the dataset are shown in the Appendix 1.

Top 10 Store Effect Differences			
Store	Model 1	Model 3	Difference
L0261	-9.3%	15.0%	24.3%
L1169	-51.7%	-27.8%	23.9%
L0273	3.7%	26.8%	23.1%
L1902	-46.0%	-25.7%	20.3%
L0389	-18.6%	0.2%	18.8%
L0867	-78.7%	-62.0%	16.8%
L0263	3.9%	19.7%	15.8%
L0847	-5.9%	9.5%	15.4%
L0315	-69.0%	-53.7%	15.3%
L2089	40.0%	54.6%	14.7%

Bottom 10 Store Effect Differences			
Store	Model 1	Model 3	Difference
L0005	48.1%	34.7%	-13.4%
L0275	-0.1%	-13.5%	-13.4%
L0012	76.3%	62.3%	-14.0%
L0231	-15.0%	-29.5%	-14.5%
L0004	62.1%	47.3%	-14.8%
L0307	-5.7%	-21.8%	-16.1%
L0242	-6.6%	-26.3%	-19.7%
L0248	-32.3%	-52.6%	-20.3%
L0303	1.9%	-18.6%	-20.5%
L0011	7.5%	-14.7%	-22.2%

Table 20: Store effects comparison – models 1 and 3

Concerning period effects, small differences were also observed between both models and every month has the same sign. The magnitude of these differences is not enough to change the previous conclusion that December, July and August are the periods with greater period effects in every model, while February and January are the periods with lower effect.

Period	sep/15	oct/15	nov/15	dec/15	jan/16	feb/16	mar/16	apr/16	may/16	jun/16	jul/16	aug/16
Effect (model 1)	-3.5%	-4.1%	-3.1%	21.8%	-6.5%	-12.0%	-1.3%	-5.2%	-5.3%	-1.3%	11.0%	9.7%
Effect (model 3)	-3.3%	-3.9%	-2.9%	22.1%	-6.2%	-11.7%	-1.0%	-4.9%	-5.0%	-0.9%	7.9%	9.9%

Table 21: Period effects comparison – models 1 and 3

In section 3.3.1., an example was given where the estimated simple regression model (model 1) was applied to a new store opening with 2,000 sqm. Now, the same example is considered and the new independent variables of the multiple regression model are introduced. The store is expected to have an attractiveness similar to some stores for which was determined a radius of influence of 10 km in the definition of their trade area. Assume that, in this store trade area, there are 50,000 residents, the population density is 190 and competition sales area is 15,000 sqm. When the simple regression model was applied in section 3.3.1., the store was estimated to sell 10,949,420 € annually. Now, applying the multiple regression model developed above, the estimated annual sales revenues are 11,307,220 €.

Considering the same hypothesis set as in section 3.3.1, i.e., that this store has similar attributes to the store L0203 that were not taken into account in this regression and that explain the differences in store effects as store accessibility, location near a commercial site, etc., the store effect estimated for store L0203 is applied to the expected value of sales revenues of the new store. In model 3, store L0203's store effect is 55.9% which is lower than that obtained in the simple regression model (68%). Therefore, this store is expected to sell 17,742,452 € annually. This value is lower than that calculated in model 1 (18,394,106 €) since the combined impact of population, population density and competition on the regression fitted value was negative when compared to that obtained considering only sales area as an independent variable.

model 1

Store Effect	
L0203	**68.0%**

month	store sales
January	1,432,620
February	1,348,562
March	1,512,851
April	1,452,797
May	1,451,028
June	1,513,596
July	1,701,649
August	1,680,815
September	1,479,154
October	1,469,254
November	1,484,676
December	1,867,103
Total	**18,394,106**

model 3

Store Effect	
L0203	**55.9%**

month	store sales
January	1,386,140
February	1,305,159
March	1,463,707
April	1,405,816
May	1,404,306
June	1,464,781
July	1,595,886
August	1,625,067
September	1,430,096
October	1,420,468
November	1,435,796
December	1,805,230
Total	**17,742,452**

Table 22: Example of period and store effects combined practical application – models 1 and 3

This example aims at highlighting the relevance of estimating cross-sectional and period fixed effects since they allow to capture the store and period differences relatively to the average values estimated in the model, enabling retailers to perform a better store assessment and consequently improving their decision-making process.

Moreover, the application of these regression models enables retailers to estimate the expected monthly sales revenues for new stores based on their sales areas and trade area characteristics, i,e., population, population density and competition sales area. Therefore, they can support their retail site location decisions on these models by estimating the expected sales revenues for the available options. On the other hand, since sales area is defined by retailers, they can estimate the expected sales revenues for the possible store's space and define the space best fitted to the store's sales potential. Finally, the model also allows to assess existing stores' performance and to predict their sales revenues in the future after changes in the characteristics of the trade area. Hence, these models can support retailers in adjusting stores' space in future revamps. Such analyses must be complemented with the costs that these decisions entail to assess their profitability for the retailer.

Similarly to what was performed in the first stage of this work, it is relevant to measure the differences among the estimated space-elasticities of the store formats operated by Sonae MC found with the application of a multiple regression model considering new independent variables that characterize store trade areas, i.e., population, population density and competition.

In this step of the analysis, a new variable is introduced, the purchasing power index of the city in which a store is located. This variable was left out from the previous model, owing to its incompatibility with the panel square least method. However, since *dummy* variables were introduced in order to identify the store format banner, the ordinary least square method is now applied because it does not measure the store fixed effects, allowing the introduction of the purchasing power index.

The descriptive statistics of each store format are displayed in the Appendix 2.

First, it is important to develop simple regression models for each variable to assess their importance at explaining store sales revenues for each store format. The statistics for each of the fitted regressions are displayed in the Appendix 3. All models show a good fit since the R^2 ranged from 75% to 85%. Moreover, all the regressions are highly significant since the p-value of their F-statistics is approximately zero.

Concerning store sales area, the table below displays the results obtained with the same simple regression model as in section 3.3.2. (model 2).

Simple Regression Model on Store Sales Area			
Store Formats	**Parameters**	**Estimate**	**P-Value**
Continente	C	5.93	0.00
	log (areas_store)	1.02	0.00
Continente Modelo	C	4.26	0.00
	log (areas_store)	1.24	0.00
Continente Bom Dia	C	7.72	0.00
	log (areas_store)	0.77	0.00

Table 23: Results of the simple regression model on store sales area, considering store formats

Concerning population, a positive impact is estimated for Continente (0.37) and for Continente Modelo (0.11) stores' sales revenues. Regarding Continente Bom Dia stores, the estimate (0.04) has a p-value of 9%, nonsignificant at a 5% significance level, but already significant at 10%. Therefore, we decided to consider it still significant and conclude that the impact for these stores is positive, although weak.

Simple Regression Model on Population			
Store Formats	**Parameters**	**Estimate**	**P-Value**
Continente	C	10.49	0.00
	log (population)	0.37	0.00
Continente Modelo	C	12.54	0.96
	log (population)	0.11	0.01
Continente Bom Dia	C	12.52	0.00
	log (population)	0.04	0.09

Table 24: Results of the simple regression model on population, considering store formats

Population density is estimated to have a positive impact on Continente, a weak positive impact on Continente Modelo and a negative impact on Continente Bom Dia store sales revenues.

Simple Regression Model on Population Density			
Store Formats	**Parameters**	**Estimate**	**P-Value**
Continente	C	13.47	0.85
	log (pop. density)	0.20	0.00
Continente Modelo	C	13.59	0.62
	log (pop. density)	0.02	0.00
Continente Bom Dia	C	13.51	0.00
	log (pop. density)	-0.06	0.00

Table 25: Results of the simple regression model on population density, considering store formats

Concerning the purchasing power index, it has a strong impact on Continente (1.11) and on Continente Modelo stores' sales revenues (0.59). The estimated parameter of this variable was found to be nonsignificant for Continente Bom Dia stores, which is understandable, since most Continente Bom Dia stores are located in the two largest cities in the country. Since the purchasing power index is related to the city in which a store is located, most Continente Bom Dia stores have a similar purchasing power index. Hence, the variability in the store sales revenues cannot be explained by the variability in the purchasing power index for this store format.

Simple Regression Model on Purchasing Power Index			
Store Formats	**Parameters**	**Estimate**	**P-Value**
Continente	C	9.58	0.00
	log (ppi)	1.11	0.00
Continente Modelo	C	11.08	0.00
	log (ppi)	0.59	0.00
Continente Bom Dia	C	13.15	0.00
	log (ppi)	´-0.04	0.49

Table 26: Results of the simple regression model on purchasing power index, considering store formats

Finally, competition sales area is estimated to have a positive impact on both Continente and Continente Modelo store sales revenues. As mentioned before, this result, although contrary to what could be expected, is explained by the fact that the best performing stores are located in highly populated areas where competition is stronger. Therefore, competition should be analyzed together with population. Nevertheless, competition was found to be nonsignificant for Continente Bom Dia stores.

Simple Regression Model on Competitors' Sales Area			
Store Formats	**Parameters**	**Estimate**	**P-Value**
Continente	C	11.74	0.03
	log (comp. sales area)	0.30	0.00
Continente Modelo	C	13.09	0.56
	log (comp. sales area)	0.07	0.00
Continente Bom Dia	C	13.23	0.00
	log (comp. sales area)	-0.03	0.20

Table 27: Results of the simple regression model on competitors' sales area, considering store formats

After having concluded that all these variables have an impact on store performance, a multiple regression model is applied using the ordinary least square method, with *dummy* variables identifying store format banners. The model is displayed below:

<u>Model 4:</u>

log sales_store $= \beta_0 + \beta_1$ continente modelo $+ \beta_2$ continente $+ \beta_3$ log areas_store $+ \beta_4$ log areas_store continente modelo $+ \beta_5$ log areas_store continente $+ \beta_6$ log population $+ \beta_7$ log population continente modelo $+ \beta_8$ log population continente $+ \beta_9$ log population density $+ \beta_{10}$ log population density continente modelo $+ \beta_{11}$ log population density continente $+ \beta_{12}$ log ppi $+ \beta_{13}$ log ppi continente modelo $+ \beta_{14}$ log ppi continente $+ \beta_{15}$ log number competitors $+ \beta_{16}$ log number competitors continente modelo $+ \beta_{17}$ log number competitors continente $+ \varepsilon$ (3.4)

The results of this regression are displayed in table 28 below.

