UNIVERSITÀ DEGLI STUDI DI
CASSINO E DEL LAZIO MERIDIONALE
Dipartimento di Economia e Giurisprudenza

Corso di Dottorato in

Istituzioni, Mercati e Comportamenti

curriculum Istituzioni e Azienda

Ciclo XXXII

Impact of the retail environment drivers on sales and demand forecasting

SSD: SECS-P/08

Coordinatore del Corso
Chiar.ma Prof.ssa Rosella Tomassoni

Dottoranda
Maja Božić

Supervisore
Chiar.mo Prof. Andrea Moretta Tartaglione

Anno accademico 2019/2020

Мојој мами, најсјајнијој звезди на небу

ACKNOWLEDGEMENTS

Concluding the Ph.D. dissertation gives me a blend of joy, pride, and sadness. This unique experience challenges me personally and professionally, changing the life and self perspectives, forming the values to become a better person. In these years I have the pleasure to meet and work with kind individuals at the University of Cassino and abroad who support and help with ideas behind the research work. I would like to express and extend my sincere thanks to all of them.

Foremost, I'd like to express my thanks to my encouraging tutor, Professor Andrea Moretta Tartaglione, who has supported me throughout this research project. I am extremely grateful for our friendly chats and your enthusiastic support to pursue the academic and personal experience, also via the Erasmus exchange program. I truly enjoyed the vibrant and positive environment in which open and original thinking is appreciated and where connecting and sharing ideas with others are stimulated, which my tutor created.

I would like to thank, profoundly, Professor Philippe Baecke from Vlerick Business School in Ghent (Belgium) who provided valuable input in the last project. His patient, detailed and determined guidance helps me to improve my knowledge and skills to finalize the dissertation. I'm highly indebted to Jonas Deprez (CEO of Daltix) who provided data and cheered the project idea.

To my research team, especially colleague Ylenia Cavacece for giving the feedback and shaping the publications, we did together.

To Professors, Renato Salvatore and Alfonso Iodice D'Enza, for the statistical support.

Thanks to my sister, Vanja Božić, for always cheerful and heartwarming support.

In final, that's my beginning and end, I thank my family. To my beloved and supportive fiancé Pierluigi who is always by my side encouraging me throughout my Ph.D. journey, providing help when required, and daughter Nicoletta who colors every step I take. I love you the most.

ABSTRACT

The Ph.D. dissertation, *Impact of the retail environment drivers on sales and demand forecasting,* aims to explore and study the influence of the retail environment patterns on sales and demand forecasting. The uncertainty of the retail environment influenced by the complexity of the supply chain and demand management put out the range of drivers that can influence the dynamics of demand and sales and the tendency to develop more functional tools to capture them.

In the first chapter, *The concept of the environment and external impact on the forecasting in retailing*, it is outlined the main research on the environment in the literature, focusing on the concept of uncertainty, complexity especially in the domain of retailing. The network analysis on citation data benefits the research on certain trends in the literature about environmental patterns on forecasting activities. Results are outlined in the theoretical framework of research trends and the concept of the retail environment for the decision-making process in forecasting.

In the second chapter, *Influence of the internal and external drivers on consumers' visits and sales productivity*, the explorative analysis focuses on analyzing the effect of the internal retail environment, as the retail format and assortment level, on the consumers' visits and it is hypothesized whether the external (competition distance, public holidays) and internal drivers (promotions, assortment level, customer visits) affect the sales productivity in different store's formats. Results show that both promotions and competition distance positively influence sales productivity and it variates based on the store format.

The third chapter, *The causal effect of the environmental and promotional variables on the sales forecasting*, studied the type of relationships between the environmental and internal drivers on the sales forecasting, looking for the causality and functional links. Using the linear and additive models shows that, additive models can allow the together nonlinear and linear functions between the macro

variables (CPI, Fuel Price, Unemployment and Temperature) and internal micro (promotions) capturing better the dynamics of effects on sales and perform better in terms of prediction.

In the previous explorative analysis, the focus is on looking at possible causal links between the external, internal drivers and sales. In the fourth chapter, *The added value of the competition information in demand forecasting*, matching closer the business reality, it has been studied the effect of the external variables in the short-term demand forecasting. Using DIY (Do-It-Yourself) stores' demand and sales data it is analyzed the influence of the competitions' promotional discounts on the weekly demand at SKUs (Stock Keeping Unit) level. Results show that if the SKUs have high sold quantity and promotions more weeks in the year the inclusion of competitors' promotional discounts may improve the demand-forecasting model, while SKUs that have the stable demand and are not often discounted the simple linear model without external competition's variables works better.

The final chapter, *Conclusions, limitations and future research*, discuss the main conclusions, academic, practical implications and future research. The research contributes both to the literature and retail practice. There is, still, the lack of studies of the external influence on the demand and sales forecasting. Retail managers may use these insights to extract important information from the environment and to apply different solutions based on the complexity of the business problem. Focal promotions and competitive promotional actions create a significant impact on consumer demand. The complex nonlinear model with the external variables may work better in a situation of enlarging retail business and where its high impact on the market creates dynamics interactions, while the simple linear models can provide efficient solutions in an already stable and easily predicted competitive market.

The limitation of this research is the missing variety of external and internal data that can help to find the optimal model. The future research will be in direction of creating the optimal model with external and internal variables that may be the efficient solutions to the large amount of SKUs and useful managerial tool for scanning the environment**.**

Keywords: environment, scanning, uncertainty, external drivers, demand forecasting, sales, assortment, retail format, competition, promotion, classification, nonlinear, weekly forecasting.

TABLE OF CONTENTS

List of tables

List of figures

INTRODUCTION

RESEARCH MOTIVATION

The development of this project has been motivated by the fact that reality in the retail business is becoming a complex process and that retailers cannot survive without interacting and adapting with the environment. Forecasting activities inside the demand management process are crucial for the efficiency of the supply chain system. The practitioners are usually focused on the good performance of the model and the fewer implementation costs, while academics are looking for advanced models considering several aspects in which it will be optimized the performance, profit, and long-term usage of the model. This research has the same idea behind it, to give the outlook of approaches how the impact of the environment on the forecasting and sales activities can be observed and quantified in the manner that will deliver optimal performance in terms of the performance, application, and utility. The main research question (*MRQ*) of what is and how can be measured the retail environment effect on the sales and demand forecasting is supported by the following sub-research questions (Figure 1):

- *RQ1*: What are the research trends on the drivers from the environment that affect forecasting activities in retailing?
- *RQ2*: How do the assortment level, retail format, internal, external factors influence consumers' visits and sales productivity?
- *RQ3*: What is the type of relationships between internal and external environmental drivers and sales?
- *RQ4*: Will the added competition's price and promotion benefit the demand forecasting model?

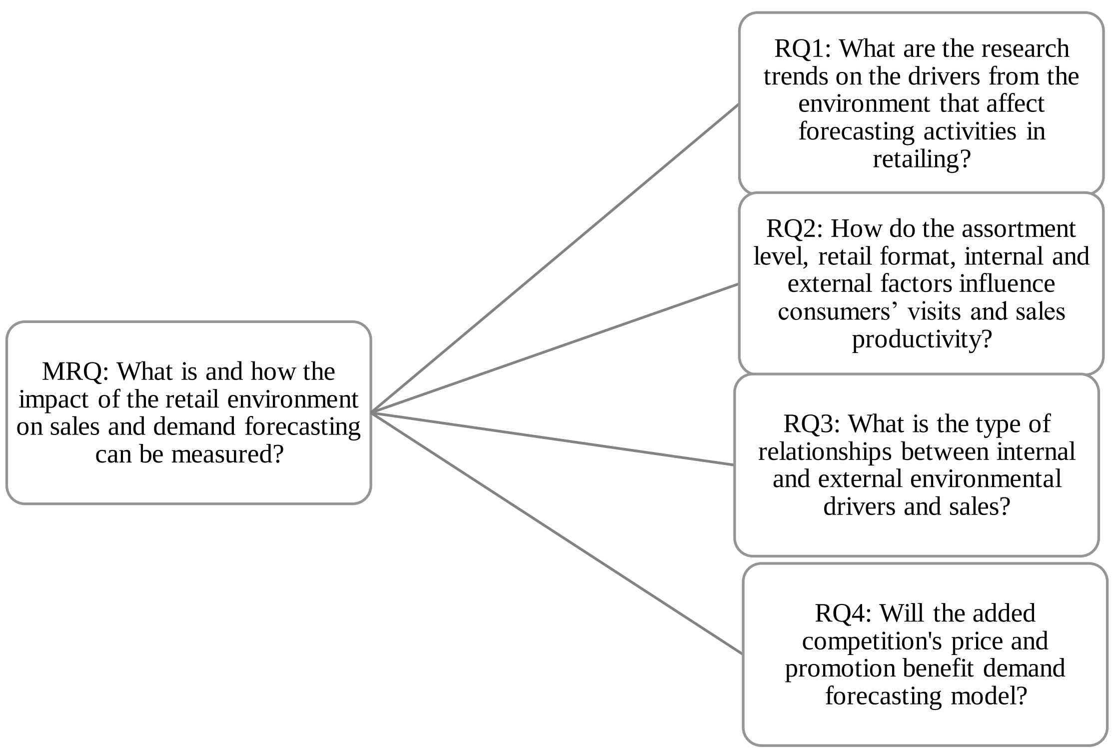
In exploring answers, the focus is to reach and identify the patterns from the environment that can be relevant to be quantified in the model. Starting from analyzing the literature which analyzing the concept of the environment and the importance of uncertainty in the retail business to the recent trends in the literature on what drivers from the environments are the matter of interests for the most of authors (RQ1). Then investigating thru data analysis how the internal factors as the type of the retail format, range of the assortment, and external factors as competition distance can affect the sales productivity in the store (RQ2); developing the model in which the aim is to explore the dynamics of the relationships between sales and internal and external environmental drivers in a retail company (RQ3) and testing on the real retail example will the added competition price and promotion benefit demand forecasting model (RQ4).

To all research questions has been dedicated the chapter that represents the independent work with literature review, conceptual framework, and methodology.

Specifically, the Ph.D. thesis is organized as follows:

Chapter I is focused on literature review and outlining the theoretical framework on the drivers from the environment, Chapter II is an explorative analysis using the empirical data to investigate hypothesis; Chapter III is the methodological research investigating the causality and type of the connection between the external drivers and sales and Chapter IV is the testing on the real retail data if the added competitor's price and promotion will benefit demand forecasting. Chapter V provides a summary of conclusions, implications, limitations of the studies, and possibilities of future research.

Figure 1. Main and sub-research questions



Source: Author's elaboration

RESEARCH OBJECTIVES AND DESCRIPTION OF STUDIES

The primary objective of the thesis is to investigate if the information about drivers from the retail environment can be in causal relationships with demand and sales, and if, this information can be quantified and used in the forecasting activities.

For that reason, in Chapter I it is explored the previous literature about the environment reflecting on the scanning, uncertainty, and complexity to highlight what studies investigate environmental drivers and how they affect the forecasting activities in the retail business. It is given the theoretical framework of the external drivers where reflecting the concept of complexity and uncertainty of the environment in retailing the decision-maker need to scan and filter right patterns and define the right relationships between environmental factors and the demand in the forecasting model.

The second objective, within Chapter II, is to analyze relationships between internal drivers as the retail format, assortment level, customers' visits, and sales productivity. With explorative analysis on data from German drug retail chain, it is possible to reveal which type of retail format the consumer often visits, what the assortment level is usually purchased, and what the internal and external factors influence sales per square meter in the diverse retail formats.

The third objective is to explore the dynamics of the relationships between sales and internal and external environmental drivers in a retail company. In Chapter III, on the example of US retail chain, using the linear and additive method (General additive Models) it is explored causal relationships between the drivers from the macro environment (CPI, Fuel Price, Unemployment), weather (Temperature), five types of the promotional drivers (Markdowns) and sales (Sales). The purpose is to support the corporate level of retail forecasting management with the advantages of the nonlinear methods.

The fourth objective, within Chapter IV, is testing the real retail data to examine whether the incorporated information of competitors' price and promotion can add value to the forecasting model and what is the size and direction of the effect of competitive price and the promotion on demand of the focal product.

Using weekly SKU level data of DIY retail chain together with external competition data, applying the time series model it is tested causality of main competitor's price and promotions on the focal demand of DIY at aim to show, as it can be the benefit in demand forecasting for the retail practitioners. The purpose is to support the operational level of the retail forecasting management.

The final objective of the Ph.D. dissertation is, how using the different statistical approaches to the retail data in presented studies, data can become a valuable asset for improving the decision-making strategies. Technical support is not only a purpose but also the strategical approach in evaluating the necessary information from the asset which retailer has already had, and that is data.

RESEARCH CONTRIBUTION

The Ph.D. thesis contributes to the academic research related to the forecasting methodology and decision-making strategies in the retail business:

1. The study highlights the inherent limits to forecasting, showing that evolution of forecasting methods is necessary and the environmental aspect should be seen as opening up new paths to reveal important insights to assist decision-making;
2. Proposing the nonlinear methods in forecasting methodology where there is a lack of research and validity to the forecasting performance and utility in the demand management;
3. Including the disaggregate effect of the competitor's price and promotion on the demand on the particular retail product. The previous research uses the aggregate effect of external information in terms of group indexes or relative percentages.

The research also has implications for the retail practitioners:

1. The results suggest managers developing the decision-making schemes in which they can scan and observe the patterns they can include in the planning activities;
2. The proposition of creating the strategies for the portfolio of the diverse retail formats and managing the size of the assortments where the results show that smaller neighbor's store and basic assortment can boost the profitability of the sales.
3. Nonlinear tools and additive models can support the corporate level of the forecasting management suggesting to consider the external variables and measure the causality and variation with the sales. The model presented contributes to the interpretation of relationships between internal and external variables in their dynamics in the retail chain environment.
4. On the operation level, for leading products in one category that is highly monitored by the competitors, it shows better off to follow the competitor's price and promotions and its disaggregate effect each SKU in the short term forecasting.

Chapter I

The concept of the environment and external impact on the forecasting in retailing

1.1 Introduction

The fact that retail business operates in a dynamic environment is no more threat and the literature discuss this topic for a long time; the increasing external complexity increases the necessity to approach the environmental change creatively. As the complexity increases the decision-making becomes complex pushing the organizations to be more agile, flexible, and capable. The ability to survive in an uncertain environment turns out the new mechanisms in organizations to adapt and prolong the life cycle in the market. In the supply chain where the trend is, lately, sharing the information among actors inside the supply chain allows the function of the forecasting to be simplified, centralizing the decision-making process. In a centralized model, sales forecasts are made by retailers to control the inventory stock and demand planning by the suppliers that make the forecasting process flexible (Ekinci and Baykasoğlu, 2016; 2019). Recent, modern technology is boosting the efficiency of the retail supply chain using machine learning tools and cognitive analysis, making the more relevant and updated demand planning. Technology that is developed nowadays creates an environment in which the retailer and supplier depend on each other and the information sharing among them becomes essential for the planning process (Chase, 2016). Using the smart shelf technology including the scanner and point of sales helping the supply operation planning system to integrate these data and investigate the true demand signals.

The evidence of the market environment suggests that successful companies apply their skills and capability in identifying opportunities in advance of their competitors and translating them into profitable opportunities early enough to develop a strong presence in front of the customer. External complexity depends on variables that are not under the control of managers and the complexity continues to rise due to the continued volatility of the factors that influence the advancement of the technology, weather conditions, political situation, demographic structure, the change in the shopping behavior. This creates dynamism in markets that are not simple to interpret, analyze or predict, making forecasting activities and marketing strategies difficult, especially in the retail sector where complexity affects decision-making processes from store organization to supplier and customer relationships (Moretta Tartaglione et al., 2018).

Moretta Tartaglione et al. (2018) discusses also that the lack of traditional forecasting plans is due to the inability to predict the changes in the environment. The environment has been seen as the system where the information is constantly circulating and updating. The traditional view of the perspective of the environment is that the 1) environment is stable (Volberda 1997; Chakravarthy 1997), 2) the organization can control the environment (Cravens 1991, White 1998). On the other hand, there is extant literature about the unstable and uncertain environment that is difficult to predict (Duncan, 1972; Dess and Beard, 1984; Daft et al., 1988; Sharfman and Dean, 1991; Harrington, 2001; Hwang, 2005; Rueda - Manzanares, et al., 2008; Oreja - Rodríguez and Yanes - Estévez, 2010; Ashill and Jobber, 2014).

In the retail industry, changes are inevitable and the demand planning process is becoming more complex. There are particular events where the distinct patterns can be easily predicted with high accuracy as the trend and the seasonal patterns, the less predictable events can be forecast using the causal factors such as price, sales promotions, and macroeconomic variables as disposable income, consumer price index, weather and other factors (Chase, 2016). The prediction of demand in the retail supply chain depends on the amount of data available and if demand for a particular product is random or no. The trend in research and practice to develop sensitive tools for catching the demand patterns when small of data is available and to detect the effects of the multiple factors. It is challenging to predict demand in highly competitive environments as competitors affect the demand patterns, which compromises the forecaster's ability to sense and predict demand due to the dynamic nature of the marketplace. Volatility and unpredictable events drive store managers to demand more flexible performance-reporting procedures (Harrauer and Schnedlitz, 2016) and to create an operational strategy that can provide demand-information in the last possible moment and to quickly respond to dynamic and volatile marketplace (Christopher et al., 2004). To boost store performance and to reduce information complexity caused by information and task overload, store managers are focusing on the optimization of processes and efficient practices and facilitating tools (Harrauer and Schnedlitz, 2016). Taking into account these statements the aim of this analysis to get more insight into the literature about the concept of the retail environment and outline what are the major elements from the environment which influence the forecasting activities. In the first part, there is the extant literature about the concept of the environment and how scanning the environment is necessary to detect the uncertain patterns and in the second part using the bibliometric analysis, textual mining of the keywords, collect, summarizes and synthesize different theoretical aspects that determine the research question:

RQ: What are the research paths on the drivers from the environment that affect forecasting activities in retailing?

The organization of the work is divided into two parts, at first the literature review of the concept of environment, its characteristics, complexity, and uncertainty, outlining the relevant research in which authors studies the influence of the environment on the retail activities and at the second part, using the social network analysis (citation analysis) on the keywords identifies the main aspects on the topic; the discussion of the results, the systematic framework on the trends in research on the major envirnmental elements affecting the demand (sales) forecasting and related conclusions.

Literature review
(1st part)

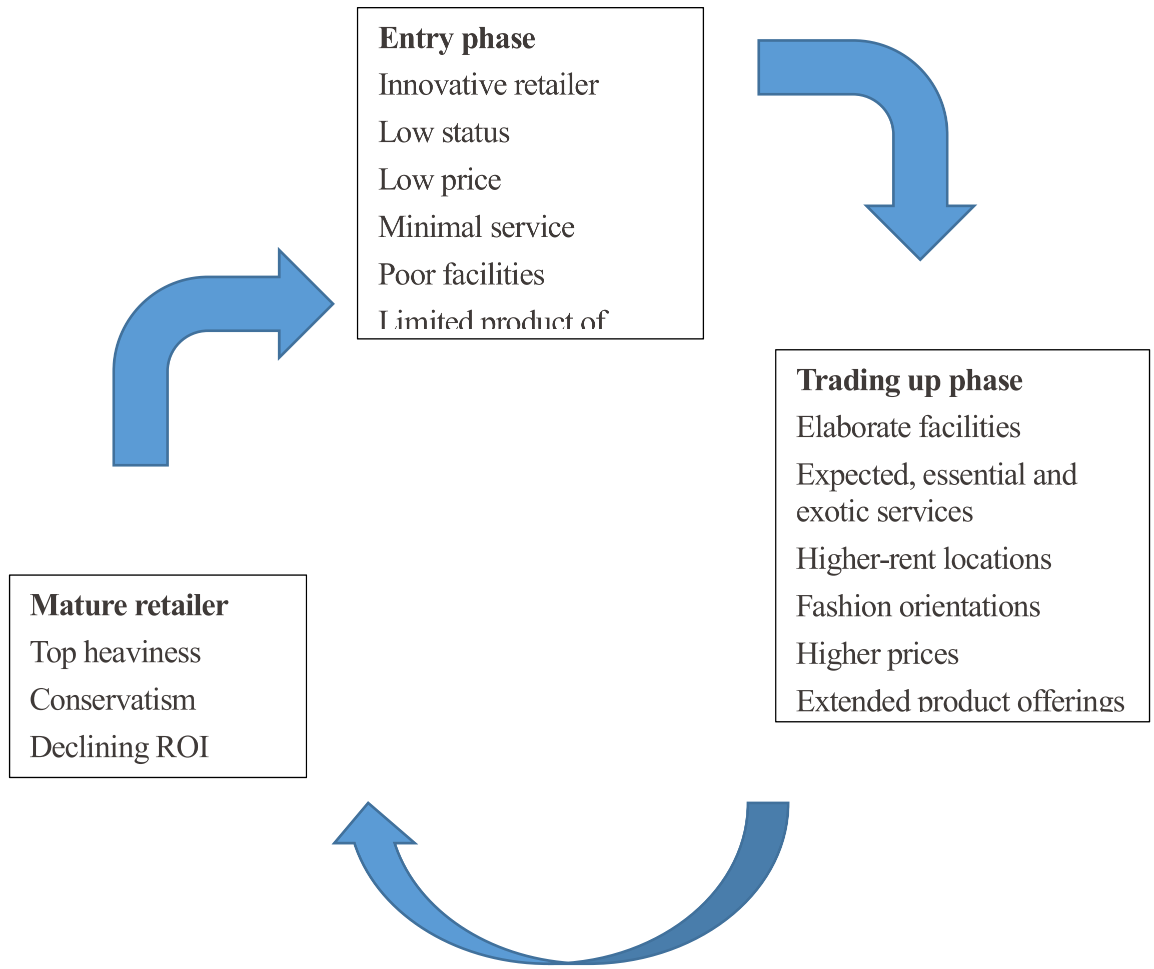
1.2 The concept of the environment in retail industry

The retail environment is the constructs of internal store surroundings and external influence from the entities that square measure in direct or indirect respect to the shop. Definition of the retail environment becomes complicated because of the overlapping effects and these days we tend to cannot speak solely concerning retail surroundings while not mentioning the external influence of e.g. inflation on sales, exchange rates, surplus or deficit, Gross Domestic Product, supplier's politics, mediation, etc. each of those components constructs specific surroundings that run across and exchange with alternative surroundings. In a recently revealed analysis (Moretta Tartaglione et al., 2019), there are mentioned completely different analysis contributions to management literature that concern the retail environment. For Waterhouse and Tiessen (1978), the environment is a longtime construct in contingency theory literature and combines internal (inner corporate) and external (outer corporate) components. Child (1975) describes environmental constructs as changes and market dynamics that square measure troublesome to predict and generate uncertainty (turbulence). Generally, organizations operate among two completely different environments (Vasconcelos and Ramirez, 2011) with the discourse surroundings; and therefore the transactional environment. The discourse surroundings consider the angle of the organization and consists of things poignant the organization, over that the organization has no power or influence. The organization will assess, appraise and perceive its discourse surroundings (Smith, 1983). Economic conditions, like inflation, economic process, exchange rates and interest rates, and socio-political problems, like mediation, law, and sociology, square measure factors that type the discourse surroundings. The transactional environment consists of the organizations and people with that the organization directly interacts (Vasconcelos and Ramirez, 2011). Winsor (1995) outlines that settled actions of the corporate are interceded by the random influences of the surroundings. Besides several authors discussing the rise of unpredictable environmental factors that ends up in larger impact upon decision-making and outcomes of a company (Ackoff, 1981; Mitroff, 1987; Perrow, 2011).

Winsor (1995) addresses the quality of environmental factors that manufacture the random impact upon selling processes which uncertainty affects social control selections that tend to be additional risky. The relation between the action of firm and ultimate market outcome is laid low with random mediators (Winsor, 1995), in alternative words, the business surroundings become additional chaotic (Waterman, 1987), and therefore the conceptualization of the selling that's being processed in those chaotic environments is consistent of the economic discipline, wherever companies behave in rational methodology s of maximization assumptive that's surroundings is ruled by theoretical account (Herbig, 1990).

The stochastic nature of influence on retailer's activity has been considered in the latest research where information coming from the environment has been considered in strategic plans of one retailer. Previous theories assume the cyclical change in retailer development. The cyclical theory of retailing was suggested by McNair (1958) that offers the Wheel of Retailing concept of business development. In Wheel of Retailing, the retailer starts from the low-price positioning increasing the profit increases also the sophistication in their activities, service, pricing policies, selling methods, then retailers invest in bigger stores, better locations increasing the operating costs that at the end becoming inflexible business due to the high cost and not easily adaptable to the new entrants with low-price policy (Figure 1-1). Brown (1988) defines the Wheel of Retailing in three steps: Entry, Trading and Vulnerability (Figure 1-1).

Figure 1-1. Wheel of retailing



Source: The wheel of retailing (Brown 1988). Adapted from Theories of retailing. Marketing Theory: A Student Text, (2nd ed, p. 348) by Moore and Doyle, 2016, Sage. Copyright 2010 by Baker and Saren.

Hollander (1960) criticized the Wheel of Retailing mentioning that in the moment of increase in operating costs the external influence of discount houses and supermarket competition is unpredictable and hard to measure. This difficulty comes from the scarce of data from a retailer, which monitoring the aggregate balance is not enough to follow the tendencies of change. The cyclical theories of change besides their wide aspects are inadequate due to the linearity in describing the retailer life's cycle and reflecting the deterministic, inflexible, and prescriptive way of retailer's development (Brown, 1988). Brown (1991) outlined that the presented model fails to capture the influence of the economic environment or for the strategic plans and management interventions. In such a moment, researchers initiate to introduce cyclical, so-called environmental theories that offer more non-flexible and efficient solutions to explain the patterns that affect changes within the retail sector. Retailers can be passive or proactive to the changes on the market, not every retailer is waiting for the failure with the "crossed hands" instead traders may deploy plans that predict difficult market trends or look for the opportunity that comes from the changed market conditions. McGoldrick (2002) noted that socio-cultural developments within the market affect the changes in retail format. Meloche et al. (1988) that the crash of retail business and change of retail format is influenced by the negative environmental change or market alteration. While by cyclical theories the retailer has the start and period in which should be actively trading and surviving, environmental theories provide an alternative explanation why market demand may shift from one retail format (premium) to another (value/discount basis) in the period of economic downturns. Takin from the real market example in the period of recession 2008 there is the increase of discount supermarkets as Lidl, Primark, and Aldi. Etgar (1984) and McGoldrick (2002) have suggested that the environmental view on retail change recognizes that an 'economic ecology' exists within retailing where the principle is that only to survive. This means in difficult periods it depends much more on the retailer's capability to respond effectively to challenging markets and to manage to survive. Pioch and Schmidt(2000) propose the conflict theory in which to predict the made adaptations by both sides result in their adopt strategies that are largely similar in terms of scope and impact. Then the differentiation at the beginning from two retail formats is disappearing that initially appeared to be so different. Cyclical and environmental theories provide wide insights and valuable definitions of the retailer's life process but no one of them provides the complete framework of the dynamics of that change. Brown (1991) and Sampson and Tigert (1994) proposed the hybrid model of cyclical, environmental and conflict theory bringing together good and outlining the weakness from the models. These models, which bring together dimensions from all three-model formats serve, in particular, to highlight the complexity and diversity of change within the retailing sector. In so doing, they provide a clear and complete system for explaining the reasons for the change and actions that should be taken. The necessity of multidimensionality of new models is to reflect the reality of the retail market that is complex and turbulent and to respond to these dynamic changes in which there can be a constant switch of intensity and level of complexity, it is required to go along with strategic, active, agile and robust plans.

1.2.1 The internal and external the retail environment

The concept of environment in the retail industry affects the previous definitions and assumes a particular value inside stores because of the emerging system of elements –*, the dynamics of the supply chain network, relationships directly connected with the store, employees, stock, layout, promotions, price policies and the role of the store in the specific place* – represent a particular context that is affected by complexity. This context determines the competitive advantage of a single store and, for this reason, the competitive position of the whole retail chain.

The external and the internal environment are connected and merge in the retail industry. In particular, focusing on the literature on relationships between the environment and the retail industry, Harrauer and Schnedlitz (2016) address the recent financial-economic crisis as the explicit cause of the turbulent market; since 2008, top managers have undertaken profound restructuring activities, which have led to significant changes in in-store activities and performance measurement. After the crisis, neither long-term forecasts nor schedules were convenient to deal with changing customer needs. Companies have had to adapt their marketing and service technology to stay competitive in the market and to incorporate the complexity of the retailing business in their performance measurement designs.

At the global level corporations compete in a market space competition, a dimension in which the space is not stable in the decision-making process but the competitive factor which complexity depends on markets defined by the time-based competition (efficiency) and over-supply conditions (the overabundance of low-cost and the more sophisticated goods)(Brondoni, 2012). The substantial role in the creation of such an environment has macroeconomic variables, industry concentration (Mintz and Currim, 2013), competition, market regulations and relationships inside the supply chain (Eksoz and Mansouri, 2014; Chase, 2013; 2016), consumer welfare, employment, and the relationships between the retail industry and production (Basker, 2007; Hausman and Leibtag, 2007; Jia, 2008; Neumark et al., 2008). The environment is characterized for instance by the store location and by the relationships between retailers and policy-makers that, as a result of bureaucracy and other legal constraints, increase the competition and the difficulties in obtaining resources.

At the store level, this affects the time it takes to organize a shop (or a retail chain) and the only way to obtain advantages is by improving the ability to sense the environmental changes dynamically (Espejo, 1993). Turbulent and unpredictable events, in particular the competitor's actions of lowering the price, growing the brand, launching the new product, or strategic customer delaying purchase drive store managers to demand more flexible performance-reporting procedures (Harrauer and Schnedlitz, 2016); strategic management is the only way in which firms can get a series of short-term advantages (D'Aveni et al., 2010). Each store is a retail store network has its individual competitive environment that arises from the different configuration of competitors and the specific

conditions in the defined geographical catchment area for the store. The specific conditions of the store's environment are characterized by contractual relationships, competitors' actions, products' life cycles, seasons, consumer plans, and tastes. Besides this, digital communication flows make it possible to dialogue simultaneously with a large number of contacts in the network about products, price promotional offers, acquiring this information signals about suppliers, products, consumers, and retailers, database management becomes significant in understanding environmental phenomena that can influence the corporate to the store level. (Brondoni, 2012).

Considering these high number of interactions in the retail network, their non-linear relationships, and their variability, the environmental impact differs from case to case in the retail industry (Harrauer and Schnedlitz, 2016); however, applying analytical methods that only sense demand signals associated with trend and seasonality at the aggregate market channel level are no longer sufficient (Chase, 2013).

1.2.2 Effects of the business environment on the retail operational activities

Operational planning is the key to the success and survival of the market (Lowson,2005; Yu, 2011; Yu and Ramathan, 2011). The operational strategies are necessary for the adaptation to the external environment and how a company will employ the resources (Prasad et al., 2001; Yu, 2011; Yu and Ramathan, 2011). Some studies find the link between the environmental factors and operations strategy (e.g. Miller and Friesen, 1983; Dess and Beard, 1984; Ward et al., 1995; Amoako-Gyampah and Boye, 2001; Anand and Ward, 2004). The business environment has been characterized by the dimension of dynamism and hostility (Yu and Ramathan, 2011). Environmental hostility has been affected by intense price competition, rising business costs, low-profit margins, severe regulatory restrictions, shortages of labor and/or raw materials, and unfavorable demographic trends, which offer few opportunities to exploit (Miller and Friesen, 1983;). Environmental dynamism refers to the extent of the unpredictability of change within the company's environment (Dess and Beard, 1984, Yu and Ramathan,2011). This change can be caused by the increase of the rate of change and innovation in the company's principal industries, the insertion of the new products and service, and the uncertainty or unpredictability of competitors' actions and customers' preferences (Lawrence and Lorsch, 1967; Miller and Friesen, 1983).

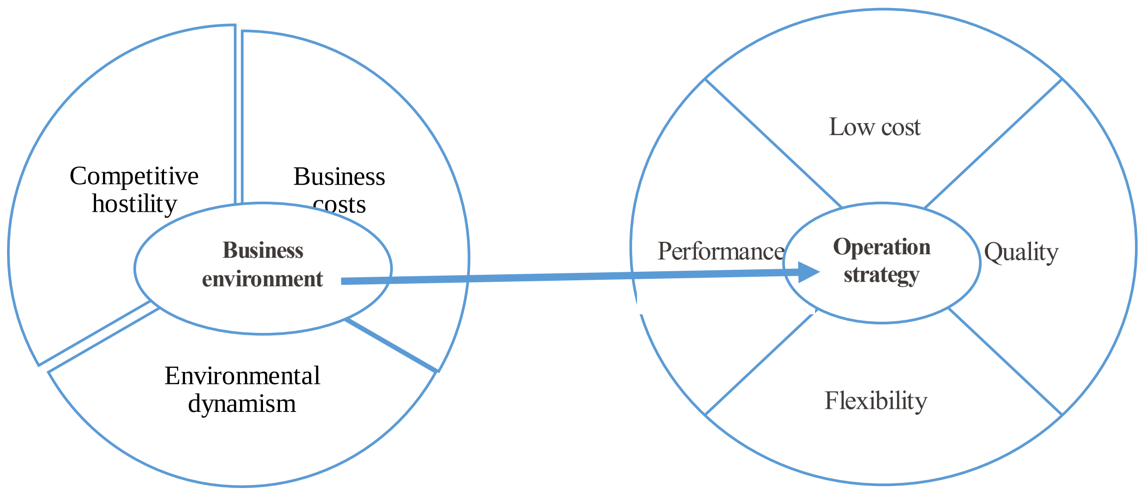
Today large retail companies are facing complexity and uncertainty when they want to operate internationally. Enlarging the business, they are dealing with the dynamic nature of the market, new business environment, unknown consumer preferences, and location choices for the retail business (Vida and Fairhurst, 1998; Dawson, 2002; Burt et al., 2003; Yu and Ramathan, 2011). In the literature, dynamism and hostility have been used to describe the business environment. Hostile environments increase the threats to the company and make complicated and complex strategic decisions (Miller and Friesen, 1983; Yu and Ramathan, 2011). Environmental hostility is defined by the intense price competition, market regulations, rising costs, deficit in raw materials and labor offer, migration trends (Miller and Friesen, 1983; Yu and Ramathan, 2011). Authors draw the attention that to sense threats from hostile business environments it is necessary to develop the analytical tools to understand and predict these challenges (Khandwalla, 1972; Yu and Ramathan, 2011). To overcome these challenges and make the clearer path to lead of the company it is necessary to scan the environment and scanning the business environment becomes the crucial capability for multinational retailers (Preble et al., 1988; Yu and Ramathan, 2011;). Oreja - Rodríguez and Yanes - Estévez (2010) noted that it is important to analyze environmental evolution to check the effectiveness of the actions and to identify the tendencies. This is the determinant for the company's performance and survival (Milikien, 1990). Oreja - Rodríguez and Yanes - Estévez (2010) developed the research whose aim is to identify the variation of the environmental variables in two years and the evolution of the dynamism perceived by the companies. To do that they used the Rasch model (Rasch, 1993) "consider the object of the study (environmental dynamism) as a latent variable in which two different entities interact: the surveyed subjects (firms) and the items of the measurement instruments (environmental variables)"[1].

Multinational companies can experience the challenging path in building the reputation in the consumer market and along this way can face challenges as high operating costs, high competitive

[1] Oreja-Rodríguez, J. R., and Yanes-Estévez, V. (2010). Environmental scanning. *Management Decision*, pp.264.

market, and consumer awareness (Lo et al., 2001; Uncles, 2010; Yu and Ramathan, 2011; Yu, 2011). It is necessary to scan, monitor, and examine the business environment before the company decides to enter the international market. Yu and Ramathan (2011) built the conceptual framework in analyzing the business environment of China retail market for the research purpose.

Figure 1-2. Effects of the business environment on the retail strategies



Source: Concepual framework of business environment and retail operation strategy". From Yu, W., and Ramanathan, R. (2012). Effects of business environment on international retail operations: Case study evidence from China. International Journal of Retail and Distribution Management, 40(3), 218–234. https://doi.org/10.1108/09590551211207175. Copyright 2012, Emerald Group Publishing Limited.

According to Yu and Ramathan (2011) business environment (Figure1-2) is defined by the three factors that companies are finding difficult to control as costs, competitive hostility, and environmental dynamism. According to the literature review of authors, it is the conclusion that operation strategy has been based on building up the capabilities that will be become the competing priorities. These competitive priorities are mainly (1) low cost; (2) quality; (3) delivery performance; and (4) flexibility. These operational strategies have been influenced by the volatility of the environment, labor availability, competition, industry environment, there are quite strong relationships between these environmental variables and operational strategies based on the competitive priorities by Ward et al. (1995).

Table 1-1. The concept of the retail environment – summary of relevant topics and authors

Topic	**Authors**
Inner corporate and external (outer corporate)	Waterhouse and Tiessen ,1978;
Market dynamics	Child,1975; Brown,1991; Sampson and Tigert, 1994; Winsor,1995;
Cyclical theories	McNair,1958; Brown,1991; Sampson and Tigert,1994;
Environmental theories	Brown, 1991; Sampson and Tigert, 1994; McGoldrick, 2002;
Retail store network	Harrauer and Schnedlitz, 2016;
Uncertainty	Lawrence and Lorsch, 1967; Miller and Friesen, 1983; Rueda-Manzares et al, 2008;
Transactional environment/discourse surroundings	Smith, 1983; Winsor, 1995; Vasconcelos and Ramirez, 2011;
Wheel of retailing	McNair, 1958; Hollander, 1960; Brown 1988;
Performance measurement	Harrauer and Schnedlitz, 2016;
Competition	Brondoni, 2012; Chase Jr, C.W., 2013; Eksoz and Mansouri, 2014; Harrauer and Schnedlitz, 2016;
Business environment	Preble et al., 1988; Vida and Fairhurst, 1998; Dawson, 2002; Burt et al., 2003; Yu and Ramathan 2011; Rueda-Manzanares et al., 2008;
Environmental dynamism	Rasch, 1993; Dess and Beard, 1984; Oreja-Rodriguez and Yanes-Estévez, 2010; Yu and Ramathan, 2011;
Operational strategies	Ward et al, 1995; Lowson, 2005; Yu, 2011; Yu and Ramathan, 2011;
Environmental hostility	Miller and Friesen, 1983; Hwang, 2005; Yu and Ramathan, 2011;

Source: Author’s elaboration

1.3 Environmental scanning, complexity and dynamics

1.3.1 Environmental scanning

The scanning of the environment is necessary for building the strategy. How the environment will be monitor, depends on how a company perceives the environment. In the literature, there was a debate about consideration of the environment in business planning. The first perspective of the environment is that the environment is independent of the individual or other environment and it is a result of perception so-called perceived environment (Simon, 1957), it is an intention to the perceived reality. The perception of the environment translates into information that is the source for the strategic planning and the condition to a greater extent (Sutcliffe and Huber, 1998; Stewart et al., 2008;). On the other hand, the second perspective is which level of environmental scanning should be considered. Daft et al. (1988) and Duncan (1972) refer to uncertainty as the principal dimension to diagnose the environment. There was not an agreement between authors on which environmental factors should be included but they all agree that environmental dynamism is the major cause of the uncertainty and easy to quantify (Child, 1972). As mention previously Duncan (1972) is more concerned about the degree of the stability to which external factors remain stable or change over time. Many authors identified environmental dynamism as difficult to predict the change (Dess and Beard, 1984; Sharfman and Dean, 1991; Harrington, 2001). The definition of environmental scanning, besides monitoring the external events, includes also the perception of the future path which the company will follow (Aguilar, 1967; Hambrick, 1982, Milliken, 1990). Even the environment scanning has been studied from a years ago there is not still enough quantified studies on how the uncertainty and dynamism can be predicted. Hwang (2005) refers to the difficulty to obtain clear information from external factors. Still, some research papers propose a longitudinal study to catch the volatility of the environment (Lenz and Engledow,1986; Barry, 2006), or to study the causality of the evolution of the environment (Emery and Trist,1965; Ansoff, 1981). These authors developed the theory of the five types of the environment and the fact that the tendency of the environment leads to an increase of the complexity and turbulence (Oreja - Rodríguez and Yanes - Estévez, 2010).