Concerning the store sales area, the results are similar to those obtained in model 2. Space-elasticities of 0.95, 1.16 and 0.81 were estimated from the multiple regression model (model 4) for Continente, Continente Modelo and Continente Bom Dia stores respectively, whereas in model 2 the estimated space-elasticity was 1.02, 1.24 and 0.77 respectively.

Concerning population, a positive impact was estimated on sales revenues for every store format. Continente stores exhibit a lower value than in the simple regression (0.17 against 0.37), while Continente Modelo stores and Continente Bom Dia stores exhibit a higher impact (0.20 in model 4 against 0.11 in the simple regression and 0.46 in model 4 against 0.04 in the simple regression, respectively).

Concerning the population density of a store trade area, it was found to be nonsignificant for all store formats.

Concerning the purchasing power index of the population living in the city where a store is located, it was found to be nonsignificant for Continente Bom Dia stores, similarly to the estimated result in the simple regression model. On the other hand, a smaller positive impact was estimated for Continente (0.48 in model 4 against 1.11 in the simple regression) and for Continente Modelo stores (0.38 in model 4 against 0.59 in the simple regression).

Concerning competitors' sales area in store trade area, a negative impact is estimated for every store format, i.e., -0.14 for Continente, -0.12 for Continente Modelo and -0.31 for Continente Bom Dia. These results are distinct from those obtained in the simple regression, where competitors' sales area had a positive impact on both Continente (0.30) and Continente Modelo stores (0.07), while in Continente Bom Dia stores the impact was not significant.

Model 4 - Results			
Store Formats	**Parameters**	**Estimate**	**P-Value**
Continente	C	4.11	0.05
	log (areas_store)	0.95	0.02
	log (population)	0.17	0.00
	log (pop. density)	-0.05	0.31
	log (ppi)	0.48	0.00
	log (comp. sales area)	-0.14	0.00
Continente Modelo	C	2.23	0.00
	log (areas_store)	1.16	0.00
	log (population)	0.20	0.00
	log (pop. density)	-0.02	0.55
	log (ppi)	0.38	0.00
	log (comp. sales area)	-0.12	0.00
Continente Bom Dia	C	5.17	0.00
	log (areas_store)	0.81	0.00
	log (population)	0.46	0.00
	log (pop. density)	-0.03	0.15
	log (ppi)	0.05	0.31
	log (comp. sales area)	-0.31	0.00

Table 28: Results – model 4

This model shows a good fit, since the R^2 is 87%, meaning that 87% and of the log stores variability of sales revenues is explained by sales area, population, population density, purchasing power index and competition sales area.

Model 4 - Statistics	
Statistics	Coefficient
R^2	0.872
Adjusted R^2	0.259
S.E. of Regression	0.259
Prob. (F-statistic)	0.000

Table 29: Statistics – model 4

The ordinary least square method allows the estimation of the period effects. Table 30 shows that the results obtained are very close to those obtained applying model 2, i.e., December stands out as the best performing month, whereas January and February exhibit the lowest performances. These results are also close to those estimated using the panel least square method, displayed in table 21.

Period	sep/15	oct/15	nov/15	dec/15	jan/16	feb/16	mar/16	apr/16	may/16	jun/16	jul/16	aug/16
Effect (model 2)	-3.5%	-4.1%	-3.1%	21.8%	-6.5%	-12.0%	-1.3%	-5.2%	-5.3%	-1.3%	11.0%	9.7%
Effect (model 4)	-3.5%	-4.1%	-3.1%	22.0%	-6.4%	-11.8%	-1.1%	-5.0%	-5.1%	-1.0%	9.3%	9.8%

Table 30: Period effects – model 2 and 4

The aforementioned example is now considered in order to show an application of models 2 and 4, that had considered differences among store formats. The store to be opened will have 2,000 square meters and will operate under the brand Continente Modelo. Moreover, its trade area has 50,000 residents, a population density of 190 and 50,000 sqm of competition sales area. It is also considered that the purchasing power index of the city where it is located is 100. Since Continente Modelo stores exhibit similar internal attributes related to their location, width and depth of assortment or costumer service provided, it is pertinent to acknowledge these differences and to consider a regression based exclusively on the stores operating under this banner.

When the simple regression model was applied in section 3.3.1., the store was estimated to sell 10,949,420 € annually, while in model 3 an annual sales revenue of 11,307,220 € was estimated. These models used the panel least square method and considered all the stores independently of their store format.

On the other hand, considering the simple regression model for Continente Modelo stores model 2), the same store is estimated to sell 10,409,511 € annually while, when the multiple regression model for Continente Modelo stores considering population, population density, competition sales area and purchasing power index as explanatory variables is applied (model 4), the store is estimated to sell 10,595,270 € annually. Table 31 below displays the estimated monthly sales revenues for this store.

model 2	
month	**store sales**
January	810,816
February	763,247
March	856,038
April	822,061
May	821,055
June	856,350
July	962,772
August	951,519
September	837,165
October	831,545
November	840,260
December	1,056,684
Total	**10,409,511**

model 4	
month	**store sales**
January	826,825
February	778,505
March	873,112
April	838,603
May	837,761
June	873,835
July	965,291
August	969,646
September	852,309
October	846,482
November	855,819
December	1,077,083
Total	**10,595,270**

Table 31: Example of period effects practical application – models 2 and 4

However, this model has the limitation of not measuring the store effects that are important to capture the impact of other variables that are not considered in the model and that explain the differences between the stores' monthly sales revenues actual values and the average monthly sales revenues fitted from the regression.

It is important to highlight some issues arising from this analysis. The table below displays the results of the space-elasticity of each store format obtained using the ordinary least square method. The results are similar to those obtained in model 2 which denotes a remarkable degree of stability since the differences between the estimated space-elasticities are small.

The results obtained agree with those found in the literature proving the positive impact of space on store sales revenues and showing differences of space-elasticity among different store formats. It has been found in the literature that the supermarket format exhibits a higher space-elasticity than the hypermarket format, while convenience stores have the lowest impact. Thurik (1988) estimated a space-elasticity of 0.51 for hypermarkets and 0.68 for supermarkets, while Castro (2011) found a space-elasticity of 0.76, 1.16 and 0.53 for hypermarkets, supermarkets and convenience stores, respectively. Similar results were also found in this study.

Moreover, in the multiple regression model space was found to be the variable with the strongest impact on store sales revenues for every store format. In fact, from all the variables considered in the model, space exhibited the largest elasticity, i.e., a percentage variation in space has the highest impact on the percentage variation of store sales revenues.

Space-Elasticity by store format			
Ordinary Least Square Method	Continente	Continente Modelo	Continente Bom Dia
Simple Regression Model (model 2)	1.02	1.24	0.77
Multiple Regression Models (model 4)	0.95	1.16	0.81

Table 32: Summary of the estimated space-elasticities for each store format using the ordinary least square method – models 2 and 4

3.3.3. Business Unit Level

In the final stage of this work the objective is to measure space-elasticity at the business unit level, also analyzing the impact of other explanatory variables in this sales-space relationship. The Portuguese retailer business units are displayed in section 3.2., in table 1.

In the first two stages of this empirical analysis, the focus was to analyze the impact of space on store sales revenues performance. The objective was to provide useful information to retailers to help them in the retail site-location decision concerning space definition of new stores and in the assessment of the fit between stores performance and their sales area. However, store space cannot be adjusted frequently and once it is defined retailers must allocate it to business units, categories of products, products and brands. This allocation process is more flexible to adjust despite having to face layout and operational constraints. Therefore, there is an opportunity for retailers to improve their store performances by redefining the space allocated to business units.

In this decision-making process, the relevant business unit attributes should be taken into account, such as the price level, promotion policies, assortment width or the service level provided, which are expected to influence business units' sales revenues. However, since it was not possible to collect such information on business units for our study, the same store external attributes used in the analysis performed at the store level are considered, i.e.

population, population density, purchasing power index and competition. Although these factors are related to stores trade area characteristics they might influence business units' performance. As an example, if the width of the assortment or the quality of the products of a business unit of a strong competitor is better than that offered by a given grocery retailer, competition is expected to influence this specific business unit performance negatively.

For this empirical analysis at the business unit level is used a panel dataset spanning information on 23 business units of 192 stores for 12 months, from September 2015 to August 2016. This dataset was processed in order not to consider business units' monthly store revenues or sales area equal to 0, which may happen, either because not every store displays all business units in their assortment or since it may not have sold a single product of the business unit in a given month. After completing this process, 45,037 observations were obtained. The most relevant statistics of the dataset are displayed in the table below.

	Business Units Analysis Dataset					
Descriptive Statistics	Sales Revenues	Sales Area	Population	Pop. Density	Competition Sales Area	PPI
Average	67,202	139	103,721	2,041	23,192	104
Median	42,719	105	62,158	712	13,308	99
Maximum	999,865	1,749	1,604,922	10,883	343,295	217
Minimum	0	0	9,564	15	800	53
St. Deviation	84,037	139	130,374	2,385	32,006	34
Skewness	3	3	4	1	4	2
Kurtosis	16	13	21	4	19	7
Sample Size	45,037	45,037	45,037	45,037	45,037	45,037

Table 33: Descriptive statistics – business unit analysis

The first step of this analysis is it to fit a simple regression model by the least square method to estimate the impact of space on business units' performance according to the following equation:

<u>Model 5</u>

$$\log \text{sales\_bu} = \beta_0 + \beta_1 \log \text{areas\_bu} + \varepsilon \qquad (3.5)$$

sales_bu: business unit monthly sales (euros);

areas_bu: business unit monthly sales area (square meters);

β_0: regression intercept;

β_1: regression coefficient;

ε: error term.