1.3.2 Environmental dynamics

The concept of dynamism interferes with many domains in the business literature as organization theory, strategic management (Hwang, 2005). It reflects the instability and turbulence of the market in terms of the economical, technological and political conditions (Emery and Trist, 1965; Aldrich, 1979; Mintzberg, 1979; Dess and Beard, 1984; Sharfman and Dean, 1991). These studies refer that the best measure of the instability is the unpredictability and the absence of the pattern (Dess and Beard, 1984), authors suggest also that dynamism should be defined as volatility hard to predict and create uncertainty for the key managers. Uncertainty can affect the organizational structure (Hwang, 2005) and more information can increase the task uncertainty. The more uncertainty increases the decision-makers need more information to achieve a certain level of performance for a company (Galbraith, 1973). The complexity of the environment and dynamism is in tight relation to the uncertainty (Lawrence and Lorsch, 1967; Thompson, 1967), whereas hostility has been tied to the resource-dependence perspective (Pfeffer and Salancik, 1978; Aldrich; 1979; Hwang, 2005). These perspective helps to understand the impact from various levels of the environment. The dynamics of changes from the environment can affect the dynamics in decision-making affecting the strategic decisions and influencing the propensity for risk-taking, proactiveness, and defensiveness (Miles and Snow, 1978; Miller and Friesen, 1983).

In the literature about the impact of environment on the organizational structure, Hickson et al. (1971) refer to the uncertainty as to the "lack of information about events, then alternatives and results are unpredictable". Many authors connect the dynamics and environmental complexity with organizational performance. The relation between uncertainty and the variables as the organizational structure (Koberg and Ungson, 1987; Kim and Burton, 2002; Hwang, 2005), a firm's strategy (Hitt et al., 1982; Gerloff et al., 1991; Choe, 2003), and economic performance (McCabe, 1990; Kim and Burton, 2002; Sawyerr et al., 2003). These studies have the same conclusion: that environmental uncertainty becomes the central problem of organizations. Forward, literature studies about the perceived environmental uncertainty.

1.3.3 *Environmental uncertainty and complexity*

Environmental uncertainty is the absence of information about future events, actions, and activities in the environment (Daft et al., 1988). Milliken (1987) defines uncertainty as to the result of the process of decision makers' understanding through which they designate meaning and interpret situations (Hwang, 2005). Uncertainty is one of the central issues of organizations' adaptations to their environments to remain competitive (Thompson, 1967).

The uncertainty produced by the environment is usually the factor that affects the top management's willingness to adapt to the changes, to update the market-orientation strategy, organization structure, and store's performance (Hwang, 2005). Uncertainty in the environment is caused by elements that interfere with each other as the new technology, new competitors, and increase in the customer's expectations. Customers have higher expectations about products due to the availability of information, fast-moving technology and overloaded choices. New technology in one industry is affecting another industry by exploiting the opportunity in the new product (Abell, 1978). New players in the competitive market can cause unpredictable changes and disrupting the status quo in the domestic industries (Hwang, 2005). These changes have caused increasing uncertainty in the marketplace (Utterback and Uhlmann, 1979; Murtha et al., 1998; Grewal and Tansuhaj, 2001; Hwang, 2005). Uncertainty charchtarerizes three elements: 1) the lack of the clarity of information, 2) the long span of feedback and 3) general uncertainty of causal relationships (Lawrence and Lorsch, 1967).

Uncertainty influences the management's decision-making process and authors tried to filter the influence in the logical categories. Miller (1993) defines uncertainty as to the lack of ability to predict the environmental variables that affect the corporate performance, on the other hand, Milliken (1987) argues that certainty is an inability to discriminate the relevant and irrelevant data. The environmental uncertainty by Miller (1993) can divide into three categories (1) general environment, (2) industry, (3) company-specific variables. The general environment includes the macroeconomic variable as socio-economic status, demographic structure, government regulations and technological standard while in the industry the environmental factors are related to the market regulations, competitors, products market, and the level of implemented technology. Uncertainty on the company level is linked to the management decisions, operations and research and development. Milliken (1987) categorizes uncertainty into three types: effect, state and response. Effect uncertainty is the lack of capacity to predict the type of the effect of a future state of the environment on the organization, and response uncertainty is the lack of ability to predict the likely consequences of a response choice. State uncertainty refers to perceived environmental uncertainty (Buchko, 1994). When managers are unable to predict and understand the changes of the elements of the environment and causal relationships between the elements at that moment the perception of the environmental uncertainty occurs. Duncan (1972) defines uncertainty as to the lack of information regarding the environmental factors related to the decision-making process, unpredictable outcome of the specific decisions and not being able to predict how probable environmental factors will affect the success or failure of the decision unit in performing the decision. Complexity and rate of change are defining uncertainty (Tung, 1979; Child,1972; Duncan, 1972; Hwang, 2005). Complexity depends on how variable environmental events are. If environmental events occur in large numbers and they are diverse then we can that environmental complexity is high. The influence of the external environment on distribution channels has been studied channels (Stern and Reve, 1980; Arndt, 1983; Achrol et al.,1983; Dwyer and Welsh, 1985; Achrol and Stern, 1988). According to those authors, environmental uncertainty produces decision-making uncertainty. In this context, the authors outline the perceived uncertainty because managers are the ones who can state if the event is certain or uncertain so their perception about decisions as well may become certain or uncertain (Dill, 1958; Emery and Trist, 1965; Duncan,1972). It has been a while in the literature timeline that authors recognize the sensitivity of retailers to the environment. Retailers have to be able to recognize the threats and opportunities that are coming from the environment and to capitalize on the opportunity (Dibb, 1996). Dibb (1996) recognizes the two types of environment: the simple-complex and static-dynamic. The difference between these two environments is that number of the factors in decision-making have been affected by the degree of the change: they change over time (static-dynamic) or remain the same (simple-complex). The uncertainty that retailers face from the environment can affect their decisions to recognize the threats and opportunities Turbulence makes the environment complex and dynamic and requires sensitive skills from the manager to capture the environmental factors. The retailer needs to pay attention to the products because they are the ones most influenced by the environment (Keegan et al., 1989). The most important factors identified by the researchers in the literature (Dibb and Simkin, 1994) are the market turbulence and competition intensity. The authors state that the turbulence and the competition intensity are perceived environmental uncertainty. If retailers do not perceive the potential danger from the market, they are likely to lose the market share becoming unprofitable (Cooper, 1979; Hayes and Abernathy, 1980; Covin and Slevin, 1989; Hwang, 2005). The managers often face the challenges coming from the complex and dynamic environment through the traditional is not any more suitable for the complex and numerous environmental factors that become less predictable. Where the uncertainty is high, environmental factors become more unpredictable. They are in constant flux (Hwang, 2005; Smart and Vertinsky, 1984) and the value of the important variables are changing unpredictably. According to Drucker (1993) and Huber (1984), turbulence is the increase of the number of events that occur within a given period. Several events create many factors that can at the same time affect one organization then the information manager receives becomes complex. Wang and Chan (1995) defined complexity as "the number and diversity of external factors facing the firm"; and dynamism as "the degree of change exhibited in those factors" (p. 34). Ansoff and McDonnell (1990) state that complexity refers to the variety of factors that management must consider when making decisions. In management literature, authors often distinguish between internal complexity, which depends on the size of the company, the number of its components, the variety of social roles and personalities; and external complexity, which depends on the variety and variability of the environment in which a company operates (Moretta Tartaglione et al., 2018). External complexity is the hardest to manage because it depends on variables that are not under the control of managers. Moretta Tartaglione et al. (2018) states that external complexity continues to rise due to the continuous changes in consumption and communication, climate change, economic and political instability, transnational competition in the global marketplace, among many other factors. This creates dynamism in markets that are not simple to interpret, analyze or predict, making forecasting activities and marketing strategies difficult, particularly in the retail sector where complexity affects decision-making processes from store organization to supplier and customer relationships.

Many authors discuss the information processing from the environment to the management. What is the mutual agreement among all authors that the more complex and dynamic environment is the more complicated is to process information (Simon 1973; Ansoff and McDonnell, 1990; Douglas and McCauley, 1999; Wang and Chan, 1995; Hwang, 2005;). An organization needs to be able to respond rapidly to changing market regulations, switch of the consumer behavior, governance policies and other external influence as demographic structure, weather impact, etc. A timely informed manager is want to be a retailer manager, able to filter in advance the information and predict the future possible outcomes. The turbulence of the market can come from the rapid change of the structure of the customers and their preferences (Kohli and Jaworski, 1990; Egeren and O'Connor, 1998) while Miller (1987) refers to the heterogeneity in production methods and marketing strategies to satisfy the customer needs.

1.3.4 Environmental uncertainty and competition intensity

The retail market has been affected by the intensity of the competition (Yasai Ardekani and Haug, 1997; Porter, 1999;2008; Hwang, 2005). The increase of uncertainty is influenced by the strategy, actions, and relationships among competitors. The competition takes many forms (Porter, 1999; 2008) as price politics, promotional discounts, new product introduction, advertising strategies, and service innovations. Competition can drive the profit in one industry and the dynamics of the profitability depend on the degree of the competitive intensity and on the basis on which competitors compete (Porter, 1999; 2008). Porter (2008) outlined that Eisenhardt (1989) characterized highly competitive environments as those with intense price and non-price competition. Such intense competition is linked to rapid and sometimes discontinuous changes in the market and competitive and technological conditions (Hwang, 2005). Competitors' actions and reactions may be highly unpredictable, and the speed of adjustment to market and technological conditions becomes the key to the survival of participants in such environments (Eisenhardt, 1989).

According to Porter (2008) on which basis the companies are competing, there is a change of the degree in intensity. The intensity of the competitor is greater if the companies are numerous and equal in size and power with the slow growth of the particular industry and high exit barriers. The intensity of competition can be at high risk when is only focused on price politics. It can easily switch the profit from the industry to the customers because for the companies are sometimes more easy to do the price cut than outline the quality of the products. In the retail industry, there are two most common types of competition: intra-type and inter-type competition. (Berry, 1995; Mishra and Raghunathan, 2004; Miller et al., 1999; Hwang, 2005;). Intra-type competition is defined as the competition between the same type of retailers i.e. two supermarkets (Mason and Mayer, 1987) while inter-type competition is the competition between the different formats of retailers selling the same merchandise (Mason and Mayer, 1987).

Uncertainty and the challenge depending on the intensity of the competition and is the risk from the external environment that has to be followed continuously. When the intensity of inter-competition is high, there will be many possible choices for the consumers, in that way the features of the products will be devalued and the power will go into the hand of consumers. Companies have to be flexible in responding to these signals and able to respond fast to consumer demands (Douglas and McCauley, 1999). Companies have to be able to monitor signal from the market and the change in consumer's preferences and behavior as to adapt the products or service that will be superior to the competitors (Egeren and O'Connor, 1998), sometimes it has to be aggressive action (Kohli and Jaworski, 1990).

The meaning of market turbulence is described as market heterogeneity (Miller, 1987), where heterogeneity presents the diverse marketing tactics and production methods to satisfy diverse customers' needs. Changing in customer preferences implies market uncertainty and turbulence. Greater is the diversification of the customers' preference the greater is the need for value creation strategies to satisfy those needs. It becomes important the ability and flexibility of companies in the market heterogeneity and constantly changing business environment. It is necessary to analyze and monitor customers' preferences, predict the demand, and monitor competition's actions. This is how a company can stay competitive and reach the business goals, making the variable portfolio of actions that will answer to the market heterogeneity.

Table 1-2. Environmental scanning, complexity and dynamics- relevant topics and authors

Topics	**Authors**
Environment perception	Simon, 1957; Aguilar, 1967; Hambrick, 1982; Milliken, 1990; Sutcliffe and Huber, 1998; Stewart et al., 2008;
Uncertainty	Lawrence and Lorsch, 1967; Thompson, 1967; Hickson at al., 1971; Duncan,1972; Milliken,1987; Daft et al., 1988; Miller,1993; Bunchko, 1994; Hwang, 2005;
Dynamism	Child, 1972; Miles and Snow, 1978; Miller and Friesen, 1983; Wang and Chan,1995; Hwang, 2005;
Instability	Emery and Trist, 1965; Aldrich, 1979; Mintzberg, 1979; Dess and Beard, 1984; Sharfman and Dean, 1991;
Complexity;	Child,1972; Duncan, 1972; Tung, 1979; Ansoff and McDonnell, 1990; Wang and Chan,1995; Hwang,

	2005; Moretta et al., 2018;
Decision-making	Ansoff and McDonnell, 1990; Dibb, 1996;
Retailers	Keegan, 1989; Dibb, 1996;
Market turbulence	Cooper, 1979; Hayes and Abernathy, 1980; Huber, 1984; Waterman, 1987; Covin and Slevin, 1989; Drucker,1993.; Simkin et al., 1994; Hwang, 2005;
Information processing	Simon, 1973; Ansoff and McDonell, 1990; Wang and Chan, 1995; Douglas and McCauley, 1999; Hwang, 2005;
Competition intensity	Eisenhardt, 1989; Yasai Ardekani and Hang, 1997; Porter, 1999;2008; Hwang, 2005;
Intra-type and inter-type competition	Mason and Mayer, 1987; Berry, 1995; Miller et al., 1999; Mishra and Raghunathan, 2004; Hwang, 2005;
Consumers behavior	Miller, 1987; Kohli and Jaworski, 1990; Egeren and O'Connor, 1998; Douglas and McCauley, 1999;
Market heterogeneity	Miller, 1987;
Flexibility of companies	Miller, 1987; Douglas and McCauley, 1999;

Source: Author's elaboration

1.4 Complexity of environment

1.4.1 Complexity

Complexity is the term in the literature that receives a lot of attention lately. The first introduction of complexity was in 1949 by Shannon and Weaver who interpreted the complexity as the "entropy" the concept links to cybernetics and used it to calculate the rate of transmitted information. The study shows that complexity enlarges with the increase of the length of information. The information is irreducible same as complex systems because no one cannot grasp information until it is completely received (Shannon and Weaver, 1949). In literature about complexity, there are two types of complexity (Ekinci and Baykasoğlu, 2019):

1. Static complexity: it is time-independent, it defines the structure of the system exploring the strength of interactions between components;
2. Dynamic complexity: investigate the behavior of one system in the measure of the time and

Authors are interested in measuring the dynamic complexity of one system. Frizelle and Woodcock (1995) used the complexity (entropy) metrics to analyze the dynamic complexity in manufacturing the environment. The results support decision-makers who will be more able to make better decisions by simplifying the operations. Sivadasan et al. (2002; 2006) analyzed via measurement of complexity how much the supply chain is at the risk from the uncertainty and under which degree of the monitoring and controlling of the complexity. In recent literature about the static complexity Battini et al. (2007) modified the entropy metrics in average mutual information (AMI) to evaluate the connectivity in the supply chain network. The connectivity in the supply network chain starts from the raw materials suppliers to the retailers. The complexity arises when the number of connections increases and calculation has performed based on the flow between the entities. Chen et al. (2013) modified the calculation of complexity to include the connection diversity and the numbers of echelons while Rensburg (2012) used the complexity metrics for the estimation of the complexity of the business units integrated with a fuzzy cognitive approach to have better understand the correlations between them. This concept helps to uncover the hidden relationships of a complex system.

1.4.2 Environment as Complex Adaptive Systems (CAS)

Complex Adaptive Systems (CAS) are systems built up from single parts or agents that interact with one another following the set of rules. Interacting between agents the system grows and learns while the behavior of the agents influences the behavior of the system completely (Wollmann and Steiner 2017). CAS learns and evolves based on the experience, becoming adaptive as the main approach for its survival, processing information with the constructs schema.

Many authors observe the environment as Complex Adaptive Systems (CASs) (Peters, 1999; Berthon, 2000; Choi et al., 2001; Begun et al., 2003; Mason, 2007; Moretta Tartaglione et al., 2018; 2019), the characteristics of CASs, individuals, behavior, self-organizing elements can be useful for the better concerning the dynamics of the market and complexity of the external environment (Bruni et al., 2016). CASs represent the networks of relationships and interactions where the entire is valid more than the sum of the parts and where the single elements react and behave changing the environment as the whole (Moretta Tartaglione and Bruni, 2016). In the retail context the supply chains have been seen as CASs (Choi et al,, 2001; Surana et al., 2005; Pathak et al., 2007; He et al., 2013; Nair et al., 2016; Nair and Reed‐Tsochas, 2019; Zhao et al., 2019;). Authors discuss the supply chain as the complex networks where is a high level of exchanging information and difficulty to control and predict the factors that influence the profitability of the supply chain. Inside the supply-chain, the flow of the materials, products, services, information, the relationships are not any more one-to-one business relationships base on the linearity in the evolution of the relationships. The supply chain has been seen as the multiple web network of the multiple emerged interactions of many suppliers and relations, in that case, the evolution of the relationships is going in a nonlinear way. Surana et al. (2005) discuss the supply chain network as the complex adaptive system:

i. Structure extending several levels; the supply chain network has multiple levels consist of suppliers, retailers, distributors, and physical entities;
ii. The heterogeneity; the variation of technologies that can offer how the demand is being satisfied from the various suppliers;
iii. Interdependence over the long term; the actors inside the supply chain rely on each other in terms of the mutual goals of reducing costs, improving qualities;
iv. Coexistence of cooperation and competition; the competition based on the conflicts about shared resources and its control;
v. Nonlinear dynamics; the relationships between the agents are often nonlinear. Nonlinearity is produced often by customer, the protocols and drift in the contractual relationships, decision-making process;
vi. Quasi-equilibrium and randomness; the supply chain network is trying to maintain the balance when is affected by the environmental changes, sometimes small event can misbalance the network and produce the structural reconfiguration;
vii. Self-organizations; a natural process of spontaneity and the self-regulative process;
viii. Adaptation and evolution: The supply chain network is receiving information from the environment, adapting and growing together with the environment dynamism. The environment is rugged and co-evolution is difficult because the entities inside the

network may change, exit, or stop interacting, in that case, the viability of the interactions among the entities change and affect the structural change of the network.

Looking at the supply chain as complex adaptive systems, the role of the retailer becomes crucial in defining the viability of the network. The exchange of the information with the consumer market and the supply market emerge the high responsibility of the control of the information that is circulating thru the position of the retailer. The control becomes the task of the managers and their decision on how much they can let and tight the organization in the emerging dynamic behavior.

The organization can assess, evaluate and understand its contextual environment (Smith, 1983). Economic conditions such as inflation, economic growth, exchange rates, and interest rates, or socio-political issues like international relations, law, and demography are examples of factors that form the contextual environment. The transactional environment is composed of the organizations and individuals with which the organization directly interacts (Vasconcelos and Ramirez, 2011).

The complexity and uncertainty that characterize the contextual environment force enterprises to improve and search continuously for new, often unconventional, solutions for shaping their decision-making processes. These solutions are organizational, technological and managerial. In the situation in which the external and internal conditions in which it functions are dynamically changing, an organization can survive and prosper only when it can adapt its changeability to the changeability of the environment in which it is operating (Gorzeń-Mitka and Okręglicka, 2014).

Moretta Tartaglione et al. (2019) discuss the complexity that has the three dimensions: variety, variability, and indeterminacy. Reflecting on the business environment, it distinguishes between the internal and external complexity. The internal complexity is determined by the size of the organizations, the constructs, and the variety of social roles and identities while the external complexity consists of the high variety and variability of environmental components that are challenging to predict and control. The acknowledge of the external environment from the perspective of the complexity (Thietart and Forgues, 1995; Cilliers, 2005; Schneider and Somers, 2006; Stacey, 2011; Blackman, 2013; Boulton et al., 2015; Devereux et al., 2020). In open systems, interacting with the environment, order and structure can create inside the system based on the energy from the environment. (Boulton et al., 2015). The causative link between environmental variables and forecasting plans becomes complex due to numerous variables and the chaotic nature of environments (Winsor, 1995). However, research trends in forecasting continue to stress the analysis and prediction of the inter-relationship between an organization and its environment (Polonsky et al., 1999). Uncertainty and high complexity can move the faster system to become flexible and responsive and to adapt to the market dynamics faster than usual. To manage complexity it is necessary to measure it (Alamoudi and Kumar, 2017). Reducing complexity is not possible but finding the methods to manage the level and the cost of complexity can help organizations to move forward and adapt rapidly.

The business environment reflecting the retail environment presents several complexity constructs (Moretta Tartaglione et al., 2018):

i. Co-determination or co-evolution taking place between firms and their environments (Achrol, 1991; Polonsky et al.,1999).

ii. Self-organisation and emergence occurring through participants in the environment (Peters, 1999; Mason, 2007).

iii. Environmental changes starting small and developing slowly and unpredictably (Mason 2007).

iv. Non-linear relationships (Black and Farias 1997; Mason, 2007).

Table 1-3. Complexity and environment – relevant topics and authors

Topics	**Authors**
Transmitted information	Shannon and Weaver, 1949;
Environment as complex adaptive system	Peters, 1999; Prendergast and Berthon, 2000; Choi et al., 2001; Begun et al., 2003; Mason, 2007; He et al., 2013 ; Bruni et al., 2016; Wollmann and Steiner, 2017; Moretta Tartaglione et al., 2018; 2019;
Connectivity in the supply chain	Choi et al., 2001; Sivadasan et al., 2002; 2006; Surana et al., 2005; Pathak et al., 2007; Battini et al., 2007; Rensburg, 2012; Chen et al., 2013; Nair et al.,

	2016; Zhao et al., 2019; Nair and Reed-Tsochas, 2019;
Static complexity	Battini et al., 2007; Ekinci and Baykasoğlu, 2016; 2019;
Dynamic complexity	Frizelle and Woodcock, 1995; Ekinci and Baykasoğlu, 2016; 2019;
Interactions	Surana et al., 2005; Moretta Tartaglione and Bruni, 2016; Wollmann and Steiner., 2017;
Contextual environment of organizations	Smith, 1983; Vasconcelos and Ramirez, 2011;
Changeability and adaptation	Gorzeń-Mitka and Okręglicka, 2014;
Forecasting plans	Winsor, 1995; Polonsky et al., 1999;
Complexity constructs	Moretta Tartaglione et al., 2018;
Complexity metrics/Measure of complexity	Frizelle and Woodcock 1995; Sivadasan et al. 2006; 2002; Battini et al., 2007; Rensburg 2012; Chen et al., 2013; Alamoudi and Kumar, 2017.

Source: Author's elaboration

1.4.3 Complexity with and within organizations

Organization can be defined as the complex systems that have adaptive and impulsive behavior because of interactions of many agents within the system (Brown and Eisenhardt, 1997). There is a certain limit to the complexity that individuals can deal with in a certain period. Due to that complexity is the variable dependent on the time (Simon, 1957). An organization that cannot cope with the change in time tends to fail being exposed to the penalty of not managing the priorities in a better way (Boisot, 2000). How organizations manage the complexity and the ability to react to changes it can be linked to the rules in natural sciences: organizations can go through the continual process of adaption and changes until they do not reach the dynamic equilibrium with their environments (Morel and Ramanujam, 1999). Some elements of the complexity match with the elements of the natural environment. The main characteristics of the complex system (Cillers, 2000;

2005) are :

i. Complex systems are made of a small but large number of elements that can be simple or either complex – e.g. individuals, organizations, processes, variables, (Morel and Ramanujam, 1999);
ii. The interaction between elements are dynamic and non-linear;
iii. Complex systems are open systems and work under non-equilibrium conditions (Morel and Ramanujam 1999);
iv. The behavior of the complex system has not been defined by the nature of the elements but of the interactions among the elements;
v. The interaction of elements produces emergent patterns of behavior (Morel and Ramanujam, 1999), while the interactions are emerging and dynamic, predictions about the system can be made by inspecting the dynamics of their relationships;
vi. Existence of feedback loops;
vii. The adaptability of the complex systems; complex systems can self-adapt without the help of the external element. Organizations have been seen as complex adaptive systems; they collect the information about surroundings and their activity with the environment and then try to incorporate their adaption and co-evolving with change in the environment (Kauffman, 1995).

Organizations tend to manage complexity in a way to reduce the information that they receive from the environment. The flow of information can be chaotic at the same time it is challenging to follow the speed of the receiving information, in that case, an organization is trying to simplify connections and the number of choices available to its members (Ashmos et al., 2002).

In the literature, it has been called the "complexity reduction" that is associated with organizational stability (Stacey, 1992). In reality, the complexity is not possible to reduce but more to manage and filter. Organizations are creating complexity by itself interconnecting and exchanging information, also with the external environment-e.g. consumers, suppliers, the political authorities, banks, labor, and other financial institutions. Stacey (1992) suggested that managers tend to establish stability managing the complexity and succeed, or experience instability and fail. Managers need to find the method for filtering and prioritizing the information from the bunch coming from external environments. There are different perspectives in the academic world regarding the perspective of complexity. The known chaos theory (Levy, 1994) assumes that because of an arisen complexity chaos is everywhere and produces chaotic situations. One of the solutions to respond to the chaotic environment is strategy development. Strategy development is a vital task of organizations (Stacey, 1993) and its planning and implementation differ from organizations to organizations in terms of adapting and absorbing complexity, how organizations are flexible, what the length of their strategic plan is, how the approach to stability with the environment is their resistance to change, and their development of mechanisms that support them at times of uncertainty.

Bititci et al. (2012) identify that the context in which PMM occurs is indeed changing and that commonly accepted PMM practices showed limited effectiveness due to the lack of a holistic approach that addresses potential external complexity.

1.4.4 Complexity of decision-making in management

A defined complexity in decision-making, as to the classic decision-making theory as non-psychological thanks to certain postulates or axioms for rational behavior, where it derives a function of real values, specific to the decision-maker subject, reflecting the context in which he operates, on value or utility, on stating that a specific choice is preferable to others if and only for the utility expected from this option results in being greater to the expected utility from alternative (Barile et al., 2011). The decision-maker faces the dimensions in which there is a loss of control, in constant searching for the tools and calculations that will find the optimal strategy to control complexity, but often falls into the limitation activities that can be done or are not reliable. Barile et al. (2011) discuss also that something different to do to decide with certainty when finding oneself ahead of emerging situations, never before tested and difficult to confront with usual tools and techniques. Usually, on every occasion in continuing to enumerate similarly as count the constituent elements and therefore "reduce" the truth in elements which are more simple, searching to get some reasonable qualification measure of complexity we are forced to work under a structure perspective instead of a system one. Even from the sensible point of view, it seems undeniable that the so-called structure approach and thus orientation to quantify factors, relations, components, variables, or others, is to be considered a privileged method to approach to gauge the emergence of complexity, leaving a major operative difficulty within the application of this sort of perspective. It derives from the observation that during a not so evident manner, the requirement for proceeding to the quantification determines that we pass inadvertently from complexity to complication. If the road between complicated and complicated is established by the numerical growth of one or more factors and not by other aspects, a classical dilemma is always: when do “few” factors become “many” and “many” become “a lot”? (Barile et al., 2011). This creates dynamism in markets that are not simple to interpret, examine or predict, making forecasting activities and marketing strategies challenging, particularly in the retail sector where complexity affects decision-making processes from store organization to supplier and customer (Moretta Tartaglione et al., 2018). The retail environment has passed tremendous change (Hamel and Prahalad, 1994; Kotter, 1996; Loewen, 1997; Conner, 1998) thus increasing complexity (Achrol, 1991). Complexity is defined as the measure of heterogeneity and variability in environmental factors (Lane and Maxfield, 1996; Chae and Hill, 1997; Chakravarthy, 1997). When complexity increases, the ability to understand and use information to plan and predict becomes more difficult (Black and Farias, 1997), the decision windows shorter, the risk of obsolescence greater, and the long-term control b making (Hummelbrunner and Jones, 2013).

In decision theory which postulates are complexity, uncertainty and unpredictability, considers strategic management (Mintzberg and Waters, 1985; Mintzberg and Westley, 2001). the decision-maker must possess the framework to filter, simplify and prioritize the knowledge. French et al. (2009) highlight the Cynefin framework (Kurtz and Snowden, 2003; Snowden, 2000, 2002; Snowden and Boone, 2007, Alexander et al., 2018) as a useful and effective summary model. This model focuses on the influence of organizational structure and therefore the environment during which operates. The concept has been supported by “sense-making” from what's known and in a way standardized This model senses the multiple factors from the environment which provides a typology of information to help managers in interpreting the character of the context of assorted decision problems. It classifies the systems into structured and unstructured ones. Within the structured systems, information awareness of an organization’s performance is straightforward to capture by having the info and controlling them in a very thanks to simplifying the manager’s decisions about the longer-term actions. In un-structured systems thanks to uncertainty and lack of predictability, the main target is on the response and emerging from the context rather than the forecasting and control at now the manager has to decentralize the decision-process e.g. engaging the stakeholders.

With the increasing complexity of the environment, decision-making processes are evolving, with a range of flexible and specific frameworks being employed to attain goals and to manage relationships efficiently to get a competitive advantage. Vicari (1998) has summarized the factors which seems to steer to growth in complexity for business organizations:

- the factors that management needs to consider, the number of knowledge that should be used, and also the number of elements that are included within the problem that must be resolved;

- the range of the information itself, the number of various problems, and therefore the number of variables affecting the problem;

- the variability of facts that a manager needs to take into consideration, the non-stable nature of the issues, and also the inconsistency of the issues over time; and

- the variables that may influence the matter which isn't known by the decision-maker.

In the management literature, decision-making in a very complex environment is required for a corporation to work effectively under conditions of uncertainty and rapid change. Strategic decision-making in complex environments requires metacognitive skills that provide leaders with a “tool bag” of decision options to use when confronting novel situations. This also requires the event of innovative and adaptable decision models (Hübner et al., 2013) beyond the linear thinking underlying the rational actor model that has characterized traditional strategic decision-making (Hummelbrunner and Jones, 2013).

The decision-maker in the retail supply chain has to look from the perspective as the whole, where there is a need to control the components where it is crucial (Reeves et al., 2020). Where is a big amount of SKU, the decision-maker creates the variety in which the optimal balance of flexibility and complexity has the role. The decision-maker has to look at the complete supply chain, the different parts of the value chain, at the beginning of producing the product there is a need for high standardization and simplicity, and more flexibility in packaging operations until the product is exposed at the shell. The complexity can be control if and only if the process is considered as the whole picture and adding and subtracting complex decisions where it is necessary (Reeves et al., 2020)

Looking at the whole picture and every time complexity is added, it has a clear strategic purpose. Focusing on controlling it where that matters most—in such case, in production and value chain, where the variety of SKUs need to be under controlled amount. Managing different parts of the value chain differently: high standardization and simplicity in the capital-intensive upstream manufacturing, and more flexibility and variety in the more labor-intensive downstream packaging operations. With the upstream components subtract some complexity every time it added a part.

Table 1-4. Complexity of organizations and decision-making in retail management - relevant topics and authors

Topics	**Authors**
Interactions of agents	Brown and Eisenhardt, 1997; Morel and Ramanujam, 1999;
Ability to react to changes/Adaptability	Kauffman, 1995; Morel and Ramanujam, 1999; Boisot, 2000;
The main characteristics of the complex system	Kauffman, 1995; Morel and Ramanujam, 1999; Cillers, 2000; 2005;
The flow of information/complexity reduction	Kauffman, 1995; Stacey, 1992; Ashmos et al., 2002;
Strategy development	Stacey, 1993;
Measure of complexity	Barile et al., 2011;
Strategic management	Mintzberg and Waters, 1985; Mintzberg and Westley, 2001;
Cynefin framework-structured and unstructured systems	Snowden, 2000, 2002; Kurtz and Snowden, 2003; Snowden and Boone, 2007; French et al., 2009; Alexander et al., 2018;
Retail environment	Achrol, 1991; Hamel and Prahalad, 1994; Kotter, 1996; Loewen, 1997; Conner, 1998; Moretta Tartaglione et al., 2018;
The measure of heterogeneity and variability	Lane and Maxfield, 1996; Chae and Hill, 1997; Chakravarthy, 1997;
Factors for a growth in complexity	Vicari, 1998;
Innovative and adaptable decision models	Hübner et al., 2013;
Traditional strategic decision-making	Hummelbrunner and Jones, 2013;
Complexity control	Barile et al., 2011; Reeves et al., 2020;
Subtracting complex decisions	Reeves et al., 2020.

Source: Author's elaboration

Bibliometric analysis
(2nd part)

1.5 Social network analysis on bibliometric data

This part of the research uses the bibliometric analysis to describe the patterns in literature with the targeted topic using quantitative data (Sarkar and Searcy, 2016). The bibliometric methodology has been based on quantifying the flow of information in two ways: 1) using the whole publication and its characteristics as the citation, author's name, keywords and source 2) identifying the links among the features, co-citation, co-authorship, co-occurrences and creating the networks of clusters (Ding et al., 2001, Perianes-Rodriguez et al., 2016). In this work the social network analyses SNA is used to recognize and classify related nodes of keywords, this method has been also used for identifying the relations between authors or research sources and institutions to assess associations and collaborations (DeNooy et al., 2005). Although, these procedures identify the relations (co-occurrences) of certain items, such as the number of times that keywords (co-word), citations (co-citation), and authors (co-authorship) are mentioned together in publications in a particular research field. This approach is mainly used to understand the underlying frame of the interrelationships between articles (Ding et al., 2001, Knutas et al. 2015). Besides, the network visualization has been supported by the VOS viewer (Van Eck and Waltman, 2010; 2014), the tool for the textual mining and visualization of the bibliometric mapping.

1.5.1 Findings of bibliometric analysis

Bibliometric analysis is performed using the keyword search of ISI Web Of Science by the topic about the external influence on the forecasting and forecasting sales in the period of 2013 to 2020, focusing on the business and marketing research domain. The following combination of keywords are used: at the first search 1) "forecasting" with "external complexity", "environmental variables", "heterogeneous market", at second 2) "forecasting sales" with "external complexity", “environmental variables”, “heterogeneous market”, “retail environment”, “competition”. Table 1-5, shows the total sum of the articles found on the Web of Science engine by the keywords combination.

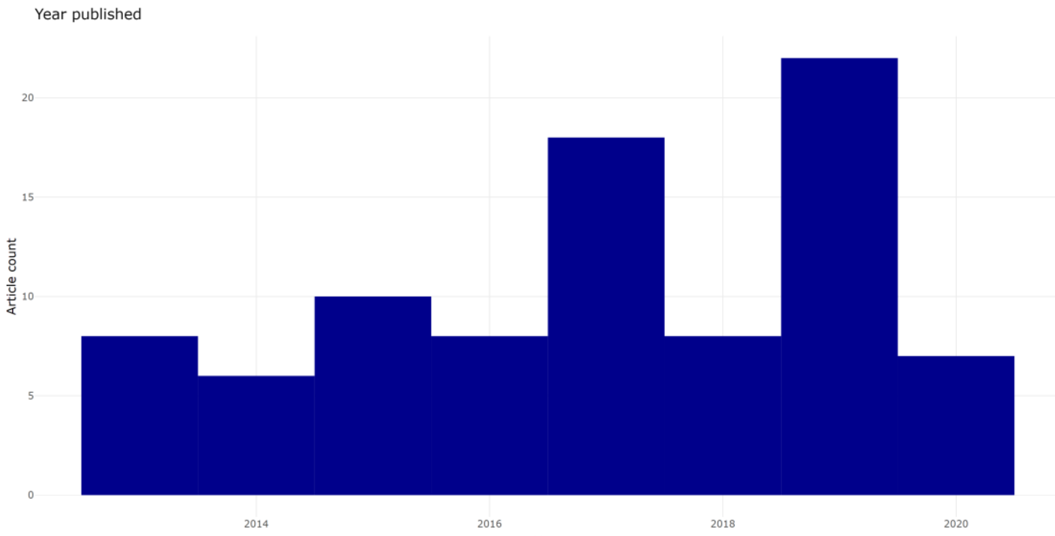
Table 1-5. The number of articles by keywords search

Keywords combination (2013-2020)	**Forecasting**	**Forecasting sales**
External complexity	81	-
Environmental variables	64	7
Heterogeneous market	198	10
Retail environment	-	17
Competition	-	55
Total	**343**	**72**

Source: Author’s elaboration

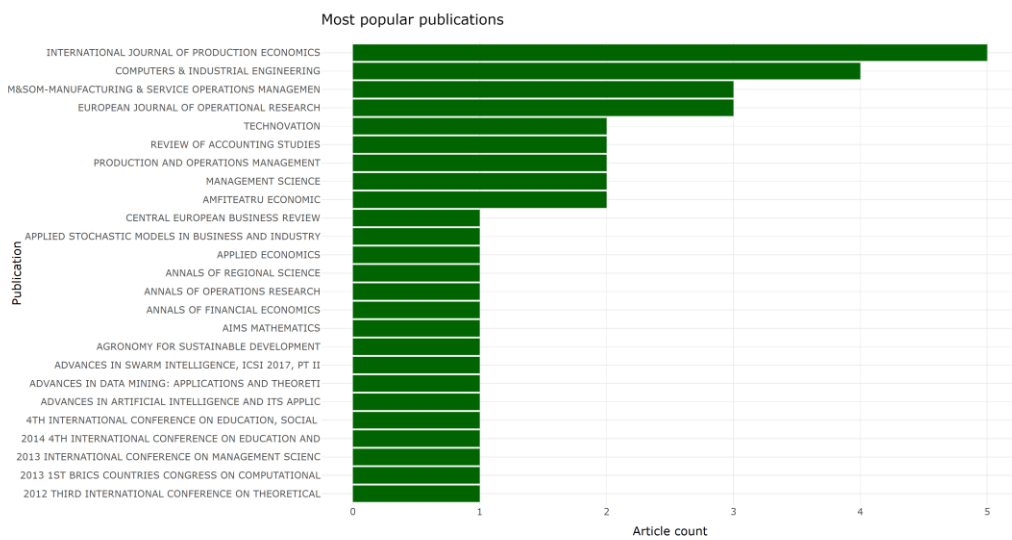
In the first research, 343 articles are identified via keyword combination with “forecasting”, at second, 72 articles are outcome via keyword combination with “forecasting sales”. The analysis, using *R* meta-language, identifies the important authors, journals and keywords in the dataset, based on the number of occurrences and citation counts. The selected articles of 72 articles are used for the analysis to detect the major bibliometric links related to the forecasting of sales (Table 1-5). The trend of publishing the articles related to the topic is gradually increasing from 2013 (Figure 1-3), showing that the major articles (22) are published in 2019, the most popular and most cited publications (Figure 1-3 and Figure 1-4) are issued by *International Journal of Production Economics, Agronomy for Sustainable Development, Computers and Industrial Engineering, European Journal of Operations Management, Management Science.*