The estimated business unit space-elasticity is 1.08 which means that, on average, a 1% increase (decrease) of business unit sales area induces a 1.08% increase (decrease) in business unit monthly sales revenues.

Model 5 Results		
Parameters	**Estimate**	**P-Value**
C	5.60	0.00
log (areas_bu)	1.08	0.00

Table 34: Results – model 5

This model shows a good fit since the R^2 is 80%, meaning that 80% of the log business unit sales variability is explained by the log sales area. Moreover, the regression is highly significant since the p-value of the F-statistic is approximately zero.

Model 5 Statistics	
Statistics	**Coefficient**
R^2	0.796
Adjusted R^2	0.796
S.E. of Regression	0.757
Prob. (F-statistic)	0.000

Table 35: Statistics – model 5

In the next step, *dummy* variables identifying business units were introduced in the model and a new model (model 6) measuring space-elasticities for each business unit was applied. The model is displayed below and is similar to that introduced above (model 5).

Model 6:

$$\log \text{sales\_bu} = \beta_0 + \sum_{i=1}^{22} \beta_{0i} \text{bu}_i + \sum_{i=1}^{22} \beta_{1i} \log \text{areas\_bu}_i + \varepsilon \qquad (3.6)$$

sales_bu: business unit monthly sales (euros);

areas_bu$_i$: business unit monthly sales area for each business unit (square meters);

$\boldsymbol{\beta_0}$: regression intercept;

$\boldsymbol{\beta_{0i}}$: regression intercept for each business unit;

$\boldsymbol{\beta_{1i}}$: regression coefficient for each business unit;

i: business unit;

ε: error term.

The full results of this model are displayed in table 36 below. Business units' space was found to have a positive impact on sales revenues in 17 out of 23 business units and space-elasticities ranged from 0.39 to 1.17. In the remaining business units, space was found to be nonsignificant. This could mean that in these business units, sales revenues are not driven by space and might depend on other attributes not considered in this model. In BU 15 – Fruits and Vegetables, this result seems understandable, since other attributes as the product appearance and quality or the atmosphere of the place where they are displayed are more valued by costumers when they shop fruits and vegetables. On the other hand, BU 02 – Sweet Savory is a business unit exhibiting a great seasonality effect, since most sales happen in festive events (Christmas and Easter). BU 08 – Dairy is a business unit in which retailers' promotional policies have a large influence on consumers' purchasing decision.

Model 6 - Results			
BU	**Parameters**	**Coef.**	**P-Value**
BU01	C	6.16	0.00
	log (areas_bu)	1.03	0.00
BU02	C	6.19	0.81
	log (areas_bu)	0.99	0.17
BU03	C	5.66	0.00
	log (areas_bu)	1.08	0.05
BU05	C	5.88	0.03
	log (areas_bu)	1.08	0.03
BU06	C	5.25	0.00
	log (areas_bu)	1.17	0.00
BU07	C	5.60	0.00
	log (areas_bu)	0.98	0.12
BU08	C	6.67	0.00
	log (areas_bu)	0.99	0.19
BU11	C	7.42	0.00
	log (areas_bu)	0.86	0.00
BU12	C	5.68	0.00
	log (areas_bu)	1.12	0.00
BU13	C	7.21	0.00
	log (areas_bu)	0.86	0.00
BU15	C	5.97	0.19
	log (areas_bu)	1.04	0.64
BU16	C	6.19	0.83
	log (areas_bu)	0.99	0.19

Model 6 - Results			
BU	**Parameters**	**Coef.**	**P-Value**
BU18	C	7.11	0.00
	log (areas_bu)	0.73	0.00
BU19	C	7.94	0.00
	log (areas_bu)	0.39	0.00
BU30	C	5.67	0.00
	log (areas_bu)	0.98	0.02
BU31	C	6.07	0.38
	log (areas_bu)	0.80	0.00
BU33	C	7.60	0.00
	log (areas_bu)	0.59	0.00
BU34	C	6.11	0.62
	log (areas_bu)	0.81	0.00
BU35	C	6.26	0.32
	log (areas_bu)	0.88	0.00
BU41	C	4.99	0.00
	log (areas_bu)	0.92	0.00
BU42	C	4.91	0.00
	log (areas_bu)	0.99	0.24
BU43	C	4.31	0.00
	log (areas_bu)	1.13	0.00
BU44	C	5.74	0.00
	log (areas_bu)	0.87	0.00

Table 36: Results – model 6

The table below shows the five business units with the highest and lowest space-elasticities. BU 06 – Home Cleaning exhibits the highest space-elasticity (1.17), while BU 19 – Cafeteria has the lowest (0.39), meaning that a 1% increase (decrease) in sales area is expected to increase (decrease) these business units' sales revenues by 1.17% and 0.39% respectively.

Model 6 - Top 5				
BU	**Description**	**Parameters**	**Coef.**	**P-Value**
BU06	Home Cleaning	C log (areas_bu)	5.25 1.17	0.00 0.00
BU43	Women Apparel	C log (areas_bu)	4.31 1.13	0.00 0.00
BU12	Fishery	C log (areas_bu)	5.68 1.12	0.00 0.00
BU03	Drinks	C log (areas_bu)	5.66 1.08	0.00 0.05
BU05	Hygiene and Beauty	C log (areas_bu)	5.88 1.08	0.03 0.03

Table 37: Top 5 business unit space-elasticities – model 6

Model 6 -Bottom 5				
BU	**Description**	**Parameters**	**Est.**	**P-value**
BU19	Cafeteria	C log (areas_bu)	7.94 0.39	0.00 0.00
BU33	Culture	C log (areas_bu)	7.60 0.59	0.00 0.00
BU18	Take Away	C log (areas_bu)	7.11 0.73	0.00 0.00
BU31	Home	C log (areas_bu)	6.07 0.80	0.38 0.00
BU34	Brico & Auto	C log (areas_bu)	6.11 0.81	0.62 0.00

Table 38: Bottom 5 business unit space-elasticities – model 6

This model shows a good fit since the R^2 is 90%, meaning that 90% of the log business unit sales variability is explained by the log sales area. Moreover, the regression is highly significant since the p-value of the F-statistic is approximately zero.

Model 6 Statistics	
Statistics	**Coefficient**
R^2	0.900
Adjusted R^2	0.900
S.E. of Regression	0.531
Prob. (F-statistic)	0.000

Table 39: Statistics – model 6

Retailers must take tactic decisions on the type of assortment to offer within stores and on how much space to allocate them. The results of these decisions are then put in practice in every store. This model allows retailers to estimate the impact on sales revenues of space reallocations among business units. As an example, assume that a given retailer increases in 1% the area of the BU 06 – Home Cleaning in every store and withdraws it from BU 33 – Culture. Assume also that the actual average monthly sales areas of BU 06 and BU 33 are 35,000 sqm and 20,000 sqm respectively. This exchange is expected to increase the monthly retailer overall sales in 390,000 €.

In the second step of this stage, new variables characterizing store trade area are introduced, i.e. population, population density, purchasing power index and competition sales area. First, simple regression models were fitted to assess the impact of these variables on each business unit sales revenues, replicating model 6 above. All the fitted models were found to be significant, i.e. the p-value of their F-statistics is approximately zero and show an acceptable fit since the R^2 ranged from 58% to 59%. These results are displayed in Appendix 4.

Briefly, population, purchasing power index and competition were found to have a positive impact on most business units' performance while population density was found to have a negative impact. The sign of the impact of these variables is similar to that found in the simple regression models at the store level. However, it should be noted that the estimated parameters of most variables were found nonsignificant to explain business units' sales revenues for a substantial number of business units.

Since the estimated parameters of these variables were found to be significant for most business units, a multiple regression model is estimated next using the ordinary least squares method and considering all the aforementioned variables. The main objective is to understand their impact on the sales-space relationship. The regression equation is next.

Model 7:

$$\log \text{sales\_bu} = \beta_0 + \sum_{i=1}^{22} \beta_{0i} \text{bu}_i + \sum_{i=1}^{22} \beta_{1i} \log \text{areas\_bu}_i + \sum_{i=1}^{22} \beta_{2i} \log \text{population}_i + \sum_{i=1}^{22} \beta_{3i} \log \text{population density}_i + \sum_{i=1}^{22} \beta_{4i} \log \text{purchasing power index}_i + \sum_{i=1}^{22} \beta_{5i} \log \text{competitors'sales area}_i + \varepsilon \qquad (3.7)$$

sales_bu: business unit monthly sales (euros);

areas_bu$_i$: business unit monthly sales area for each business unit (square meters);

population$_i$**:** number of residents living in a store trade area for each business unit;

population density$_i$**:** number of residents per square meter living in a store trade area for each business unit;

competitors' sales area $_i$**:** sales area of competitor grocery retailer stores in store trade area (square meters) for each business unit;

β_{0i}**:** regression intercept for each business unit;

β_{1i}**:** regression coefficient of sales area for each business unit;

β_{2i}**:** regression coefficient of population for each business unit;

β_{3i}**:** regression coefficient of population density for each business unit;

β_{4i}**:** regression coefficient of purchasing power index for each business unit;

β_{5i}**:** regression coefficient of competitors' sales area for each business unit;

i: business unit;

ε: error term.

The full results of the fitted model are displayed in Appendix 5.

In this multiple regression model, denoted as model 7, business unit sales area was found to have a positive impact on business unit sales revenues for 18 out of the 23 business units. Business unit space-elasticity ranges from 0.20 to 1.12 which means that in these business units, a 1% increase (decrease) in their sales area induces a sales increase (decrease) between 0.12% and 1.12%. These results are similar to those estimated in the simple regression model, in which 17 out of the 23 business units were found to be affected by sales area and space-elasticities ranged from 0.39 to 1.17. However, those business units where space impact was found to be nonsignificant differed between models 6 and 7. The comparison of the estimated space-elasticities in both models is also shown below in table 42. There are two business units in which space impact is nonsignificant in both models – UN 15 Fruits and Vegetables and UN 32 Children Apparel. Above all, the estimated space-elasticities for each business unit are similar in both models exhibiting a great degree of stability. Therefore, we conclude that store trade area characteristics do not influence the sales-space relationship. Moreover, business unit sales area is the variable with the strongest impact on business unit sales revenues for most business units.

On the other hand, recall that competition was estimated to have a positive impact on business unit sales revenue in a simple regression model (Appendix 4). However, when considered in a multiple model with population and purchasing power index, its impact is negative, even though its estimated parameter is nonsignificant for 7 out of 23 business units. This sign change of the estimated impact between the simple regression model and the multiple regression models concerning sales area competition was also noted in previous sections at the store level.

Model 7- Top 5 Space-Elasticities				
BU	**BU Description**	**Parameters**	**Coefficient**	**P-Value**
BU 06	Home Cleaning	C	2.83	0.00
		log (areas_bu)	1.12	0.00
		log (population)	0.34	0.74
		log (pop. density)	-0.08	0.00
		log (ppi)	0.30	0.00
		log (comp. sales area)	-0.21	0.99
BU 43	Women Apparel	C	4.46	0.83
		log (areas_bu)	1.10	0.00
		log (population)	0.41	0.34
		log (pop. density)	-0.11	0.00
		log (ppi)	-0.34	0.00
		log (comp. sales area)	-0.24	0.71
BU 01	Savory	C	4.36	0.00
		log (areas_bu)	1.01	0.00
		log (population)	0.33	0.00
		log (pop. density)	0.00	0.67
		log (ppi)	0.07	0.15
		log (comp. sales area)	-0.21	0.00
BU 08	Dairy	C	4.68	0.31
		log (areas_bu)	0.96	0.05
		log (population)	0.28	0.29
		log (pop. density)	-0.03	0.07
		log (ppi)	0.16	0.18
		log (comp. sales area)	-0.15	0.09
BU 30	Leisure	C	6.54	0.00
		log (areas_bu)	0.95	0.00
		log (population)	0.08	0.00
		log (pop. density)	-0.12	0.00
		log (ppi)	-0.08	0.04
		log (comp. sales area)	-0.05	0.00

Table 40: Top 5 business unit space-elasticities – model 7

Model 7- Bottom 5 Space-Elasticities				
BU	**BU Description**	**Parameters**	**Coefficient**	**P-Value**
BU 19	Cafeteria	C	2.89	0.01
		log (areas_bu)	0.20	0.00
		log (population)	0.50	0.07
		log (pop. density)	-0.35	0.00
		log (ppi)	1.18	0.00
		log (comp. sales area)	-0.30	0.28
BU 33	Culture	C	1.73	0.00
		log (areas_bu)	0.59	0.00
		log (population)	0.21	0.02
		log (pop. density)	-0.12	0.00
		log (ppi)	1.01	0.00
		log (comp. sales area)	-0.03	0.00
BU 18	Take Away	C	2.94	0.00
		log (areas_bu)	0.62	0.00
		log (population)	0.26	0.17
		log (pop. density)	-0.10	0.00
		log (ppi)	0.64	0.00
		log (comp. sales area)	-0.06	0.00
BU 31	Home	C	1.60	0.00
		log (areas_bu)	0.80	0.00
		log (population)	0.25	0.10
		log (pop. density)	-0.11	0.00
		log (ppi)	0.71	0.00
		log (comp. sales area)	-0.08	0.00
BU 13	Cheese and Cold Meats	C	5.86	0.00
		log (areas_bu)	0.80	0.00
		log (population)	0.25	0.12
		log (pop. density)	-0.07	0.00
		log (ppi)	0.12	0.46
		log (comp. sales area)	-0.13	0.03

Table 41: Bottom 5 business unit space-elasticities – model 7

Space-elasticities at business unit level					
BU	BU description	model 6		model 7	
		Est.	P-value	Est.	P-value
BU01	Savory	1.03	0.00	1.01	0.00
BU02	Sweet Savory	0.99	0.17	0.95	0.02
BU03	Drinks	1.08	0.05	1.03	0.64
BU05	Hygiene and Beauty	1.08	0.03	1.02	0.70
BU06	Home Cleaning	1.17	0.00	1.12	0.00
BU07	Frozen	0.98	0.12	0.93	0.01
BU08	Dairy	0.99	0.19	0.96	0.05
BU11	Butchery	0.86	0.00	0.85	0.00
BU12	Fishery	1.12	0.00	1.05	0.21
BU13	Cheese and Cold Meats	0.86	0.00	0.80	0.00
BU15	Fruits and Vegetables	1.04	0.64	0.98	0.34
BU16	Bakery	0.99	0.19	0.92	0.00
BU18	Take Away	0.73	0.00	0.62	0.00
BU19	Cafeteria	0.39	0.00	0.20	0.00
BU30	Leisure	0.98	0.02	0.95	0.00
BU31	Home	0.80	0.00	0.80	0.00
BU33	Culture	0.59	0.00	0.59	0.00
BU34	Brico & Auto	0.81	0.00	0.80	0.00
BU35	Pet and Care	0.88	0.00	0.86	0.00
BU41	Baby Apparel	0.92	0.00	0.94	0.00
BU42	Children Apparel	0.99	0.24	1.00	0.68
BU43	Women Apparel	1.13	0.00	1.10	0.00
BU44	Men Apparel	0.87	0.00	0.83	0.00

Table 42: Summary of the estimated space-elasticities at the business unit level – models 6 and 7

This model shows a good fit since the R^2 is 91%, meaning that 91% of the log business units' variability is explained by the sales area, population, population density, purchasing power index and competitors' sales area.

Model 7 Statistics	
Statistics	Coefficient
R^2	0.911
Adjusted R^2	0.911
S.E. of Regression	0.502
Prob. (F-statistic)	0.000

Table 43: Statistics – model 7

The results obtained in models 5, 6 and 7 agree with the results obtained in the literature concerning the positive impact of space on product categories. However, some important differences occur. At the category level, Desmet and Renaudin (1998) study results showed that the average value of space-elasticity for all the product categories was 0.205, ranging from -0.44 to 0.80. Castro (2011) estimated an impact ranging from -0.24 to 1.84 at the category level. In this study, the estimated average space-elasticity was 1.08, ranging from 0.39 to 1.17. Therefore, no negative space-elasticity values were found, supporting the hypothesis of a positive impact of space on business unit sales revenues. Furthermore, the estimated values are substantially higher than those found by Desmet and Renaudin (1998) which also supports the positive magnitude of space impact on business units' sales revenues.

These models can support retailers' decisions on the space allocation process to business units, since they provide an estimate of the expected impact on revenues caused by business units' space changes. As an example, this retailer knows that changing the percentage of current space allocated to BU 06 – Home Cleaning is expected to increase (decrease) more than proportionally the percentage of this business unit's sales revenues, since the estimated impact found by model 6 was 1.17. On the other hand, the BU 33 – Culture exhibits a space-elasticity of 0.59, meaning that a 1% increase (decrease) of the business unit's sales area is expected to increase (decrease) its sales revenues by 0.59%. This assessment of retailer business units' space-elasticities provides the retailer important information to estimate trade-offs of space exchanges among business units.

However, these models are static which is a limitation in this study. In fact, they provide a picture of the business units' space-elasticities. If the retailer has the goal to increase in 50% the store sales area of BU 06 – Home cleaning, whose space-elasticity is 1.17, he cannot expect to increase sales revenues by 58.5%. In fact, as mentioned in Chapter 2, space-elasticity has decreasing marginal returns and takes the form of an S-Curve.

Nevertheless, these models were able to show the positive impact of space on business unit performance and the strong degree of stability of business unit space-elasticity when considered with other explanatory variables. Furthermore, it also provides the retailer an assessment of the expected trade-offs derived from space changes.

Chapter 4

Conclusions

In this work, we aimed at understanding and measuring the relationship between retailers' space and sales revenues. The focus was placed in two specific management activities: the space allocation process at the store and business unit levels. For this purpose, datasets provided by a Portuguese leading grocery retailer were used in order to apply a methodology divided into three stages.

In the first stage, we studied the direct relationship between space and sales revenues at the store level. Applying a simple regression model, using the panel least square method, a significant positive relationship was found between these two variables and a space-elasticity of 0.67 was calculated. Then, we considered the differences among different store formats and we applied a simple regression model using the ordinary least square method to measure the impact of space on store sales revenues for each store format. We also found a significant positive impact of space on sales revenues for every store format.

In the second stage, new independent variables related with stores trade area were introduced in the prior simple regression models. These variables intended to characterize two main factors influencing the store market potential: demand and competition. A geographical information system was used to calculate stores' attractiveness and to define stores' trade area. The values of these new variables were calculated for each store trade area helping to define the market potential for each store considered in the analysis. Adding these variables to store sales area allows an estimation of the store sales potential. To this purpose, multiple regression models were applied. These models allow store sales estimation helping retailers' in retail site location decisions, store performance assessment and in the store space definition. Firstly, a panel least square method was used to estimate a similar positive impact of space-elasticity (0.66). Secondly, an ordinary least square method was used to estimate the space impact for each store format. The results obtained were also similar to those found in the first stage.