Figure 1-3. The trend of issued publications on the topic 2013-2020



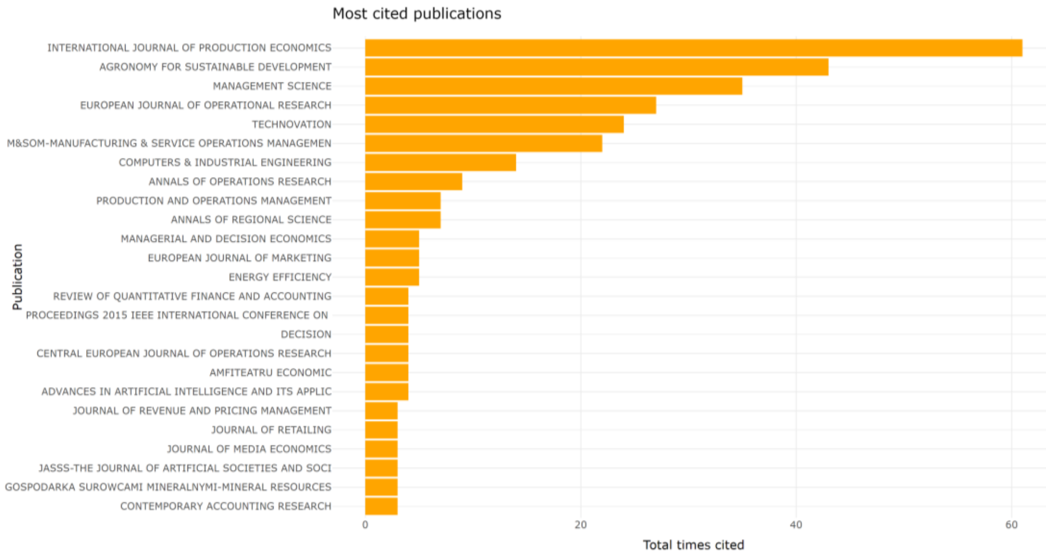
Source: Author's elaboration

Figure 1-4. The most popular journals by the number of issued publications 2013-2020



Source: Author's elaboration

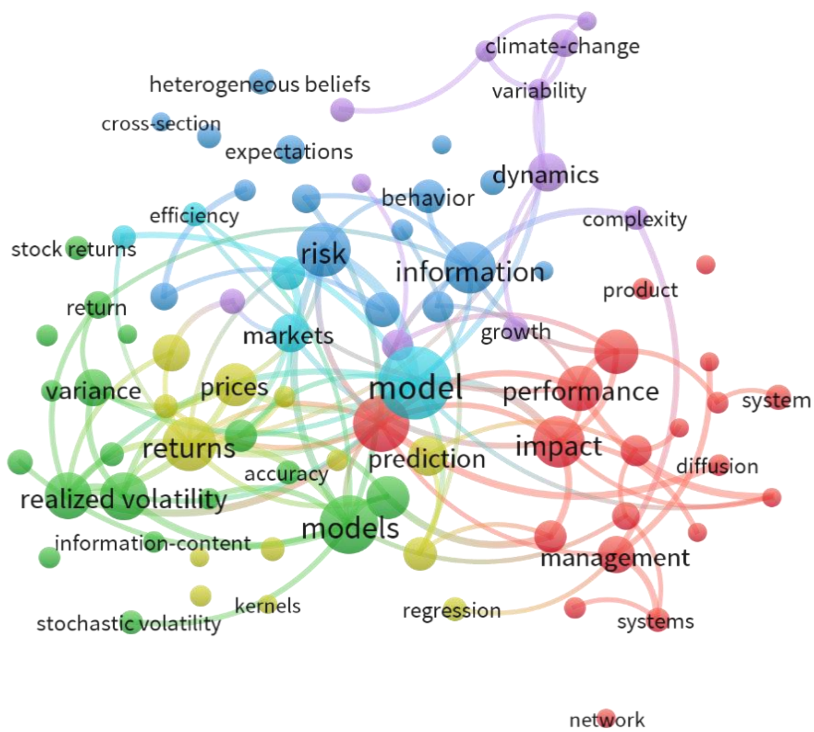
Figure 1-5. The most cited publications in journals 2013-2020



Source: Author's elaboration

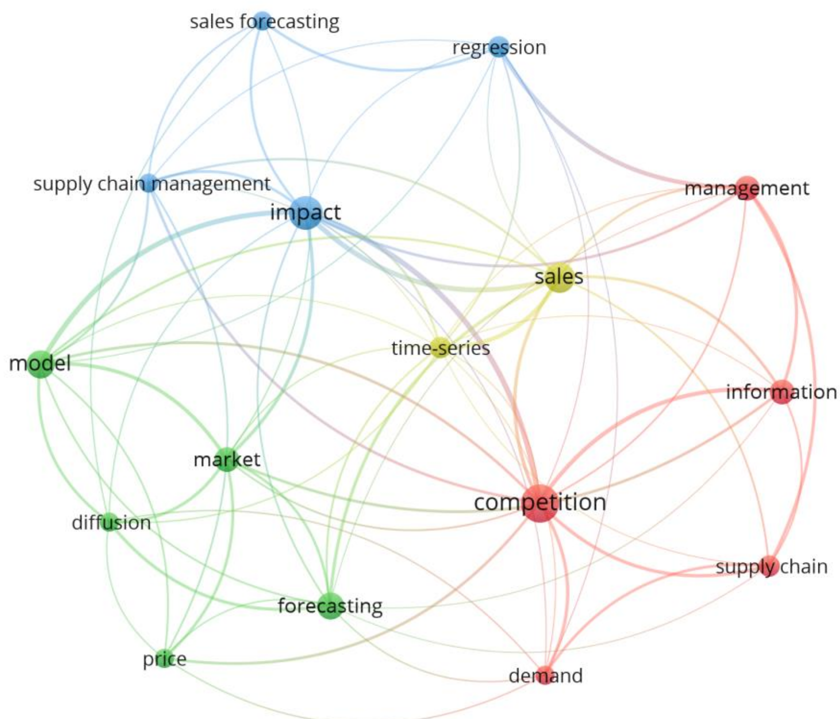
In the next step, from the completed data of articles related to "forecasting" and "forecasting sales" textual mining software VOSviewer (Van Eck and Waltman, 2010; 2014) using the fractional counting of co-occurrence of the author keywords define the cluster of the most interacted occurred keywords. Fractal counting means, each term cited in a publication has the same influence in a bibliographic coupling analysis, which means that each keyword is considered to be equally representative of what the publications about, not only the most cited keyword. In Figure1-6., there are outlined four main clusters of keywords defined by colors. In red color, the cluster with the highest co-occurred keywords with the strong association strength are "complexity", "system", "environment", "factor", "level", in green, the cluster of "market", "volatility", "performance", "evidence" and "price" and in blue the cluster of "heterogeneity", "agent" and "expectations". The presence of the high association strength of the relationship between the cluster "red" and cluster "green" is evident. Deepening keyword analysis, selecting only the articles related to the "forecasting sales", network analysis is done on the words or phrases that frequently appear in the titles of an article's references, but do not appear in the title of the article itself. That is the unique algorithm behind the search tool Web of Science, extracting the so-called KeyWords Plus improves the power of cited-reference searching by searching across disciplines for all the articles that have cited references in common. In this case, according to Figure 1-7 there are clearly defined clusters of highly correlated words, the most apparent is the cluster red with the word “competition”, “information”, “management”, “supply chain” and “demand”, in the green cluster “model”, “market”, “forecasting”, “diffusion”, “price”, in the blue, there are “impact”, “regression”, “sales forecasting”, “supply-chain management” and within the yellow cluster the words most occurred and high association strength “sales” and “time series.”. In the final figure (Figure 1-8) it is shown the most popular keywords by the number of times cited. The top-cited keywords are “forecasting”, “demand forecasting”, “concentration”, “agriculture”, “food waste”, “FMCG market”, “external factors” etc.

Figure 1-6. Clusters of the most interacted occurred keywords



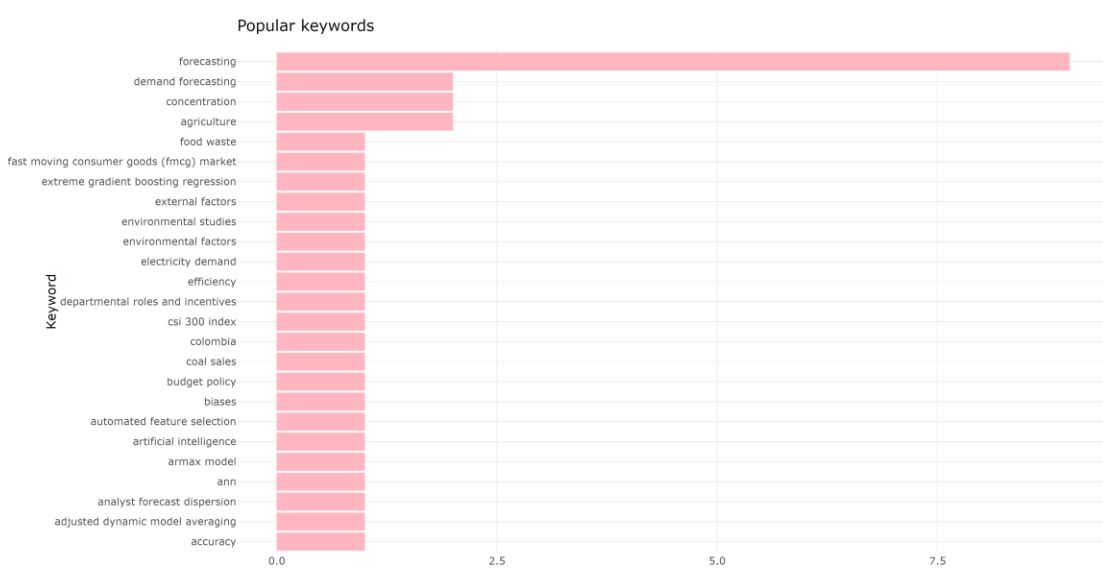
Source: Author’s elaboratio

Figure 1-7. Clusters of the most interacted keyword searching across disciplines for all the articles that cite common references



Source: Author's elaboration

Figure 1-8. The most cited keywords



Source: Author's elaboration

Expanding the social network analysis with the detailed analysis of the high-scored papers (Table 1-7) along with relevant articles through Google Scholar (Table 1-6), it emerges that the first problem for the authors is the choice of the environmental factors to include in the forecasting methods (Moretta Tartaglione et al., 2018).

Taking into account the high-ranked papers and looking at the most popular and cited keywords it is possible to see what is the main impact on forecasting in research works. Everingham et al. 2016 investigated the *accuracy* of the prediction model to observe how the *seasonal climate* influencing the annual variation in sugarcane productivity generated from the predictor variables. Climate change is set apart as an influential factor (Moretta Tartaglione et al., 2018) thus *climate change* affects and makes changes worldwide. It affects production plants, farms, water resources, translating influence via marketing channels until the final consumers. Climate change is a hardly controlled factor, involving in the future strategies and plans, companies can better prepare for weather surprises and predict better the impact on the forecasting models.

In developing countries, the farmers make the production and sales decisions basing on the informal source of information as family, neighbors, or simply traditional habit. Lacking accurate *information* on *climate* and *prices* can lead to inefficiencies in the production and sales of *agricultural* products that can affect the profit of the small-scale farmers (Camacho and Conover, 2019).

Market volatility, its stochastic *behavior*, the level of the sales agent's activities determine the dynamics of the *demand,* in other words, the interrelations between the manufacturer and the seller determine the demand signals and the *accuracy* of the prediction (Kung and Chen, 2014). Authors identify that the accuracy of the sales forecasting models influence the profitability of the manufacturer, in the case when it is high, it is beneficial for both, then when it is low both sides must tend to improve The relation inside the *supply-chain* and alignment of the forecasting models, considering the market volatility, have the high *impact* on the accuracy on the *forecasting demand* and keeping it discrete does not bring the profitability either to manufacturer or sales agent. Study of the dynamics of the relationship between sales and internal (promotions) and external environmental operators (Consumer Price Index, Unemployment rate, Fuel Price, Temperature) in a retail company using a systems outlook to support retail management decisions with nonlinear methods (Moretta Tartaglione et al., 2019).

The inventory in the retail store stock can affect the demand and the inventory record inaccuracy can produce *bias* in the prediction of the demand.

Mersereau (2015) investigates the unknown on the forecasting demand in an *environment* with censored observations, the result is the "systemic downward bias" in the prediction of demand under the assumption of the inventory record inaccuracies. Demand in the fashion industry is highly volatile and the rapid changes make the *supply chain* depend on *efficient* and quick forecasts that will involve multiple stock keeping units (SKUs). The volatile fashion market leans on the different demand forecasting *systems* that will align the *supply chain management* and benefit the interests of multiple actors. The need for the computational challenging *models* for analyzing the big amount of the sales data (Ren et al., 2017).

The literature on forecasting demand has outlined the importance of the *price*. Discount or the promotional strategies as the direct effect on the change of the demand. Promotions are becoming the shed light on how to predict the demand for perishable products (Donselaar et al., 2016). The *competition* arises due to the technological advantages of the products and fast switching the generations of the novel products. That implies cost reductions due to the price war and simulate price reductions. Tsai (2013) studies how conventional multi-generational models ignore the price effect on the market potential substitutes and introduces the novel *diffusion* model including the *heterogeneous* price elasticity and the consumers' affinities to the LCD TV sizes. The necessity for the "adaptable, interpretable, accurate and stable" *forecasting algorithm* for promotions (Aguilar-Palacios et al., 2019) that will perform well under different product categories, to be flexible to the market *change* and diverse geographical locations and at the same time to be significant, accurate and operative for *supply chains*.

1.5.2 Findings of literature review

In the first part of this work there is the extinguished literature review on the environment, its characteristics and how environment elements are defining the market changes and retailing, in the second part the SNA on the literature about forecasting demand or sales results in the network of the keywords that are the most obstacles for the forecasting sales (demand) management. In following part, using Google Scholar, it will be outlined the research findings on the relevant research about the forecasting in retailing and what the most influential external factors are.

As retail business is becoming a complex set of activities also its forecasting sales is becoming a complex activity that does *not lean on a one factor or one method*, as the history of sales, but many factors that are indirect or direct linked to the sales: orders, promotions, customer behaviors, logistic information, competition distance, the price of gas, etc (Moretta Tartaglione et al., 2018).

The focus is on the cooperation inside the supply chain among actors to achieve efficient and more aligned forecasting management. The authors propose a model on forecasting management model and represent supply chain model role, demand management, *collaborative coordination* (Albarune and Habib, 2015). The internal short-term production planning controls (Towers and Ashford, 2001) the real-time production planning and partnerships are principal for the transformation process between customer, producer and supplier from the perspective of the supply chain. Defining the assortment size and shelf space planning is important for retailers to better meet consumer demand and to respond to increased product variety. Study of the effectiveness of rebates and returns contracts when a retailer can improve her demand information quality by employing *costly forecasting effort* where contracts play two roles: Giving incentives to impact the retailer's forecasting decision and obtaining information by the forecast to characterize production decisions (Taylor and Xiao, 2009). In food supply chains it leans on *collaborative forecasting* between manufacturers and retailers. In a market where the retailer doesn't share actual forecast with the manufacturer at the first stage of a natural disaster, his and the retailer's profit increase. The retailer might share its actual forecast with a manufacturer to get a *lower wholesale price*. The grocery planning framework integrates retail specifics, hierarchical and sequential aspects of decision making and that is why it also creates the basis for research and development of *advanced decision support systems* (Hübner et al., 2013). Also, retailer's profit under the noninformation sharing case is always higher than under the information sharing case (Zhang, 2009). Eksoz et al. (2014) identify trends, gaps and areas for future research involving partners' integration, *information sharing* and forecasting processes in the supply chain. The addition of an *online alternative channel* by an established brick-and-mortar retailer raises problems and the aim of making the customer view the alternatives raises the question of complementarity versus substitution (Fildes et al., 2019).

Time management differs from to type of the retail industry. For instance, in the fashion industry, operational decisions struggle with the *tight schedule and limited data* when corresponding forecasting method has to be completed (Choi et al., 2014).

The proposition of an analytical framework for the design of supply chains that are resilient to *nonlinear system dynamics* and providing a systematic procedure for the analysis of the impact of nonlinear control structures on systems behavior (Spiegler, 2013).

Study of the dynamics of the relationship between sales and internal and external environmental operators in a retail company using a systems outlook to support retail management decisions with *nonlinear methods* (Moretta Tartaglione et al., 2019).

Mintz and Currim (2013) discuss how often companies have to adapt their marketing and service technology to stay competitive in the market and cover the complexity of retailing business in their performance measurement design (*complexity*).

Demand forecasting is affected strongly by *consumer behavior* that is hard to predict and how the information from the consumer is processed is becoming the crucial point in the supply chain. Retailers can vary products or prices based on anticipated consumer behavior and it refers not only to demand forecasts but also to optimization methods to steer consumers' behavior (Moretta Tartaglione et al., 2018). *The location of the decoupling point* has a misleading impact on planning separating the planning tasks into forecast-driven and order-driven processes (Hubner et al., 2013).

Study of the *consumers' socio-demographic profile* impact on their decision to buy new or pre-owned cars beyond different car segments, finding that economic (financial income, access to credit), individual (age, level of education), household (parking space) aspects have an impact on car segment and car type choice (Prieto and Caemmerer, 2013). Customer-generated content from different internet sources can potentially reduce demand forecasting errors in the short term and refine response to demand volatility in retail pharmacy positions (Papanagnou and Matthews-Amune, 2018). Examination of the benefit of modeling consumer demand with distributions when calculating the planned orders of products with low and irregular demand (Sillanpää and Liesiö, 2018). The author explores whether observing *Internet browsing habits* can inform practitioners about aggregate consumer behavior in an emerging market. Scanner, point of the sales are the source of data for the demand management to integrate and where can be detected the true demand signals. The trend in estimating demand is to lean on the promotions internally and the highest percentage of the demand forecasting is based on promotions. As promotions are internal drivers, the company is interested in the monitoring of the competitor's promotions. Huang et al. (2014) investigate the value *of competitive information* including the competitive price and the competitive promotion in forecasting product sales at the UPC (universal product code) level for retailers. In the presence of thousands of products, the focus is how to overcome *high-dimensionality* by investigating the value of inter and intra category SKU-level promotional information (Ma et al., 2016). Seasonality and weather impact are two factors usually interrelated in the retail sector. More seasonal clothing is sold when the temperature changing is extreme, but the temperature changes from day to day or week to week do not affect the sale of seasonal clothing. The effect of *weather* on seasonal sales clothing depends on the geographical location of a store, type of products and/or type of store (Bahng and Kincade, 2012). Intra-annual weather changes notably affect retail sales and the unveiling of retail categories to weather are not the same depending on the season, the reactions of retail categories to the same weather changes are differing greatly (Bertrand and Parnaudeau, 2017). *Climate and weather change* such as precipitation, humidity rate, wind, and temperature is predominant explanatory variables that are affecting retail sales. To respond to weather changes, retail managers should develop risk-strategies. Arunraj and Ahrens (2016) aim at developing a model to analyze *the relationship between weather and retail shopping behavior* (i.e., store traffic and sales). The proposition of a methodology that controls the impact of the short-term and the long-term weather unpredictability on the forecast (Verstraete et al., 2019). Considering uncontrollable factors such as weather or climate change into the forecasting model can make a more reliable model reflecting the specific moments when the switching in demand can be related to the weather conditions.

1.6 Discussion

Literature overview of the concept of the environment in the first part helps to determine what creates the retail environment and to patch the trend on what is affecting the forecasting activities. The results of the network analysis on the keywords search allow identifying the most current drivers from the environment that researchers considering in their studies. The studies are systematically organized in Table 1-6, according to the type of the external factor, the role of the forecast horizon, and the solution that were be able to be found. The common base for almost all research are that environment is uncertain, complex and hard to manage. That is reflected in the *retail supply chain* in terms of the *information sharing* among and complexity of the *relations inside supply chain*, the need for the *collaborative coordination*; *market volatility*, the necessity to monitor *competition's actions*; the influence of the *macroeconomic variables, climate and weather conditions*; on consumer demand side, the *socio-demographic profile* of consumers, *irregular demand, internet habits* of consumers. The impact of the alternative *online channel*, *internet sources* and *high dimensionality* of the *big sales data* for what the stable flexible, *forecasting algorithm* are needed.

Table 1-6. Theoretical framework of research on the external influence on the forecasting in retailing

Authors	[2]**Type of external factor**	**The role of forecast horizon**	**Solution**
Mintz and Currim (2013)	Competition/Industry concentration	Long-term	The inclusion of firm strategy, environmental characteristics and the use of financial metrics are positively associated with marketing-mix performance.
Chase Jr, C.W.(2013)	Relations inside supply chain/Unexpected external events	Long-term	Unconstrained demand forecasting.
Carriere-Swallow and Lebbe (2013)	Internet	Short-term	Improve the efficiency of nowcasting models.
Hubner et al. (2013)	Coordination of the supply-chain planning activities	Long/Mid/Short-term	Advanced decision support systems
Hubner et al. (2013)	Location of the decoupling point	Short-term	Optimization methods to steer consumer behavior, demand to forecast order-driven processes.
Huang et al. (2014)	Competitive information	Short term	Model with competitive information is accurate in the promotional period.
Tsai (2013)	Competition among the multilevel generations of the technological products	Short-term	Diffusion model that incorporate the heterogonous price elasticity and consumer behavior.
Spiegler (2013)	Resilience of supply chain	Long-term	Use of nonlinear control structures
Eksoz et al. (2014)	Shared information in a supply chain	Long-term	Collaborative forecasts between retailers and manufacturers.
Choi et al. (2014)	Limited data/Seasonal cycle	Short-term	Efficient forecasting models that will help operational

[2] General definition of a nature of factor (can be a group of factors).

			decisions with a tight schedule.
Albarune and Habib (2015)	Collaborative coordination	Long-term	Efficient and aligned forecasting model
Mersereau (2015)	Inaccuracy of inventory stock	Short-term	Method that sharply reduces the bias
Arunraj and Ahrens (2016)	Weather	Long-term	Developed a model that analyzes the relationships between a weather and retail shopping behavior.
Harrauerand Schnedlitz (2016)	Financial crisis	Short term	More flexible performance–reporting procedures.
Everingham et al. (2016)	Seasonal climate	Long-term	Accurate crop prediction can yield better demand prediction
Ma et al. (2016)	High dimensionality	Short-term	Advanced forecasting model to overcome high number of variables
Camacho and Conover (2019)	Accurate information on weather	Short/Long-term	Better use of information on weather and price in small-size farmers.
Fildes et al. (2019)	Online channel, high dimensionality, little data	Long/Mid/Short-term	Systemized theoretical and practical approaches
Kung and Chen (2014)	Market volatility	Long-term	Improving the accuracy of the retailer's forecasting model benefit both manufacturer and retailer.
Prieto and Caemmerer (2013)	Socio-demographic profile of consumer	Long-term	Income, age, education, household impact choice demand for car.
Moretta Tartaglione et al. (2019)	Consumer Price Index, Unemployment rate, Fuel Price, Temperature	Short/Long term	Nonlinear analysis apprehend better the impact of internal and external environmental drivers on retail sales.
Bertrand and Parnaudeau (2017)	Climate and weather change	Short/Long-term	Precipitation, humidity rate, wind and temperature are affecting retail sales.
Ren et al. (2017)	Big amount of sales data	Short-term	The need for the computational challenging models.
Aguliar-Palacios et al. (2019)	Promotions under different product categories and different locations	Short-term	The need for stable, accurate and flexible forecasting algorithm.
Papanagnou and Matthews-Amune (2018)	Internet sources	Short-term	Customer content can reduce the demand forecasting errors.
Sillanpää and Liesiö (2018)	Irregular demand	Short-term	The use of distributions increase the precision in the replenishment order.
Verstraete et al. (2019)	Weather conditions	Short/Long term	The proposed framework adopts the methods that are effective for predicting energy

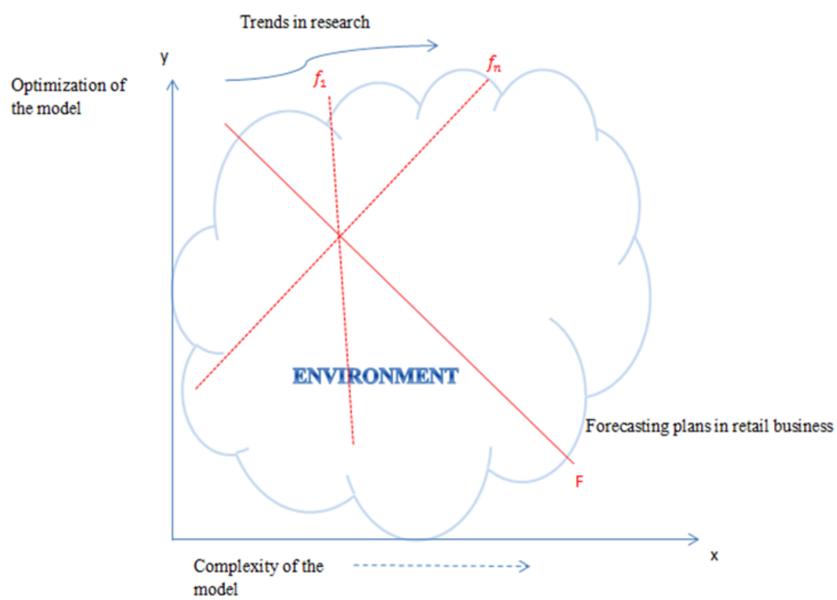
			demand and conveys the challenges that the retailer is facing.

Source: Author's elaboration

Matching the perspectives on how to handle the environment in the forecasting process in one framework and looking for the relations and trends, authors (Moretta Tartaglione et al., 2018) explain the inverse relation (Figure 1-9) between two aspects –complexity and optimization-, in the turbulent environment when complexity (*x*) increases it is challenging to optimize forecasting activities in retail considering many volatile factors, then optimization (*y*) decreases and vice versa, optimization (*y*) increases when complexity (*x*) decreases; in other words, it is simple to maintain forecasting models (*F*), not considering the environmental impact. In the middle point, complexity and optimization are delivering the "optimal" forecasting model. This trend can change considering the development of advanced technical tools cognitive analysis, support vector machines, artificial intelligence that can handle efficiently the high dimensionality of big data. The trend will move these perspectives (*f1-fn*): by increasing market volatility, new forecasting tools will be developed to try to maintain complexity efficiently so that the complex model becomes the optimal model (*fn*).

The relationship between complexity and optimization will tend to become linear, and the complex model with a high number of external factors will be the optimal, stable, and efficient model. The results of the network analysis on the keywords search allow identifying the most current drivers from the environment that researchers considering in their studies. The studies are systematically organized in Table 1-6, according to the type of the external factor, the role of the forecast horizon, and the solution that was able to be found. The common base for almost all research is that environment is uncertain, complex, and hard to manage. That is reflected in the *retail supply chain* in terms of the *information sharing* among and complexity of the *relations inside the supply chain*, the need for the *collaborative coordination*; *market volatility*, the necessity to monitor *competition's actions*; the influence of the *macroeconomic variables*, *climate and weather conditions*; on the consumer demand side, the *socio-demographic profile* of consumers, *irregular demand*, *internet habits* of consumers. The impact of the alternative *online channel*, *internet sources*, and *high dimensionality* of the *big sales data* for what the stable flexible, *forecasting algorithm* is needed.

Figure 1-9. Future of the forecasting plans in relation to the optimization and complexity of the models.

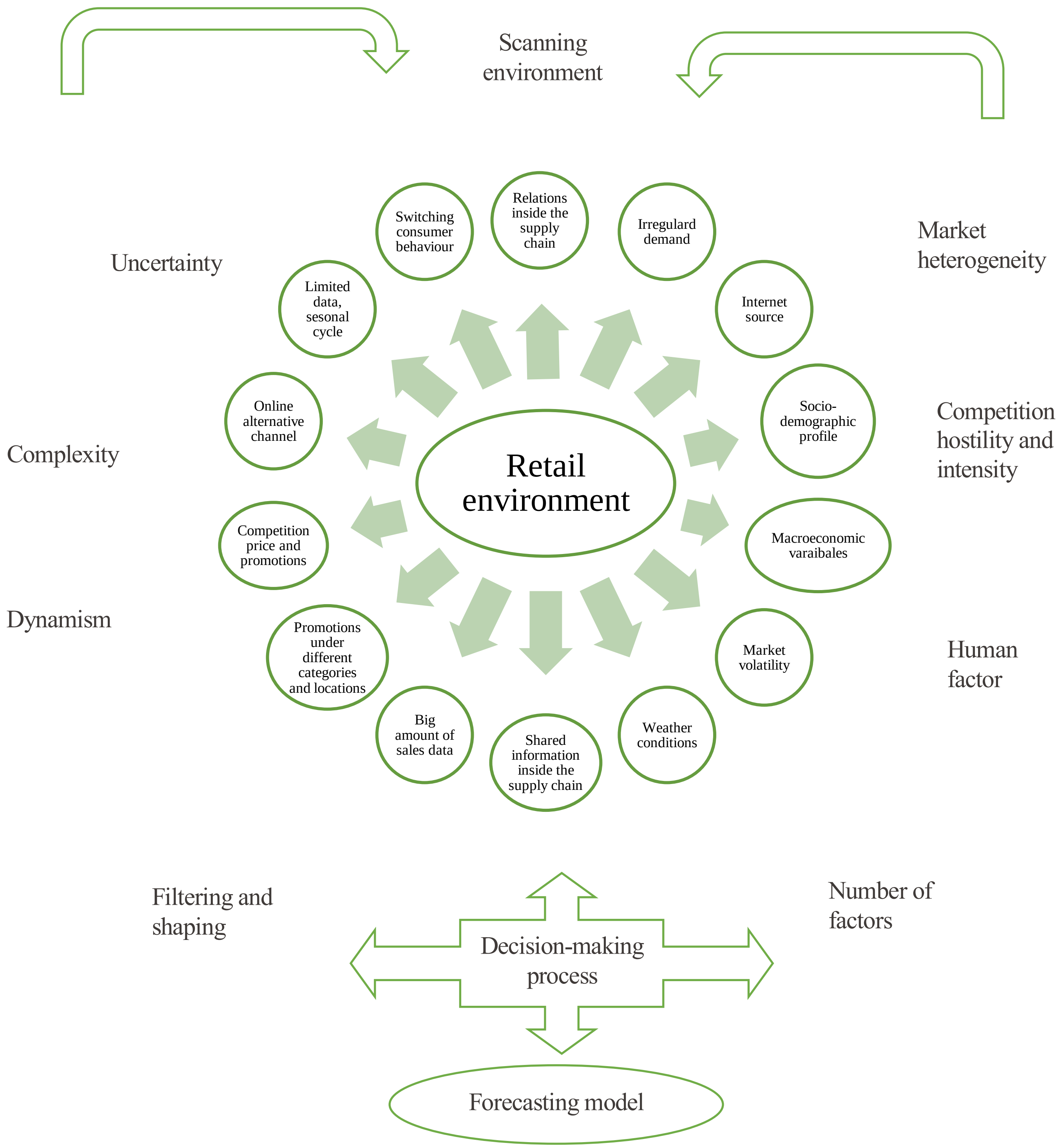


Source: Figure 4: The environmental impact on forecasting plans in retail business through relation between complexity and optimization of the forecasting plan. Reprinted from Moretta Tartaglione, A. M., Bruni, R., Bozic, M., and Cavacece, Y. (2018). How the environmental complexity affects forecasting in retail? Actual research trends. *Journal of Marketing Trends (1961-7798)*, *5*(2). Copyright 2009, Marketing Trends Association (Paris – Venice Association).

1.7 Conclusion

The retail environment has dynamic and unpredictable behavior and is influencing forecasting activities. The base for every forecasting plan to detect the pattern or relationship to insert in the model and scanning the correct relationships with the environment is becoming a valid tool to refine the demand management. The trend in research on the impact of the drivers from the environment shows that the focus is to build a coordinated forecasting system inside the retail supply chain allowing both manufacturers and retailers to benefit from the information shared from the forecasting. The accurate forecasting models can reduce the cost of inventory and replenishment stock, regulate prices and promotional strategies, creating an accurate shelf display and improve the anticipated decision-making process. Considering the findings from bibliometric analysis and the relevant papers about the topic, the trend is passing from the past research which generalizes the concept of the environment and defining its main characteristics: complexity, uncertainty, and dynamism to recent where research observe market as heterogeneous where the level of the competition intensity depends on the competition hostility with the necessity for the continuous scanning of the environment (Figure 1-10). Reflecting on the environment in retailing the decision-maker has to scan and filter the right patterns and define the right relationships between environmental factors and the demand in the forecasting model. The same decision-maker may have shifted attitudes in the modeling strategy and that brings to the main cause of unpredictability and it is the human factor that associate with trends and self-fulfilling aims to change the influence of the outcome (Chase, 2016). The trend of future forecasting models will be to 1) improve the collaboration and information sharing inside the supply chain 2) centralize the forecasting decisions on the retailer's side, 3) manage better the flux of data 4) to detect the pattern of the volatile demand 5) include the competition's price and promotions 6) monitor the impact on weather 6) develop the flexible forecasting algorithm 7) consider the pattern from the online channel and internet sources 8) detecting sharply the consumer profile and anticipate shopping behavior.

Figure 1-10. The concept of the retail environment for the decision-making process in forecasting



Source: Author's elaboration

1.8 Appendix

Table 1-7. The most cited papers from the Social Network Analysis (SNA)

YearPublished	TimesCited	Article
2016	43	ACCURATE PREDICTION OF SUGARCANE YIELD USING A RANDOM FOREST ALGORITHM | EVERINGHAM Y, 2016, AGRON SUSTAIN DEV, V 36, DOI 10.1007/S13593-016-0364-Z |
2016	13	ANALYSIS AND FORECASTING OF DEMAND DURING PROMOTIONS FOR PERISHABLE ITEMS | VAN DK, 2016, INT J PROD ECON, V172, P 65, DOI 10.1016/J.IJPE.2015.10.022 |
2013	31	PROVISION OF INCENTIVES FOR INFORMATION ACQUISITION: FORECAST-BASED CONTRACTS VS. MENUS OF LINEAR CONTRACTS | |
2013	12	MODELING DIFFUSION OF MULTI-GENERATIONAL LCD TVS WHILE CONSIDERING GENERATION-SPECIFIC PRICE EFFECTS AND CONSUMER BEHAVIORS | TSAI B, 2013, TECHNOVATION, V 33, P 345, DOI 10.1016/J.TECHNOVATION.2013.05.002 |
2015	9	DEMAND ESTIMATION FROM CENSORED OBSERVATIONS WITH INVENTORY RECORD INACCURACY | MERSEREAU AJ, 2015, MandSOM-MANUF SERV OP, V 17, P 335, DOI 10.1287/MSOM.2015.0520 |
2017	9	A COMPARATIVE STUDY ON FASHION DEMAND FORECASTING MODELS WITH MULTIPLE SOURCES OF UNCERTAINTY | REN S, 2017, ANN OPER RES, V257, P 335, DOI 10.1007/S10479-016-2204-6 |
2015	7	PROMOTION PLANNING AND SUPPLY CHAIN CONTRACTING IN A HIGH-LOW PRICING ENVIRONMENT | BREITER A, 2015, PROD OPER MANAG, V 24, P 219, DOI 10.1111/POMS.12250 |
2013	5	THE FUTURE OF UTILITY CUSTOMER-FUNDED ENERGY EFFICIENCY PROGRAMS IN THE USA: PROJECTED SPENDING AND SAVINGS TO 2025 | BARBOSE GL, 2013, ENERG EFFIC, V 6, P 475, DOI 10.1007/S12053-012-9187-1 |
2016	5	CONSUMERS PRE-LAUNCH AWARENESS AND PREFERENCE ON MOVIE SALES | MOON S, 2016, EUR J MARKETING, V 50, P1024, DOI 10.1108/EJM-06-2015-0324 |
2015	4	AN EMPIRICAL INVESTIGATION OF DYNAMIC ORDERING POLICIES | LARSON CR, 2015, MANAGE SCI, V 61, P 2118, DOI 10.1287/MNSC.2014.2077 |
2019	4	LONG-TERM LOAD FORECASTING: MODELS BASED ON MARS, ANN AND LR METHODS | NALCACI G, 2019, CENT EUR J OPER RES, V 27, P1033, DOI 10.1007/S10100-018-0531-1 |
2015	4	APPLICATION OF NEURAL NETWORKS TECHNIQUE IN PREDICTING IMPULSE BUYING AMONG SHOPPERS IN INDIA | PRASHAR S, 2015, DECISION, V 42, P 403, DOI 10.1007/S40622-015-0109-X |
2018	3	RELATING PRODUCT PRICES TO LONG-RUN MARGINAL COST: EVIDENCE FROM SOLAR PHOTOVOLTAIC MODULES | Bronze |
2018	3	COMPARING PREDICTION MARKET MECHANISMS: AN EXPERIMENT-BASED AND MICRO VALIDATED MULTI-AGENT SIMULATION | KLINGERT FM, 2018, JASSS-J ARTIF SOC S, V 21, DOI 10.18564/JASSS.3577 |
2014	3	IMPACT OF RESELLERS AND SALES AGENTS FORECASTING ACCURACY IN A MULTILAYER SUPPLY CHAIN | KUNG L, 2014, NAV RES LOG, V 61, P 207, DOI 10.1002/NAV.21578 |
2019	3	THE FORECAST OF COAL SALES TAKING THE FACTORS INFLUENCING THE DEMAND FOR HARD COAL INTO ACCOUNT | RYBAK A, 2019, GOSPOD SUROWCAMI MIN, V 35, P 129, DOI 10.24425/GSM.2019.128203 |