Therefore, the impact of space on store sales revenues was found positive in both stages of the model and exhibited a high degree of stability. Moreover, among the variables considered, store space was found to have the strongest impact on store sales revenues.

In the final stage of the methodology, the scope of the analysis was extended to the business unit level. Firstly, a simple regression model was applied using the least squares method. The results found a significant positive impact of space on business unit sales revenues (1.08). Then, the differences among business units were determined and the correlation between space and sales revenues was found significant for 17 of 23 business units. Space-elasticities ranged from 0.39 to 1.17. Finally, store trade area characteristics were introduced in the prior model and the estimated space-elasticities were found to be similar to those obtained in the prior model and were found significant for 18 of 23 business units. Therefore, similarly to what was concluded in the analysis performed at the store level, space was found to have a positive impact on business unit sales revenues and space-elasticities showed a high degree of stability. Furthermore, space was found to be the variable with the strongest impact on store sales revenues for most business units.

Throughout this methodology, we met the objectives defined for this work and we demonstrated the importance of space management decisions for retailers' performance.

However, this study presents some limitations. Firstly, sales revenues were considered the only performance measure. Stores' performance depends on other variables, besides sales revenues, in order to become profitable and achieve long-term sustainability. Therefore, profitability indicators should be considered, such as space cost, in the store space allocation decision to assess stores' expected profitability. Secondly, it was not possible to consider other important store attributes influencing stores' performance such as store accessibility, visibility, customers' preferences, etc. This limitation is stronger in the business unit analysis since specific business unit attributes were not considered, such as product quality, assortment, service level, price level, promotion policies, etc. The introduction of these variables could help explain the non-significance of the space impact on sales revenues for some business units and assess the impact of those variables on the sales-space relationship. Finally, the fitted models are static and datasets must be updated with some frequency to accommodate changes in the variables. This limitation is stronger in the business unit analysis since store attributes do not change often. Furthermore, these models estimate a static space-elasticity for stores or business units and do not consider decreasing marginal returns. In fact, the estimated space-elasticity provides a picture of the present and cannot be applied when a store increases its area by 100%, instead of 1%, due to decreasing marginal returns.

Nevertheless, this work is the first to approach the space allocation problem at the store level considering store trade area characteristics. It has overcome the trade area measurement issue that had been identified has a major difficulty faced by academics and practitioners in store performance evaluation. The results obtained provide important insights to retailers concerning the impact of store trade area characteristics in their performance. The fitted models can support retail site location decisions by estimating sales revenues based on potential new stores characteristics. Furthermore, they also provide an assessment of stores and business units' space performance, with regard to their expected sales revenue variations induced by space changes. This assessment might lead retailers to take actions to improve their performance. Foremost, this work has shown that retailers' space is a resource that has a significant impact on retailers' performance and therefore retailers must find more effective ways of managing it.

Bibliography

Abbott, H., & Palekar, U. S.: 2008, Retail replenishment models with display-space elastic demand, *European Journal of Operational Research*, **186**(2), 586-607.

Alarcón Lorenzo, S.: 2011, The trade credit in the Spanish agrofood industry, *Mediterranean Journal of Economics, Agriculture and Environment (New Medit)*, **10**(2), 51-57.

Applebaum, W.: 1932, The Secondary Commercial Centers of Cincinnati. University of Cincinnati, Institute of Industrial Research.

Baviera-Puig, A., Castellanos, J., Buitrago, J. M., & Rodríguez, J. E.: 2011, Geomarketing: Determinación de las áreas de influencia de los supermercados. In VIII Congreso de Economía Agraria, Madrid, Spain.

Baviera-Puig, A., Buitrago-Vera, J., & Mas-Verdú, F.: 2012, Trade areas and knowledge-intensive services: The case of a technology centre. *Management Decision*, **50**(8), 1412-1424.

Borin, N., Farris, P. W., & Freeland, J. R.: 1994, A model for determining retail product category assortment and shelf space allocation, *Decision sciences*, **25**(3), 359-384.

Bridson, K., Evans, J., & Hickman, M.: 2008, Assessing the relationship between loyalty program attributes, store satisfaction and store loyalty, *Journal of Retailing and consumer Services*, **15**(5), 364-374.

Briesch, R. A., Chintagunta, P. K., & Fox, E. J.: 2009, How does assortment affect grocery store choice? , *Journal of Marketing Research*, **46**(2), 176-189.

Carpenter, J. M., & Moore, M.: 2006, Consumer demographics, store attributes, and retail format choice in the US grocery market. International Journal of Retail & Distribution Management, **34**(6), 434-452.

Castro, A. C. B.: 2011, As Vendas e o Espaço no Retalho: Modelos Econométricos Aplicados a um Grupo de Distribuição Alimentar Português.

Cho, H., & Fiorito, S. S.: 2010, Self-service technology in retailing. the case of retail kiosks, *Symphonya. Emerging Issues in Management*, (1), 43-55.

Church, R. L.: 2002, Geographical information systems and location science, *Computers & Operations Research*, **29**(6), 541-562.

Clarke, I., Mackaness, W., & Ball, B.: 2003, Modelling intuition in retail site assessment (MIRSA): making sense of retail location using retailers' intuitive judgements as a support for decision-making, *The International Review of Retail, Distribution and Consumer Research*, **13**(2), 175-193.

Cox, K. K.: 1970, The effect of shelf space upon sales of branded products, *Journal of Marketing Research*, 55-58.

Corstjens, M., & Doyle, P.: 1981, A model for optimizing retail space allocations, *Management Science*, **27**(7), 822-833.

Curhan, R. C.: 1972, The relationship between shelf space and unit sales in supermarkets, *Journal of Marketing Research*, 406-412.

Davidson, W. R., Sweeney, D. J., & Stampfl, R. W.: 1984, Retailing management., Wiley.

Davies, R. L.: 1977, Store location and store assessment research: the integration of some new and traditional techniques. *Transactions of the Institute of British Geographers*, 141-157.

Desmet, P., & Renaudin, V.: 1998, Estimation of product category sales responsiveness to allocated shelf space, *International Journal of Research in Marketing*, **15**(5), 443-457.

Dhar, S. K., & Hoch, S. J.: 1997, Why store brand penetration varies by retailer, *Marketing Science*, **16**(3), 208-227.

Fernie, J., Fernie, S., & Moore, C.:2015, Principles of retailing. Routledge.

García-Palomares, J. C., Gutiérrez, J., & Latorre, M.: 2012, Optimizing the location of stations in bike-sharing programs: A GIS approach, *Applied Geography*, **35**(1-2), 235-246.

Hernandez, T., & Bennison, D.: 2000, The art and science of retail location decisions, *International Journal of Retail & Distribution Management*, **28**(8), 357-367.

Hernandez, T.: 2007, Enhancing retail location decision support: The development and application of geovisualization, *Journal of Retailing and Consumer Services*, **14**(4), 249-258.

Hoffman, K. D., & Turley, L. W.: 2002, Atmospherics, service encounters and consumer decision making: An integrative perspective, *Journal of Marketing theory and practice*, **10**(3), 33-47.

Jallais, J., Orsoni, J., Fady, A., & do Céu Pedreño, M.:1993, O marketing da distribuição: aplicação ao ponto de venda.

Jayasankara Prasad, C., & Ramachandra Aryasri, A.: 2011, Effect of shopper attributes on retail format choice behaviour for food and grocery retailing in India, *International Journal of Retail & Distribution Management*, **39**(1), 68-86.

Jones, M. A. (1999).: Entertaining shopping experiences: an exploratory investigation, *Journal of retailing and consumer services*, **6**(3), 129-139.

Kamakura, W. A., & Kang, W.: 2007, Chain-wide and store-level analysis for cross-category management, *Journal of Retailing*, **83**(2), 159-170.

Marques, M., Castro, J.M., Silva, V., Carvalho, A. and Monteiro, R.: 2014, Risco sísmico em Portugal – desenvolvimentos do projeto PRISE", *5as Jornadas Portuguesas de Engenharia de Estruturas*, pp. 1-16

Mitchell, V. W., & Harris, G.: 2005, The importance of consumers' perceived risk in retail strategy, *European Journal of marketing*, **39**(7/8), 821-837.

Nilsson, E., Gärling, T., Marell, A., & Nordvall, A. C.: 2015, Importance ratings of grocery store attributes, *International Journal of Retail & Distribution Management*, **43**(1), 63-91.

Nogales, A. F., & Suarez, M. G.: 2005, Shelf space management of private labels: a case study in Spanish retailing, *Journal of retailing and consumer services*, **12**(3), 205-216.

Phillips, H., & Bradshaw, R.: 1993, How customers actually shop: customer interaction with the point of sale, *Market Research Society Journal.*, **35**(1), 1-10.

Prasad, C., & Aryasri, A.: 2011, Effect of shopper attributes on retail format choice behaviour for food and grocery retailing in India. International Journal of Retail & Distribution Management, **39**(1), 68-86.

Ramaseshan, B., Achuthan, N. R., & Collinson, R.: 2009, A retail category management model integrating shelf space and inventory levels, *Asia-Pacific Journal of Operational Research*, **26**(04), 457-478.

Reutterer, T., & Teller, C.: 2009, Store format choice and shopping trip types, *International Journal of Retail & Distribution Management*, **37**(8), 695-710.

Rogers, D. S., & Green, H. L.: 1979, A new perspective on forecasting store sales: applying statistical models and techniques in the analog approach, *Geographical Review*, 449-458.

Roig-Tierno, N., Baviera-Puig, A., Buitrago-Vera, J., & Mas-Verdu, F.: 2013, The retail site location decision process using GIS and the analytical hierarchy process, *Applied Geography*, 40, 191-198.