YearPublished	TimesCited	Article
2019	2	FORECASTING PROMOTIONAL SALES WITHIN THE NEIGHBOURHOOD | AGUILAR-PALACIOS C, 2019, IEEE ACCESS, V 7, P74759, DOI 10.1109/ACCESS.2019.2920380 |
2017	2	A HIGH INCIDENCE OF PYTHIUM AND PHYTOPHTHORA DISEASES RELATED TO RECORD-BREAKING RAINFALL IN SOUTH FLORIDA | CAMPOVERDE EV, 2017, HORTTECHNOLOGY, V 27, P 78, DOI 10.21273/HORTTECH03514-16 |
2019	2	FAST MOVING CONSUMER GOODS (FMCG) MARKET IN ROMANIA FEATURES AND TRENDS | STANCIU S, 2019, AMFITEATRU ECON, V 21, P 778, DOI 10.24818/EA/2019/S13/778 |
2018	1	OPTIMAL PRODUCTION PLANNING UTILIZING LEFTOVERS FOR AN ALL-YOU-CARE-TO EAT FOOD SERVICE OPERATION | BIRISCI E, 2018, J CLEAN PROD, V171, P 984, DOI 10.1016/J.JCLEPRO.2017.10.052 |
2019	1	THE IMPACT OF RECEIVING SMS PRICE AND WEATHER INFORMATION ON SMALL SCALE FARMERS IN COLOMBIA | CAMACHO A, 2019, WORLD DEV, V123, DOI 10.1016/J.WORLDDEV.2019.06.020 |
2019	1	A SIAMESE NEURAL NETWORK APPLICATION FOR SALES FORECASTING OF NEW FASHION PRODUCTS USING HETEROGENEOUS DATA | CRAPAROTTA G, 2019, INT J COMPUT INT SYS, V 12, P1537, DOI 10.2991/IJCIS.D.191122.002 |
2019	1	COORDINATING JUDGMENTAL FORECASTING: COPING WITH INTENTIONAL BIASES | PENNINGS CL, 2019, OMEGA-INT J MANAGE S, V 87, P 46, DOI 10.1016/J.OMEGA.2018.08.007 |
2020	0	HETEROGENEOUS BELIEFS AND RETURN VOLATILITY AROUND SEASONED EQUITY OFFERINGS | HIBBERT AM, 2020, J FINANC ECON, V137, P 571, DOI 10.1016/J.JFINECO.2020.03.003 |
2020	0	COST OPTIMIZATION MODEL FOR ITEMS HAVING FUZZY DEMAND AND DETERIORATION WITH TWO-WAREHOUSE FACILITY UNDER THE TRADE CREDIT FINANCING | KUMAR BA, 2020, AIMS MATH, V 5, P1603, DOI 10.3934/MATH.2020109 |
2014	0	ANALYSIS ON LOCATION SELECTION FOR RETAILERS BASED ON GIS | LI Y, 2014, ADV EDUC RES, V 68, P 115 |
2019	0	DECOMPOSITION AND FORECASTING TIME SERIES IN BUSINESS ECONOMY USING PROPHET FORECASTING MODEL | Other Gold |
2019	0	DEMAND FORECASTING AT RETAIL STAGE FOR SELECTED VEGETABLES: A PERFORMANCE ANALYSIS | PRIYADARSHI R, 2019, J MODEL MANAG, V 14, P 1042, DOI 10.1108/JM2-11-2018-0192 |
2016	0	INVESTIGATION OF FACTORS INFLUENCING EMPLOYEE PERFORMANCE A CASE OF SALES FORECASTING | SINDELAR J, 2016, INT J ORGAN ANAL, V 24, P 340, DOI 10.1108/IJOA-07-2013-0687 |
2019	0	FORECASTING REALIZED VOLATILITY DYNAMICALLY BASED ON ADJUSTED DYNAMIC MODEL AVERAGING (AMDA) APPROACH: EVIDENCE FROM CHINAS STOCK MARKET | YUAN P, 2019, ANN FINANC ECON, V 14, DOI 10.1142/S2010495219500222 |

Source: Author’s elaboration

1.9 References

1. Abell, D. F. (1978). Strategic Windows: The time to invest in a product or market is when a 'strategic window'is open. *Journal of marketing*, *42*(3), 21-26. https://doi.org/10.1177/002224297804200307
2. Achrol, R. S., and Stern, L. W. (1988). Environmental determinants of decision-making uncertainty in marketing channels. *Journal of marketing research*, *25*(1), 36-50. https://doi.org/10.1177/002224378802500104
3. Achrol, R. S., Reve, T., and Stern, L. W. (1983). The environment of marketing channel dyads: a framework for comparative analysis. *Journal of marketing*, *47*(4), 55-67. https://doi.org/10.1177/002224298304700407
4. Achrol, R.S. (1991). Evolution of the marketing organization: new forms for turbulent environments. *Journal of Marketing*, 55(October), 77-93. https://doi.org/10.1177/002224299105500406
5. Ackoff, R. L. (1981). *Creating the corporate future: Plan or be planned for*. University of Texas Press
6. Aguilar, F. J. (1967). *Scanning the business environment*. New York: Macmillan.
7. Aguilar-Palacios, C., Muñoz-Romero, S., and Rojo-Álvarez, J. L. (2019). Forecasting Promotional Sales Within the Neighbourhood. *IEEE Access*, *7*, 74759-74775. 10.1109/ACCESS.2019.2920380
8. Alamoudi, D., and Kumar, A. (2017). Information System Complexity and Business Value. *Int J Econ Manag Sci*, *6*(400), 2. https://doi.org/10.4172/2162-6359.1000400
9. Albarune, A. R. B., and Habib, M. M. (2015). A study of forecasting practices in supply chain management. *International Journal of Supply Chain Management*, *4*(2), 55-61.
10. Alexander, A., Kumar, M., and Walker, H. (2018). A decision theory perspective on complexity in performance measurement and management. *International Journal of Operations and Production Management*. https://doi.org/10.1108/IJOPM-10-2016-0632
11. Aldrich HE. 1979. *Organizations and Environments*. Englewood Cliffs, NJ: Prentice-Hall.
12. Amoako-Gyampah, K., and Boye, S. S. (2001). Operations strategy in an emerging economy: the case of the Ghanaian manufacturing industry. *Journal of Operations Management*, *19*(1), 59-79. https://doi.org/10.1016/S0272-6963(00)00046-2
13. Anand, G., and Ward, P. T. (2004). Fit, flexibility and performance in manufacturing: coping with dynamic environments. *Production and Operations Management*, *13*(4), 369-385. https://doi.org/10.1111/j.1937-5956.2004.tb00224.x
14. Ansoff, I. (1981). *Strategic Management*, Halsted Press. New York, NY.
15. Ansoff, I., and McDonnell, E. (1990). Implanting Strategic Management, Prentice Hall. *New York*.
16. Arndt, J. (1983). The political economy paradigm: foundation for theory building in marketing. *Journal of marketing*, *47*(4), 44-54. https://doi.org/10.1177/002224298304700406
17. Arunraj, N. S. and Ahrens, D. (2016). Estimation of non-catastrophic weather impacts for retail industry. *International Journal of Retail and Distribution Management*, *44*(7), 731-753. https://doi.org/10.1108/IJRDM-07-2015-0101
18. Ashill, N. J., and Jobber, D. (2014). The effects of the external environment on marketing decision-maker uncertainty. *Journal of Marketing Management*, *30*(3-4), 268-294. https://doi.org/10.1080/0267257X.2013.811281
19. Ashmos, D. P., Duchon, D., McDaniel, Jr, R. R., and Huonker, J. W. (2002). What a mess! Participation as a simple managerial rule to 'complexify'organizations. *Journal of Management studies*, *39*(2), 189-206. https://doi.org/10.1111/1467-6486.00288
20. Bahng, Y., and Kincade, D. H. (2012). The relationship between temperature and sales. *International Journal of Retail and Distribution Management*. https://doi.org/10.1108/09590551211230232
21. Banker, R.D., Lee, S.-Y. Potter, G. and Srinivasan, D. (1996). Contextual analysis of performance impacts of outcome-based incentive compensation. *Academy of Management Journal*, 39(4), 920-948. https://doi.org/10.5465/256717
22. Barile, S., Bassano, C., Calabrese, M., Confetto, M. G., Di Nauta, P., Piciocchi, P., ... and Vollero, A. (2011). *Contributions to theoretical and practical advances in management. A Viable Systems Approach (VSA)* (pp. 1-288). International Printing.
23. Barry, E. J., Kemerer, C. F., and Slaughter, S. A. (2006). Environmental volatility, development decisions, and software volatility: a longitudinal analysis. *Management Science*, *52*(3), 448-464.

https://doi.org/10.1287/mnsc.1050.0463
24. Basker, E. (2007). The causes and consequences of Wal-Mart's growth. *Journal of Economic Perspectives*, *21*(3), 177-198. https://doi.org/10.1257/jep.21.3.177
25. Battini, D., Persona, A., and Allesina, S. (2007). Towards a use of network analysis: quantifying the complexity of Supply Chain Networks. *International Journal of Electronic Customer Relationship Management*, *1*(1), 75-90. https://doi.org/10.1504/IJECRM.2007.014427
26. Begun, J.W., Zimmerman, B., Dooley, K. (2003). Health care organizations as complex adaptive systems. *Advances in health care organization theory*, 1, 253-288.
27. Berry, L. L. (1995). Relationship marketing of services—growing interest, emerging perspectives. *Journal of the Academy of marketing science*, *23*(4), 236-245. https://doi.org/10.1177/009207039502300402
28. Bertrand, J. L., and Parnaudeau, M. (2017). No more blaming the weather: a retailer’s approach to measuring and managing weather variability. *International Journal of Retail and Distribution Management*. https://doi.org/10.1108/IJRDM-09-2016-0152
29. Bititci, U., Garengo, P., Dörfler, V., and Nudurupati, S. (2012). Performance measurement: challenges for tomorrow. *International journal of management reviews*, *14*(3), 305-327. https://doi.org/10.1111/j.1468-2370.2011.00318.x
30. Black, J., & Farias, G. (1997, July). Genesis of complexity cycles. In *Eighth Annual International Conference of The Society for Chaos Theory in Psychology & Life Sciences, Boston University, Boston* (Vol. 31).
31. Blackman, T. (2013). Rethinking policy-related research: Charting a path using qualitative comparative analysis and complexity theory. *Contemporary Social Science*, *8*(3), 333-345. https://doi.org/10.1080/21582041.2012.751500
32. Boisot, M. (2000). Is there a complexity beyond the reach of strategy?. *Emergence*, *2*(1), 114-134. https://doi.org/10.1207/S15327000EM0201_08
33. Boulton, J. G., Allen, P. M., & Bowman, C. (2015). *Embracing complexity: Strategic perspectives for an age of turbulence*. OUP Oxford.
34. Brondoni, S. M. (2012). Innovation and Imitation: Corporate Strategies for Global Competition. *Symphonya. Emerging Issues in Management*, (1), 10-24. http://dx.doi.org/10.4468/2012.1.02brondoni
35. Brown, S. (1988). The wheel of the wheel of retailing. *International Journal of Retailing*, *3*(1), 16-37.
36. Brown, S. (1991). Variations on a marketing enigma: the wheel of retailing theory. *Journal of Marketing Management*, *7*(2), 131-155. https:/doi.org/10.1080/0267257X.1991.9964146
37. Brown, S. L., and Eisenhardt, K. M. (1997). The art of continuous change: Linking complexity theory and time-paced evolution in relentlessly shifting organizations. *Administrative science quarterly*, 1-34. https://doi.org/10.2307/2393807
38. Bruni, R., Carrubbo, L., Cavacece, Y., Di Muro, M. (2016). Market Dynamics: A Complex Adaptive System View. P*roceedingd Reser Conference*, Naples.
39. Buchko, A. A. (1994). Conceptualization and measurement of environmental uncertainty: An assessment of the Miles and Snow perceived environmental uncertainty scale. *Academy of Management Journal*, *37*(2), 410-425. https://doi.org/10.5465/256836
40. Burt, S., Dawson, J., and Sparks, L. (2003). Failure in international retailing: research propositions. *The International Review of Retail, Distribution and Consumer Research*, *13*(4), 355-373. https://doi.org/10.1080/0959396032000129471
41. Camacho, A., and Conover, E. (2019). The impact of receiving SMS price and weather information on small scale farmers in Colombia. *World Development*, *123*, 104596. https://doi.org/10.1016/j.worlddev.2019.06.020
42. Carrière-Swallow Y. and Labbé F. (2010). Nowcasting With Google Trends in an Emerging Market. *Working Papers Central Bank of Chile* 588, Central Bank of Chile.
43. Casti, J. (1997), *Would-Be Worlds*. John Wiley and Sons, New York, NY.
44. Chae, M. and Hill, J.S. (1997). High versus low formality marketing planning in global industries: determinants and consequences. *Journal of Strategic Marketing*, 5(3), 22. https://doi.org/10.1080/096525497346848
45. Chakravarthy, B. (1997). A new strategy framework for coping with turbulence. *Sloan Management Review*, (Winter), 69-82.
46. Chase Jr, C. W. (2013), *Demand-driven forecasting: a structured approach to forecasting*. John Wiley and Sons.
47. Chase, C. W. (2016). *Next generation demand management: People, process, analytics, and technology*. John Wiley and Sons.
48. Chen, J. C., Cheng, C. H., and Huang, P. B. (2013). Supply chain management with lean production and RFID

application: A case study. *Expert Systems with applications*, *40*(9), 3389-3397. https://doi.org/10.1016/j.eswa.2012.12.047
49. Child, J. (1972). Organizational structure, environment and performance: The role of strategic choice. *sociology*, *6*(1), 1-22. https://doi.org/10.1177%2F003803857200600101
50. Child, J. (1975). Managerial and organizational factors associated with company performance. Part II. A contingency analysis. *Journal of Management Studies*, 12(1-2), 12-27. https://doi.org/10.1111/j.1467-6486.1975.tb00884.x
51. Choe, J. M. (2003). The effect of environmental uncertainty and strategic applications of IS on a firm's performance. *Information and Management*, *40*(4), 257-268. https://doi.org/10.1016/S0378-7206(02)00008-3
52. Choi, T. M., Hui, C. L., Liu, N., Ng, S. F., and Yu, Y. (2014). Fast fashion sales forecasting with limited data and time. *Decision Support Systems*, *59*, 84-92.https://doi.org/10.1016/j.dss.2013.10.008
53. Choi, T. Y., Dooley, K. J., and Rungtusanatham, M. (2001). Supply networks and complex adaptive systems: control versus emergence. *Journal of operations management*, *19*(3), 351-366. https://doi.org/10.1016/S0272-6963(00)00068-1
54. Christopher, M., Lowson, R., and Peck, H. (2004). Creating agile supply chains in the fashion industry. *International Journal of Retail and Distribution Management*. 32 (8), pp. 367-376. https://doi.org/10.1108/09590550410546188
55. Cilliers, P. (2000). Rules and complex systems. *Emergence, A Journal of Complexity Issues in Organizations and Management*, *2*(3), 40-50. https://doi.org/10.1207/S15327000EM0203_04
56. Cilliers, P. (2005). Knowing complex systems. *Managing Organisational Complexity: Philosophy, Theory, Application, IAP, Connecticut*, 7-19.
57. Conner, D.R. (1998), *Leading at the Edge of Chaos: How to Create the Nimble Organization*, John Wiley, New York, NY.
58. Cooper, R. G. (1979). The dimensions of industrial new product success and failure. *Journal of marketing*, *43*(3), 93-103. https://doi.org/10.1177/002224297904300310
59. Covin, J. G., and Slevin, D. P. (1989). Strategic management of small firms in hostile and benign environments. *Strategic management journal*, *10*(1), 75-87. https://doi.org/10.1002/smj.4250100107
60. Cravens, D.W. (1991), *Strategic Marketing*, 3rd ed., Irwin, Homewood, IL.
61. Daft, R. L., Sormunen, J., and Parks, D. (1988). Chief executive scanning, environmental characteristics, and company performance: An empirical study. *Strategic management journal*, *9*(2), 123-139. https://doi.org/10.1002/smj.4250090204
62. D'Aveni, R. A., Dagnino, G. B., and Smith, K. G. (2010). The age of temporary advantage. *Strategic management journal*, *31*(13), 1371-1385. https://doi.org/10.1002/smj.897
63. Dawson, J. (2002). 11 Towards a model of the impacts of retail internationalisation. *The internationalisation of retailing in Asia*, 189.
64. DeNooy, W., Mrvar, A., Batagelj, V., Mehri, D., Auto, I., Lee, C. S., ... and Voters, M. P. (2005). Structural analysis in the social sciences. *Cambridge: Cambridge University Press.*
65. Dess, G. G., and Beard, D. W. (1984). Dimensions of organizational task environments. *Administrative science quarterly*, 52-73. https://doi.org/10.2307/2393080
66. Devereux, L., Melewar, T. C., Dinnie, K., and Lange, T. (2020). Corporate identity orientation and disorientation: A complexity theory perspective. *Journal of Business Research*, *109*, 413-424. https://doi.org/10.1016/j.jbusres.2019.09.048
67. Dibb, S. (1996). The impact of the changing marketing environment in the Pacific Rim: four case studies. *International Journal of Retail and Distribution Management*. https://doi.org/10.1108/09590559610131691
68. Dibb, S., and Simkin, L. (1994). Implementation problems in industrial market segmentation. *Industrial Marketing Management*, *23*(1), 55-63. https://doi.org/10.1016/0019-8501(94)90027-2
69. Dill, W. R. (1958). Environment as an influence on managerial autonomy. *Administrative science quarterly*, 409-443. https://doi.org/10.2307/2390794
70. Ding, Y., Chowdhury, G. G., & Foo, S. (2001). Bibliometric cartography of information retrieval research by using co-word analysis. *Information processing & management*, *37*(6), 817-842. https://doi.org/10.1016/S0306-4573(00)00051-0
71. Douglas, C. A., and McCauley, C. D. (1999). Formal developmental relationships: A survey of organizational practices. *Human Resource Development Quarterly*, *10*(3), 203-220. https://doi.org/10.1002/hrdq.3920100302