Silva, A. L., & Cardoso, M. G: 2005, Predicting supermarket sales: The use of regression trees, *Journal of Targeting*, Measurement and Analysis for Marketing, **13**(3), 239-249.

Sinigaglia, N.: 1997, Measuring retail units efficiency: a technical approach (Doctoral dissertation, Louvain School of Management).

Thurik, R.: 1988, Les grandes surfaces en France: étude de la relation ventes/surface du magasin, *Recherche et Applications en Marketing (French Edition)*, **3**(3), 21-37.

Wong, A., & Dean, A.: 2009. Enhancing value for Chinese shoppers: The contribution of store and customer characteristics, *Journal of retailing and consumer services*, **16**(2), 123-134.

Wood, S., & Tasker, A.: 2008, The importance of context in store forecasting: the site visit in retail location decision-making, *Journal of Targeting, Measurement and Analysis for Marketing*, **16**(2), 139-155.

Wood, S., & Reynolds, J.: 2012, Leveraging locational insights within retail store development? Assessing the use of location planners' knowledge in retail marketing, Geoforum, **43**(6), 1076-1087.

Woodside, A. G., & Trappey, R. J.: 1992, Finding out why customers shop your store and buy your brand: Automatic cognitive processing models of primary choice, *Journal of Advertising Research*.

Yang, M. H., & Chen, W. C.: 1999, A study on shelf space allocation and management, *International journal of production economics*, **60**, 309-317.

Appendix

1. Store Effects for the Regression Models estimated using the panel least square method (models 1 and 3)

Store Effects			
Store	**Model 1**	**Model 3**	**Difference**
L0261	-9.3%	15.0%	24.3%
L1169	-51.7%	-27.8%	23.9%
L0273	3.7%	26.8%	23.1%
L1902	-46.0%	-25.7%	20.3%
L0389	-18.6%	0.2%	18.8%
L0867	-78.7%	-62.0%	16.8%
L0263	3.9%	19.7%	15.8%
L0847	-5.9%	9.5%	15.4%
L0315	-69.0%	-53.7%	15.3%
L2089	40.0%	54.6%	14.7%
L0927	21.5%	35.1%	13.6%
L0272	5.5%	18.9%	13.4%
L2081	-51.2%	-37.9%	13.3%
L0294	39.8%	53.0%	13.2%
L0317	-84.1%	-70.9%	13.1%
L2087	-50.3%	-37.1%	13.1%
L0328	-54.8%	-43.6%	11.2%
L0253	-35.6%	-25.0%	10.7%
L1984	24.7%	34.9%	10.2%
L0379	20.0%	30.1%	10.1%

Store Effects			
Store	**Model 1**	**Model 3**	**Difference**
L0309	6.7%	16.6%	9.9%
L0249	2.6%	12.1%	9.5%
L1503	-18.1%	-9.1%	9.0%
L2076	-37.5%	-28.6%	8.9%
L0265	-47.6%	-39.1%	8.5%
L0331	8.6%	16.7%	8.2%
L2085	-50.4%	-42.2%	8.2%
L0845	29.5%	37.6%	8.2%
L0335	-38.4%	-30.4%	8.0%
L2084	-9.5%	-1.6%	7.9%
L0313	-32.6%	-24.8%	7.8%
L0330	-57.5%	-50.0%	7.6%
L1004	-7.3%	0.2%	7.5%
L0324	-31.8%	-24.3%	7.5%
L0010	21.1%	28.5%	7.4%
L0325	-25.0%	-17.6%	7.3%
L0260	-37.8%	-30.5%	7.3%
L0228	-24.8%	-17.8%	7.0%
L1057	-31.8%	-24.9%	6.9%
L0296	1.4%	8.3%	6.9%

Store Effects			
Store	**Model 1**	**Model 3**	**Difference**
L0206	2.8%	9.6%	6.8%
L0305	-71.7%	-65.1%	6.6%
L1704	-59.0%	-52.4%	6.6%
L1504	-15.7%	-9.3%	6.4%
L0245	15.4%	21.6%	6.2%
L0283	-8.9%	-2.8%	6.1%
L0233	-11.5%	-5.6%	5.9%
L0235	-15.8%	-9.9%	5.9%
L0319	30.8%	36.6%	5.8%
L1392	-1.1%	4.6%	5.7%
L0251	4.9%	10.6%	5.7%
L1415	7.9%	13.3%	5.4%
L0250	22.8%	28.1%	5.3%
L0271	24.7%	29.8%	5.2%
L0304	-38.0%	-32.8%	5.1%
L1058	39.3%	44.3%	4.9%
L0266	-15.1%	-10.2%	4.9%
L0280	-6.9%	-2.0%	4.9%
L0284	-40.0%	-35.1%	4.9%
L0314	-61.7%	-57.5%	4.2%

Store Effects			
Store	**Model 1**	**Model 3**	**Difference**
L0322	-11.3%	-7.2%	4.2%
L0216	-10.2%	-6.2%	4.0%
L1000	30.5%	34.5%	3.9%
L0293	1.7%	5.4%	3.7%
L0222	12.1%	15.6%	3.5%
L0232	17.1%	20.4%	3.3%
L0318	-60.0%	-56.8%	3.2%
L0201	13.2%	16.3%	3.1%
L0226	0.6%	3.7%	3.0%
L0312	-49.5%	-46.6%	3.0%
L0262	21.9%	23.9%	2.1%
L0327	6.9%	8.9%	2.0%
L0297	-16.3%	-14.7%	1.6%
L1501	-35.1%	-33.5%	1.5%
L0256	53.9%	55.4%	1.5%
L1707	-1.2%	0.2%	1.4%
L0277	-34.0%	-32.7%	1.4%
L0298	-18.0%	-16.9%	1.1%
L0289	3.5%	4.5%	1.0%
L2090	-46.0%	-45.0%	1.0%

Store Effects			
Store	**Model 1**	**Model 3**	**Difference**
L1397	12.3%	13.1%	0.8%
L0334	-22.7%	-21.9%	0.8%
L0218	17.7%	18.3%	0.6%
L0244	32.4%	33.0%	0.6%
L0282	-49.8%	-49.2%	0.6%
L2083	-15.3%	-14.9%	0.4%
L0278	-1.0%	-0.6%	0.4%
L0240	0.1%	0.4%	0.3%
L0213	-12.0%	-11.9%	0.1%
L1051	8.0%	8.1%	0.1%
L0375	-7.3%	-7.4%	-0.1%
L0329	-22.6%	-22.9%	-0.3%
L1702	-36.9%	-37.2%	-0.3%
L0259	11.4%	10.9%	-0.5%
L0202	51.8%	51.3%	-0.5%
L0234	-24.0%	-24.7%	-0.7%
L1393	4.9%	4.1%	-0.8%
L0258	29.0%	28.1%	-0.9%
L0211	2.6%	1.5%	-1.1%
L1056	-31.1%	-32.1%	-1.1%

Store Effects			
Store	**Model 1**	**Model 3**	**Difference**
L0439	-17.3%	-18.4%	-1.1%
L0013	28.0%	26.8%	-1.2%
L0842	7.5%	6.3%	-1.2%
L0238	1.3%	-0.1%	-1.3%
L0257	24.2%	22.8%	-1.4%
L0323	13.4%	12.0%	-1.4%
L0219	-6.4%	-7.9%	-1.4%
L0237	-32.8%	-34.3%	-1.5%
L1706	27.7%	26.2%	-1.5%
L0381	-30.7%	-32.4%	-1.6%
L0291	-23.6%	-25.3%	-1.7%
L0288	44.9%	43.0%	-1.8%
L0320	-15.3%	-17.2%	-1.8%
L0241	29.7%	27.5%	-2.2%
L0459	35.6%	33.3%	-2.2%
L1978	-12.5%	-14.8%	-2.3%
L0247	5.5%	3.2%	-2.3%
L0268	-26.6%	-29.0%	-2.4%
L2082	-36.9%	-39.3%	-2.4%
L0321	-42.3%	-44.7%	-2.5%

Store Effects			
Store	**Model 1**	**Model 3**	**Difference**
L1391	-36.8%	-39.4%	-2.6%
L0229	-8.0%	-10.7%	-2.7%
L0471	-87.9%	-90.7%	-2.8%
L0239	4.7%	1.8%	-2.9%
L0221	-20.0%	-22.9%	-2.9%
L1053	-0.2%	-3.2%	-2.9%
L0209	53.3%	50.4%	-2.9%
L0333	-8.7%	-11.9%	-3.2%
L0214	55.6%	52.3%	-3.3%
L0224	14.9%	11.1%	-3.8%
L0236	11.2%	7.1%	-4.0%
L1703	-13.8%	-17.9%	-4.1%
L0207	43.5%	39.3%	-4.2%
L0003	85.5%	81.3%	-4.3%
L0446	-12.3%	-16.7%	-4.3%
L0462	11.4%	7.1%	-4.4%
L0279	13.4%	8.7%	-4.6%
L0215	32.5%	27.6%	-4.9%
L0940	-41.5%	-46.5%	-5.0%
L0269	31.4%	26.3%	-5.1%

Store Effects			
Store	**Model 1**	**Model 3**	**Difference**
L0217	43.7%	38.6%	-5.1%
L0466	-11.9%	-17.1%	-5.2%
L0230	8.3%	3.0%	-5.2%
L0001	71.6%	65.9%	-5.7%
L0270	-12.0%	-18.0%	-6.1%
L0208	-10.0%	-16.1%	-6.1%
L0461	35.1%	28.8%	-6.3%
L0006	65.0%	58.6%	-6.4%
L0252	2.9%	-3.5%	-6.4%
L0299	14.6%	8.2%	-6.5%
L1008	-23.3%	-29.9%	-6.5%
L0460	7.8%	1.2%	-6.6%
L0210	41.0%	34.3%	-6.7%
L0212	41.7%	34.9%	-6.8%
L0220	21.4%	14.5%	-6.9%
L0295	-9.5%	-16.6%	-7.1%
L0205	34.3%	27.2%	-7.1%
L0267	7.6%	0.2%	-7.4%
L0468	-13.0%	-20.4%	-7.4%
L0290	-9.7%	-17.3%	-7.6%

Store Effects			
Store	**Model 1**	**Model 3**	**Difference**
L0204	40.8%	32.7%	-8.1%
L0463	85.1%	76.9%	-8.1%
L0340	-36.1%	-44.5%	-8.4%
L0008	27.2%	18.6%	-8.5%
L0281	31.4%	22.9%	-8.6%
L0009	55.6%	46.9%	-8.7%
L0494	-2.6%	-11.2%	-8.7%
L0223	51.9%	43.0%	-8.8%
L0016	-5.5%	-14.8%	-9.3%
L1055	-45.6%	-55.0%	-9.4%
L1054	11.9%	1.8%	-10.1%
L0002	66.2%	56.1%	-10.1%
L0458	16.4%	6.0%	-10.3%
L0014	20.5%	10.0%	-10.5%
L0255	0.9%	-9.8%	-10.7%
L0464	72.2%	61.4%	-10.8%

Store Effects			
Store	**Model 1**	**Model 3**	**Difference**
L0326	-8.1%	-19.9%	-11.7%
L0246	40.9%	29.1%	-11.8%
L0203	68.0%	55.9%	-12.1%
L0465	32.8%	20.5%	-12.3%
L0843	8.5%	-3.9%	-12.4%
L0007	70.0%	56.7%	-13.3%
L0005	48.1%	34.7%	-13.4%
L0275	-0.1%	-13.5%	-13.4%
L0012	76.3%	62.3%	-14.0%
L0231	-15.0%	-29.5%	-14.5%
L0004	62.1%	47.3%	-14.8%
L0307	-5.7%	-21.8%	-16.1%
L0242	-6.6%	-26.3%	-19.7%
L0248	-32.3%	-52.6%	-20.3%
L0303	1.9%	-18.6%	-20.5%
L0011	7.5%	-14.7%	-22.2%

2. Descriptive Statistics of each Portuguese retailer store format

	Continente					
Descriptive Statistics	Store Sales	Store Sales Area	Population	Pop. Density	Competitors Sales Area	PPI
Average	3,176,247	7,017	185,781	1,901	43,153	117
Median	2,776,823	7,410	112,241	623	26,644	103
Maximum	9,255,335	15,822	1,604,922	8,406	343,295	217
Minimum	841,056	3,069	22,841	37	4,497	85
St. Deviation	1,595,194	2,778	180,791	2,190	42,976	35
Skewness	1	1	2	1	2	2
Kurtosis	0	0	8	1	6	3
Sample Size	480	480	480	480	480	480

	Continente Modelo					
Descriptive Statistics	Store Sales	Store Sales Area	Population	Pop. Density	Competitors Sales Area	PPI
Average	944,455	2,069	55,350	2,000	11,434	89
Median	886,325	2,049	45,220	551	7,950	86
Maximum	3,398,974	3,860	732,699	65,397	132,215	137
Minimum	286,742	1,131	9,564	15	800	53
St. Deviation	330,842	328	45,821	5,717	10,237	16
Skewness	1	1	5	7	3	0
Kurtosis	4	8	51	60	27	0
Sample Size	1,344	1,344	1,344	1,344	1,344	1,344

	Continente Bom Dia					
Descriptive Statistics	Store Sales	Store Sales Area	Population	Pop. Density	Competitors Sales Area	PPI
Average	468,830	996	149,239	5,228	34,315	136
Median	416,781	934	101,497	5 354	21,349	112
Maximum	1,198,333	1,619	864,812	10,883	214,363	217
Minimum	207,988	474	17,718	94	800	57
St. Deviation	205,812	329	157,534	2,492	41,596	46
Skewness	1	0	3	0	3	1
Kurtosis	1	-1	10	0	8	-1
Sample Size	432	432	432	432	432	432

3. Statistics of the Simple Regression Models of the sales area and trade area variables introduced in model 4

sales area	
Statistics	Coefficient
R^2	0.850
Adjusted R^2	0.849
S.E. of Regression	0.280
Prob. (F-statistic)	0.000

population	
Statistics	Coefficient
R^2	0.760
Adjusted R^2	0.758
S.E. of Regression	0.354
Prob. (F-statistic)	0.000

pop. density	
Statistics	Coefficient
R^2	0.745
Adjusted R^2	0.743
S.E. of Regression	0.365
Prob. (F-statistic)	0.000

ppi	
Statistics	Coefficient
R^2	0.753
Adjusted R^2	0.753
S.E. of Regression	0.358
Prob. (F-statistic)	0.000

competitors sales area	
Statistics	Coefficient
R^2	0.746
Adjusted R^2	0.744
S.E. of Regression	0.365
Prob. (F-statistic)	0.000

4. Results and Statistics of the Simple Regression Models of the Trade Area Variables Introduced in Model 7

Simple Regression on population			
BU	**Parameters**	**Est.**	**P-value**
BU01	C log (population)	8.82 0.23	0.00 0.00
BU02	C log (population)	8.86 0.23	0.93 0.96
BU03	C log (population)	8.65 0.25	0.67 0.60
BU05	C log (population)	8.73 0.23	0.83 0.98
BU06	C log (population)	9.00 0.18	0.67 0.15
BU07	C log (population)	7.18 0.26	0.00 0.40
BU08	C log (population)	8.98 0.22	0.70 0.81
BU11	C log (population)	9.29 0.16	0.24 0.05
BU12	C log (population)	9.22 0.18	0.32 0.16
BU13	C log (population)	8.28 0.25	0.17 0.55
BU15	C log (population)	7.95 0.30	0.03 0.05
BU16	C log (population)	7.97 0.24	0.03 0.81

Simple Regression on population			
BU	**Parameters**	**Est.**	**P-value**
BU18	C log (population)	5.90 0.35	0.00 0.00
BU19	C log (population)	6.89 0.23	0.02 0.94
BU30	C log (population)	8.24 0.05	0.15 0.00
BU31	C log (population)	8.74 0.09	0.84 0.00
BU33	C log (population)	7.81 0.20	0.01 0.35
BU34	C log (population)	10.28 -0.10	0.00 0.00
BU35	C log (population)	8.31 0.15	0.20 0.02
BU41	C log (population)	-2.36 0.78	0.00 0.00
BU42	C log (population)	-0.13 0.80	0.00 0.00
BU43	C log (population)	-8.53 1.34	0.00 0.00
BU44	C log (population)	-0.21 0.84	0.00 0.00

Statistics	
Statistics	Coefficient
R^2	0.593
Adjusted R^2	0.593
S.E. of Regression	1.071
Prob. (F-statistic)	0.000

Simple Regression on Pop. Density			
BU	**Parameters**	**Est.**	**P-value**
BU01	C log (pop.density)	11.62 -0.04	0.00 0.02
BU02	C log (pop.density)	11.74 -0.05	0.77 0.50
BU03	C log (pop.density)	11.65 -0.03	0.17 0.93
BU05	C log (pop.density)	11.92 -0.09	0.85 0.00
BU06	C log (pop.density)	11.65 -0.10	0.76 0.00
BU07	C log (pop.density)	10.09 0.00	0.00 0.10
BU08	C log (pop.density)	11.75 -0.04	0.92 0.66
BU11	C log (pop.density)	11.45 -0.06	0.30 0.33
BU12	C log (pop.density)	11.74 -0.08	0.65 0.05
BU13	C log (pop.density)	11.26 -0.03	0.04 0.72
BU15	C log (pop.density)	11.23 0.01	0.00 0.03
BU16	C log (pop.density)	10.83 -0.03	0.00 0.83

Simple Regression on Pop. Density			
BU	**Parameters**	**Est.**	**P-value**
BU18	C log (pop.density)	9.62 0.02	0.00 0.01
BU19	C log (pop.density)	10.45 -0.11	0.00 0.14
BU30	C log (pop.density)	11.47 -0.40	0.00 0.00
BU31	C log (pop.density)	11.63 -0.28	0.04 0.00
BU33	C log (pop.density)	10.86 -0.13	0.00 0.00
BU34	C log (pop.density)	11.57 -0.36	0.00 0.00
BU35	C log (pop.density)	11.06 -0.17	0.61 0.00
BU41	C log (pop.density)	6.76 -0.08	0.00 0.03
BU42	C log (pop.density)	7.22 0.30	0.00 0.00
BU43	C log (pop.density)	13.51 -0.85	0.07 0.00
BU44	C log (pop.density)	7.49 0.32	0.00 0.00

Statistics	
Statistics	Coefficient
R^2	0.590
Adjusted R^2	0.589
S.E. of Regression	1.075
Prob. (F-statistic)	0.000

Simple Regression on PPI			
BU	**Parameters**	**Est.**	**P-value**
BU01	C log (ppi)	10.70 0.15	0.00 0.74
BU02	C log (ppi)	10.54 0.19	0.77 0.15
BU03	C log (ppi)	9.96 0.32	0.17 1.00
BU05	C log (ppi)	10.60 0.15	0.85 0.29
BU06	C log (ppi)	10.86 0.03	0.76 0.03
BU07	C log (ppi)	8.19 0.41	0.00 0.83
BU08	C log (ppi)	10.64 0.17	0.92 0.11
BU11	C log (ppi)	11.25 -0.04	0.30 0.46
BU12	C log (ppi)	10.94 0.06	0.65 0.14
BU13	C log (ppi)	9.59 0.32	0.04 0.00
BU15	C log (ppi)	8.21 0.67	0.00 0.02
BU16	C log (ppi)	8.69 0.42	0.00 0.00

Simple Regression on PPI			
BU	**Parameters**	**Est.**	**P-value**
BU18	C log (ppi)	5.77 0.86	0.00 0.00
BU19	C log (ppi)	4.97 0.96	0.00 0.00
BU30	C log (ppi)	12.96 -0.90	0.00 0.00
BU31	C log (ppi)	11.79 -0.44	0.04 0.08
BU33	C log (ppi)	8.35 0.36	0.00 0.00
BU34	C log (ppi)	14.03 -1.06	0.00 0.03
BU35	C log (ppi)	10.42 -0.11	0.61 0.00
BU41	C log (ppi)	0.91 1.16	0.00 0.00
BU42	C log (ppi)	-2.08 2.42	0.00 0.00
BU43	C log (ppi)	9.28 -0.45	0.07 0.00
BU44	C log (ppi)	-1.87 2.46	0.00 0.00

Statistics	
Statistics	Coefficient
R^2	0.574
Adjusted R^2	0.574
S.E. of Regression	1.096
Prob. (F-statistic)	0.000

Simple Regression on competititors' sales area			
BU	**Parameters**	**Est.**	**P-value**
BU01	C	9.75	0.00
	log (comp. sales area)	0.17	0.00
BU02	C	9.71	0.90
	log (comp. sales area)	0.18	0.85
BU03	C	9.53	0.47
	log (comp. sales area)	0.20	0.39
BU05	C	9.54	0.50
	log (comp. sales area)	0.18	0.73
BU06	C	9.69	0.86
	log (comp. sales area)	0.14	0.25
BU07	C	8.09	0.00
	log (comp. sales area)	0.21	0.24
BU08	C	9.82	0.81
	log (comp. sales area)	0.17	0.97
BU11	C	9.99	0.41
	log (comp. sales area)	0.11	0.06
BU12	C	9.92	0.57
	log (comp. sales area)	0.14	0.27
BU13	C	9.18	0.06
	log (comp. sales area)	0.20	0.40
BU15	C	8.96	0.01
	log (comp. sales area)	0.25	0.02
BU16	C	8.71	0.00
	log (comp. sales area)	0.20	0.37

Simple Regression on competititors' sales area			
BU	**Parameters**	**Est.**	**P-value**
BU18	C	6.97	0.00
	log (comp. sales area)	0.29	0.00
BU19	C	7.81	0.00
	log (comp. sales area)	0.17	0.98
BU30	C	8.13	0.00
	log (comp. sales area)	0.07	0.00
BU31	C	8.88	0.00
	log (comp. sales area)	0.10	0.01
BU33	C	8.32	0.00
	log (comp. sales area)	0.18	0.89
BU34	C	9.95	0.49
	log (comp. sales area)	-0.08	0.00
BU35	C	8.91	0.01
	log (comp. sales area)	0.11	0.04
BU41	C	-0.20	0.00
	log (comp. sales area)	0.68	0.00
BU42	C	2.16	0.00
	log (comp. sales area)	0.69	0.00
BU43	C	-4.70	0.00
	log (comp. sales area)	1.16	0.00
BU44	C	2.03	0.00
	log (comp. sales area)	0.74	0.00

Statistics	
Statistics	Coefficient
R^2	0.590
Adjusted R^2	0.590
S.E. of Regression	1.075
Prob. (F-statistic)	0.000

5. Full Results of the Multiple Regression Model fitted at the business unit level (model 7)

Model 7- results				
BU	**BU Description**	**Parameters**	**Estimate**	**P-value**
BU 01	Savory	C	4.36	0.00
		log (areas_bu)	1.01	0.00
		log (population)	0.33	0.00
		log (pop. density)	0.00	0.67
		log (ppi)	0.07	0.15
		log (comp. sales area)	-0.21	0.00
BU 02	Sweet Savory	C	4.01	0.00
		log (areas_bu)	0.95	0.02
		log (population)	0.28	0.36
		log (pop. density)	-0.04	0.01
		log (ppi)	0.19	0.07
		log (comp. sales area)	-0.14	0.04
BU 03	Drinks	C	3.54	0.01
		log (areas_bu)	1.03	0.64
		log (population)	0.26	0.14
		log (pop. density)	-0.05	0.00
		log (ppi)	0.26	0.00
		log (comp. sales area)	-0.13	0.03
BU 05	Hygiene and Beauty	C	3.29	0.00
		log (areas_bu)	1.02	0.70
		log (population)	0.28	0.36
		log (pop. density)	-0.08	0.00
		log (ppi)	0.33	0.00
		log (comp. sales area)	-0.13	0.02
BU 06	Home Cleaning	C	2.83	0.00
		log (areas_bu)	1.12	0.00
		log (population)	0.34	0.74
		log (pop. density)	-0.08	0.00
		log (ppi)	0.30	0.00
		log (comp. sales area)	-0.21	0.99
BU 07	Frozen		2.56	0.00
		log (areas_bu)	0.93	0.01
		log (population)	0.29	0.42
		log (pop. density)	-0.03	0.05
		log (ppi)	0.27	0.00
		log (comp. sales area)	-0.10	0.00
BU 08	Dairy	C	4.68	0.31
		log (areas_bu)	0.96	0.05
		log (population)	0.28	0.29
		log (pop. density)	-0.03	0.07
		log (ppi)	0.16	0.18
		log (comp. sales area)	-0.15	0.09
BU 11	Butchery		6.30	0.00
		log (areas_bu)	0.85	0.00
		log (population)	0.19	0.00
		log (pop. density)	-0.01	0.80
		log (ppi)	0.08	0.91
		log (comp. sales area)	-0.13	0.03

Model 7- results				
BU	**BU Description**	**Parameters**	**Estimate**	**P-value**
BU 12	Fishery	C	4.82	0.15
		log (areas_bu)	1.05	0.21
		log (population)	0.31	0.68
		log (pop. density)	-0.08	0.00
		log (ppi)	0.01	0.34
		log (comp. sales area)	-0.18	0.31
BU 13	Cheese and Cold Meats	C	5.86	0.00
		log (areas_bu)	0.80	0.00
		log (population)	0.25	0.12
		log (pop. density)	-0.07	0.00
		log (ppi)	0.12	0.46
		log (comp. sales area)	-0.13	0.03
BU 15	Fruits and Vegetables	C	2.90	0.00
		log (areas_bu)	0.98	0.34
		log (population)	0.18	0.00
		log (pop. density)	-0.03	0.06
		log (ppi)	0.54	0.00
		log (comp. sales area)	-0.09	0.00
BU 16	Bakery	C	3.97	0.23
		log (areas_bu)	0.92	0.00
		log (population)	0.24	0.08
		log (pop. density)	-0.10	0.00
		log (ppi)	0.39	0.00
		log (comp. sales area)	-0.13	0.03
BU 18	Take Away	C	2.94	0.00
		log (areas_bu)	0.62	0.00
		log (population)	0.26	0.17
		log (pop. density)	-0.10	0.00
		log (ppi)	0.64	0.00
		log (comp. sales area)	-0.06	0.00
BU 19	Cafeteria		2.89	0.01
		log (areas_bu)	0.20	0.00
		log (population)	0.50	0.07
		log (pop. density)	-0.35	0.00
		log (ppi)	1.18	0.00
		log (comp. sales area)	-0.30	0.28
BU 30	Leisure	C	6.54	0.00
		log (areas_bu)	0.95	0.00
		log (population)	0.08	0.00
		log (pop. density)	-0.12	0.00
		log (ppi)	-0.08	0.04
		log (comp. sales area)	-0.05	0.00
BU 31	Home		1.60	0.00
		log (areas_bu)	0.80	0.00
		log (population)	0.25	0.10
		log (pop. density)	-0.11	0.00
		log (ppi)	0.71	0.00
		log (comp. sales area)	-0.08	0.00

Model 7- results				
BU	**BU Description**	**Parameters**	**Estimate**	**P-value**
BU 33	Culture	C	1.73	0.00
		log (areas_bu)	0.59	0.00
		log (population)	0.21	0.02
		log (pop. density)	-0.12	0.00
		log (ppi)	1.01	0.00
		log (comp. sales area)	-0.03	0.00
BU 34	Brico & Auto	C	3.79	0.08
		log (areas_bu)	0.80	0.00
		log (population)	0.14	0.00
		log (pop. density)	-0.08	0.00
		log (ppi)	0.48	0.00
		log (comp. sales area)	-0.09	0.00
BU 35	Pet and Care	C	4.63	0.39
		log (areas_bu)	0.86	0.00
		log (population)	0.25	0.13
		log (pop. density)	-0.05	0.01
		log (ppi)	0.21	0.04
		log (comp. sales area)	-0.18	0.40
BU 41	Baby Apparel	C	7.43	0.00
		log (areas_bu)	0.94	0.00
		log (population)	0.56	0.00
		log (pop. density)	-0.22	0.00
		log (ppi)	-0.83	0.00
		log (comp. sales area)	-0.36	0.00
BU 42	Children Apparel	C	3.18	0.08
		log (areas_bu)	1.00	0.68
		log (population)	0.41	0.52
		log (pop. density)	-0.08	0.04
		log (ppi)	0.58	0.00
		log (comp. sales area)	-0.51	0.00
BU 43	Women Apparel		4.46	0.83
		log (areas_bu)	1.10	0.00
		log (population)	0.41	0.34
		log (pop. density)	-0.11	0.00
		log (ppi)	-0.34	0.00
		log (comp. sales area)	-0.24	0.71
BU 44	Men Apparel	C	3.36	0.14
		log (areas_bu)	0.83	0.00
		log (population)	0.13	0.14
		log (pop. density)	0.09	0.02
		log (ppi)	0.50	0.00
		log (comp. sales area)	-0.19	0.81