72. Drucker, P. F. (1993). *Managing in turbulent times*. Routledge.
73. Duncan, R. B. (1972). Characteristics of organizational environments and perceived environmental uncertainty. *Administrative science quarterly*, 313-327. https://doi.org/10.2307/2392145
74. Dwyer, F. R., and Welsh, M. A. (1985). Environmental relationships of the internal political economy of marketing channels. *Journal of Marketing Research*, *22*(4), 397-414. https://doi.org/10.1177/002224378502200405
75. Eisenhardt, K. M. (1989). Making fast strategic decisions in high-velocity environments. *Academy of Management journal*, *32*(3), 543-576. https://doi.org/10.5465/256434
76. Ekinci, E., and Baykasoglu, A. (2016). Modelling complexity in retail supply chains. Kybernetes. https://doi.org/10.1108/K-12-2014-0307
77. Ekinci, E., and Baykasoğlu, A. (2019). Complexity and performance measurement for retail supply chains. *Industrial Management and Data Systems*. 119 (4), pp. 719-742. https://doi.org/10.1108/IMDS-08-2018-0342
78. Eksoz C., Mansouri S. A. and Bourlakis M. (2014). Collaborative forecasting in the food supply chain: conceptual framework. *International Journal of Production Economics*, 158, 120-135. https://doi.org/10.1016/j.ijpe.2014.07.031
79. Emery, F. E., and Trist, E. L. (1965). The causal texture of organizational environments. *Human relations*, *18*(1), 21-32. https://doi.org/10.1177/001872676501800103
80. Espejo, R. (1993). Domains of interaction between a social system and its environment. *Systems practice*, *6*(5), 517-525. https://doi.org/10.1007/BF01059509
81. Etgar,M. (1984) 'The retail ecologymodel: a comprehensivemodel of retail change', in J. Sheth (ed.) *Research in Marketing*,Vol. 7, Greenwich, CT: JAI Press, pp. 41–62.
82. Everingham, Y., Sexton, J., Skocaj, D., and Inman-Bamber, G. (2016). Accurate prediction of sugarcane yield using a random forest algorithm. *Agronomy for sustainable development*, *36*(2), 27. https://doi.org/10.1007/s13593-016-0364-z
83. Fildes, R., Ma, S., and Kolassa, S. (2019). Retail forecasting: Research and practice. *International Journal of Forecasting*. https://doi.org/10.1016/j.ijforecast.2019.06.004
84. French, S., Maule, J., and Papamichail, N. (2009). *Decision behaviour, analysis and support*. Cambridge University Press.
85. Frizelle, G., and Woodcock, E. (1995). Measuring complexity as an aid to developing operational strategy. *International Journal of Operations and Production Management*. https://doi.org/10.1108/01443579510083640
86. Galbraith, J. (1973). Designing complex organizations. *Reading, Mass*, 15(2).
87. Gerloff, E. A., Muir, N. K., and Bodensteiner, W. D. (1991). Three components of perceived environmental uncertainty: An exploratory analysis of the effects of aggregation. *Journal of Management*, *17*(4), 749-768. https://doi.org/10.1177/014920639101700408
88. Gorzeń-Mitka, I., & Okręglicka, M. (2014). Improving decision making in complexity environment. *Procedia Economics and Finance*, *16*, 402-409. https://doi.org/10.1016/S2212-5671(14)00819-3
89. Grewal, R., and Tansuhaj, P. (2001). Building organizational capabilities for managing economic crisis: The role of market orientation and strategic flexibility. *Journal of marketing*, *65*(2), 67-80. https://doi.org/10.1509/jmkg.65.2.67.18259
90. Hambrick, D. C. (1982). Environmental scanning and organizational strategy. *Strategic management journal*, *3*(2), 159-174. https://doi.org/10.1002/smj.4250030207
91. Hamel, G. and Prahalad, C.K. (1994), *Competing for the Future*, Harvard Business School Press, Boston, MA.
92. Harrauer V. and Schnedlitz P. (2016). Impact of environment on performance measurement design and processing in retailing. *International Journal of Retail and Distribution Management*, 44 (3), 320 – 335. https://doi.org/10.1108/IJRDM-03-2015-0048
93. Harrington, R. (2001). Environmental uncertainty within the hospitality industry: Exploring the measure of dynamism and complexity between restaurant segments. *Journal of Hospitality and Tourism Research*, *25*(4), 386-398. https://doi.org/10.1177/109634800102500403
94. Hausman, J., and Leibtag, E. (2007). Consumer benefits from increased competition in shopping outlets: Measuring the effect of Wal-Mart. *Journal of Applied Econometrics*, *22*(7), 1157-1177. https://doi.org/10.1002/jae.994

95. Hayes, R. H., and Abernathy, W. J. (1980). Managing our way to economic decline. *Harvard Business Review. 58*(4).
96. He, Z., Wang, S., and Cheng, T. C. E. (2013). Competition and evolution in multi-product supply chains: An agent-based retailer model. *International Journal of Production Economics*. *146*(1), pp. 325-336. https://doi.org/10.1016/j.ijpe.2013.07.019
97. Herbig, P. A. (1990). Marketing chaos–when randomness can be deterministic. *Journal of International Marketing and Market Research*, *16*(2), 65-84.
98. Hickson, D. J., Hinings, C. R., Lee, C. A., Schneck, R. E., and Pennings, J. M. (1971). A strategic contingencies' theory of intraorganizational power. *Administrative science quarterly*, 216-229. https://doi.org/10.2307/2391831
99. Hitt, M. A., Ireland, R. D., and Palia, K. A. (1982). Industrial firms' grand strategy and functional importance: Moderating effects of technology and uncertainty. *Academy of Management Journal*, *25*(2), 265-298. https://doi.org/10.5465/255990
100. Hollander, S. C. (1960). The wheel of retailing. *Journal of Marketing*, *25*(1), 37-42. https://doi.org/10.1177/002224296002500106
101. Huang, T., Fildes, R., and Soopramanien, D. (2014). The value of competitive information in forecasting FMCG retail product sales and the variable selection problem. *European Journal of Operational Research*, *237*(2), 738-748. https://doi.org/10.1016/j.ejor.2014.02.022
102. Huber, G. P. (1984). The nature and design of post-industrial organizations. *Management science*, *30*(8), 928-951. https://doi.org/10.1287/mnsc.30.8.928
103. Hübner, A. H., Kuhn, H. and Sternbeck, M. G. (2013*).* Demand and supply chain planning in grocery retail: an operations planning framework. *International Journal of Retail and Distribution Management*, *41*(7), 512-530. https://doi.org/10.1108/IJRDM-05-2013-0104
104. Hummelbrunner, R., and Jones, H. (2013). A guide to managing in the face of complexity. *London: ODI*.
105. Hwang, E. J. (2005). *Strategic management and financial performance in South Korean apparel retail stores* (Doctoral dissertation, Virginia Tech). http://hdl.handle.net/10919/29855
106. Jia, P. (2008). What happens when Wal-Mart comes to town: An empirical analysis of the discount retailing industry. *Econometrica*, *76*(6), 1263-1316. https://doi.org/10.3982/ECTA6649
107. Kauffman, J. M. (1995). Why we must celebrate a diversity of restrictive environments. *Learning Disabilities Research and Practice*.
108. Keegan, D. P., Eiler, R. G., and Jones, C. R. (1989). Are your performance measures obsolete?. *Strategic Finance*, *70*(12), 45.
109. Khandwalla, P. (1972). Environment and its impact on the organization. *International Studies of Management and Organization*, 2 (3), pp. 297-313. https://doi.org/10.1080/00208825.1972.11656125
110. Khandwalla, P.N. (1977), *The Design of Organizations*, Harcourt Brace Jovanovich, New York, NY.
111. Kim, J., and Burton, R. M. (2002). The effect of task uncertainty and decentralization on project team performance. *Computational and Mathematical Organization Theory*, *8*(4), 365-384. https://doi.org/10.1023/A:1025472702927
112. Knutas, A., Hajikhani, A., Salminen, J., Ikonen, J. and Porras, J. (2015, June). Cloud-based bibliometric analysis service for systematic mapping studies. *Proceedings of the 16th International Conference on Computer Systems and Technologies* (pp. 184-191). ACM- http://hammer.nailsproject.net/
113. Koberg, C. S., and Ungson, G. R. (1987). The effects of environmental uncertainty and dependence on organizational structure and performance: A comparative study. *Journal of Management*, *13*(4), 725-737. https://doi.org/10.1177/014920638701300412
114. Kohli, A. K., and Jaworski, B. J. (1990). Market orientation: the construct, research propositions, and managerial implications. *Journal of marketing*, *54*(2), 1-18. https://doi.org/10.1177/002224299005400201
115. Kotter, J.P. (1996). Kill complacency. *Fortune*, 5(August), 122-124.
116. Kumar, V., and Karande, K. (2000). The effect of retail store environment on retailer performance. *Journal of business research*, *49*(2), 167-181.https://doi.org/10.1016/S0148-2963(99)00005-3
117. Kung, L. C., and Chen, Y. J. (2014). Impact of reseller's and sales agent's forecasting accuracy in a multilayer supply chain. *Naval Research Logistics (NRL)*, *61*(3), 207-222. https://doi.org/10.1002/nav.21578
118. Kurtz, C. F., and Snowden, D. J. (2003). The new dynamics of strategy: Sense-making in a complex and complicated world. *IBM systems journal*, *42*(3), 462-483. https://doi.org/10.1147/sj.423.0462
119. Lane, D., and Maxfield, R. (1996). Strategy under complexity: Fostering generative relationships. *Long range planning*, *29*(2), 215-231. https://doi.org/10.1016/0024-6301(96)00011-8

120. Lawrence, P. R., and Lorsch, J. W. (1967). Differentiation and integration in complex organizations. *Administrative science quarterly*, 1-47. https://doi.org/10.2307/2391211
121. Lenz, R. T., and Engledow, J. L. (1986). Environmental analysis units and strategic decision-making: A field study of selected 'leading-edge'corporations. *Strategic Management Journal*, *7*(1), 69-89. https://doi.org/10.1002/smj.4250070107
122. Lenz, R. T., and Engledow, J. L. (1986). Environmental analysis: The applicability of current theory. *Strategic Management Journal*, *7*(4), 329-346. https://doi.org/10.1002/smj.4250070404
123. Levy, D. (1994). Chaos theory and strategy: Theory, application, and managerial implications. *Strategic management journal*, *15*(S2), 167-178. https://doi.org/10.1002/smj.4250151011
124. Liao, Z., and Long, S. (2019). Can interfirm trust improve firms' cooperation on environmental innovation? The moderating role of environmental hostility. *Business Strategy and the Environment*, *28*(1), pp. 198-205. https://doi.org/10.1002/bse.2249
125. Lightfoot, W. (2003). Multi-channel mistake: the demise of a successful retailer. *International Journal of Retail and Distribution Management*. 31 (4), pp. 220-229. https://doi.org/10.1108/09590550310469211
126. Linstone, H. A. (1999). Complexity science: implications for forecasting. *Technological Forecasting and Social Change*, *62*(1-2), 79-90. https://doi.org/10.1016/S0040-1625(99)00034-7
127. Lo, H. K., and Yip, C. W. (2001). Fare deregulation of transit services: winners and losers in a competitive market. *Journal of advanced transportation*, *35*(3), 215-235. https://doi.org/10.1002/atr.5670350304
128. Loewen, J. (1997), *The Power of Strategy: A Practical Guide for South African Managers*, Zebra: Sandton.
129. Lowson, R. H. (2005). Retail operations strategies: empirical evidence of role, competitive contribution and life cycle. *International Journal of Operations and Production Management*, *25*(7), 642-680. https://doi.org/10.1108/01443570510605081
130. Ma, S., Fildes, R., and Huang, T. (2016). Demand forecasting with high dimensional data: The case of SKU retail sales forecasting with intra-and inter-category promotional information. *European Journal of Operational Research*, *249*(1), 245-257. https://doi.org/10.1016/j.ejor.2015.08.029
131. Mai, E. S., Yang, J., and Chen, H. (2011). Primary product network size on complementary product sales. *International Journal of Retail and Distribution Management*. 39 (11), pp. 851-866. https://doi.org/10.1108/09590551111177963
132. Mason, J. B. (1987). Morris l. *Mayer. Modern retailing, theory and practice. Fifth edition. 725pp*.
133. Mason, R.B., (2007). The external environment's effect on management and strategy. A complexity theory approach. *Management Decision,* 45(1), 10-28. https://doi.org/10.1108/00251740710718935
134. McCabe, D. L. (1990). The assessment of perceived environmental uncertainty and economic performance. *Human Relations*, *43*(12), 1203-1218. https://doi.org/10.1177%2F001872679004301202
135. McGoldrick, P. J. (2002). *Retail marketing*. McGraw-Hill.
136. McNair, M. P. (1958). *Significant trends and developments in the postwar period* (pp. 1-25). Pittsburgh, PA: University of Pittsburgh Press.
137. Mersereau, Adam J. Demand estimation from censored observations with inventory record inaccuracy. *Manufacturing and Service Operations Management* 17, no. 3 (2015): 335-349. https://doi.org/10.1287/msom.2015.0520
138. Miles, R. E., Snow, C. C., Meyer, A. D., and Coleman Jr, H. J. (1978). Organizational strategy, structure, and process. *Academy of management review*, *3*(3), 546-562. https://doi.org/10.5465/amr.1978.4305755
139. Miller, C. E., Reardon, J., and McCorkle, D. E. (1999). The effects of competition on retail structure: An examination of intratype, intertype, and intercategory competition. *Journal of Marketing*, *63*(4), 107-120. https://doi.org/10.1177/002224299906300409
140. Miller, D. (1987). The structural and environmental correlates of business strategy. *Strategic management journal*, *8*(1), 55-76. https://doi.org/10.1002/smj.4250080106
141. Miller, D., and Friesen, P. H. (1983). Strategy-making and environment: the third link. *Strategic management journal*, *4*(3), 221-235. https://doi.org/10.1002/smj.4250040304
142. Miller, D., and Friesen, P. H. (1984). A longitudinal study of the corporate life cycle. *Management science*, *30*(10), 1161-1183. https://doi.org/10.1287/mnsc.30.10.1161
143. Miller, K. D. (1993). Industry and country effects on managers' perceptions of environmental uncertainties. *Journal of International Business Studies*, *24*(4), 693-714. https://doi.org/10.1057/palgrave.jibs.8490251

144. Milliken, F. J. (1987). Three types of perceived uncertainty about the environment: State, effect, and response uncertainty. *Academy of Management review*, *12*(1), 133-143. https://doi.org/10.5465/amr.1987.4306502
145. Milliken, F. J. (1990). Perceiving and interpreting environmental change: An examination of college administrators' interpretation of changing demographics. *Academy of management Journal*, *33*(1), 42-63. https://doi.org/10.5465/256351
146. Mintz, O., and Currim, I. S. (2013). What drives managerial use of marketing and financial metrics and does metric use affect performance of marketing-mix activities?. *Journal of Marketing*, *77*(2), 17-40. https://doi.org/10.1509/jm.11.0463
147. Mintzberg, H., and Waters, J. A. (1985). Of strategies, deliberate and emergent. *Strategic management journal*, *6*(3), 257-272. https://doi.org/10.1002/smj.4250060306
148. Mintzberg, H., & Westley, F. (2001). It's not what you think. *MIT Sloan Management Review*, *42*(3), 89-93.
149. Mishra, B. K., and Raghunathan, S. (2004). Retailer-vs. vendor-managed inventory and brand competition. *Management Science*, *50*(4), 445-457. https://doi.org/10.1287/mnsc.1030.0174
150. Mitroff, I., Mitroff, I. I., Mohrman, S. A., and Little, G. (1987). *Business not as usual: Rethinking our individual, corporate, and industrial strategies for global competition*. Jossey-Bass.
151. Morel, B., and Ramanujam, R. (1999). Through the looking glass of complexity: The dynamics of organizations as adaptive and evolving systems. *Organization Science*, *10*(3), 278-293. https://doi.org/10.1287/orsc.10.3.278
152. Moretta Tartaglione A., Bruni R., (2016). International networks as complex adaptive systems. *Mercati and Competitività*, 3(2016), 111-131. https://doi.org/10.3280/MC2016-003007
153. Moretta Tartaglione, A., Bruni, R., Bozic, M., and Cavacece, Y. (2018). How the environmental complexity affects forecasting in retail? Actual research trends. *Journal of Marketing Trends (1961-7798)*, *5*(2).
154. Moretta Tartaglione, A., Bruni, R., and Bozic, M. (2019). Exploring the retail industry environment using nonlinear analysis. *International Journal of Retail and Distribution Management*, Vol. 47 No. 4, pp. 453-47. https://doi.org/10.1108/IJRDM-09-2017-0224
155. Murtha, T. P., Lenway, S. A., and Bagozzi, R. P. (1998). Global mind-sets and cognitive shift in a complex multinational corporation. *Strategic Management Journal*, *19*(2), 97-114. https://doi.org/10.1002/(SICI)1097-0266(199802)19:2<97::AID-SMJ943>3.0.CO;2-2
156. Nair, A., and Reed-Tsochas, F. (2019). Revisiting the complex adaptive systems paradigm: Leading perspectives for researching operations and supply chain management issues. https://doi.org/10.1002/joom.1022
157. Nair, A., Yan, T., Ro, Y. K., Oke, A., Chiles, T. H., and Lee, S. Y. (2016). How environmental innovations emerge and proliferate in supply networks: A complex adaptive systems perspective. *Journal of Supply Chain Management*, *52*(2), 66-86. https://doi.org/10.1002/joom.1022
158. Neumark, D., Zhang, J., and Ciccarella, S. (2008). The effects of Wal-Mart on local labor markets. *Journal of Urban Economics*, *63*(2), 405-430. https://doi.org/10.1016/j.jue.2007.07.004
159. Nguimkeu, P. (2016). Some effects of business environment on retail firms. *Applied Economics*, *48*(18), pp. 1647-1654. https://doi.org/10.1080/00036846.2015.1105923
160. Oreja-Rodríguez, J. R., and Yanes-Estévez, V. (2010). Environmental scanning. *Management Decision.* 48 (2), pp. 260-276. https://doi.org/10.1108/00251741011022617
161. Papanagnou, C. I., and Matthews-Amune, O. (2018). Coping with demand volatility in retail pharmacies with the aid of big data exploration. *Computers and Operations Research*, *98*, 343-354. https://doi.org/10.1016/j.cor.2017.08.009
162. Pathak, S. D., Day, J. M., Nair, A., Sawaya, W. J., and Kristal, M. M. (2007). Complexity and adaptivity in supply networks: Building supply network theory using a complex adaptive systems perspective. *Decision sciences*, *38*(4), 547-580. https://doi.org/10.1111/j.1540-5915.2007.00170.x
163. Perianes-Rodriguez, A., Waltman, L., and Van Eck, N. J. (2016). Constructing bibliometric networks: A comparison between full and fractional counting. *Journal of Informetrics*, *10*(4), 1178-1195. https://doi.org/10.1016/j.joi.2016.10.006
164. Perrow, C. (2011). *Normal accidents: Living with high risk technologies-Updated edition*. Princeton university press.
165. Peters, E.E. (1999), *Patterns in the Dark: Understanding Risk and Financial Crisis with Complexity Theory*, Wiley, New York, NY.
166. Pioch, E. A., and Schmidt, R. A. (2000). Consumption and the retail change process: a comparative analysis of toy retailing in Italy and France. *The International Review of Retail, Distribution and Consumer Research*, *10*(2),

183-203. https://doi.org/10.1080/095939600342352
167. Polonsky, M.J., Suchard, H.T. and Scott, D.R. (1999). The incorporation of an interactive external environment: an extended model of marketing relationships. *Journal of Strategic Marketing*, 7,41-55. https://doi.org/10.1080/096525499346521
168. Porter, M. E. (1999). Michael Porter on competition. *Antitrust Bull.*, *44*, 841.
169. Porter, M. E. (2008). *On competition*. Harvard Business Press.
170. Prasad, S., Babbar, S., and Motwani, J. (2001). International operations strategy: current efforts and future directions. *International Journal of Operations and Production Management*. https://doi.org/10.1108/01443570110390381
171. Preble, J. F., Rau, P. A., and Reichel, A. (1988). The environmental scanning practices of US multinationals in the late 1980's. *Management International Review*, 4-14. https://www.jstor.org/stable/40227901
172. Prieto, M., and Caemmerer, B. (2013). An exploration of factors influencing car purchasing decisions. *International Journal of Retail and Distribution Management*. https://doi.org/10.1108/IJRDM-02-2012-0017
173. Rasch, G. (1993). *Probabilistic models for some intelligence and attainment tests*. MESA Press, 5835 S. Kimbark Ave., Chicago, IL 60637; e-mail: MESA@ uchicago. edu; web address: www. rasch. org; tele.
174. Reeves, M., Levin, S., Fink, T., and Levina, A. (2020). Taming Complexity. *Harvard Business Review*, *98*(1), 112-121.
175. Ren, S., Chan, H. L., & Ram, P. (2017). A comparative study on fashion demand forecasting models with multiple sources of uncertainty. *Annals of Operations Research*, *257*(1), 335-355. https://doi.org/10.1007/s10479-016-2204-6
176. Rueda-Manzanares, A., Aragón-Correa, J. A., and Sharma, S. (2008). The influence of stakeholders on the environmental strategy of service firms: The moderating effects of complexity, uncertainty and munificence. *British Journal of management*, *19*(2), 185-203. https://doi.org/10.1111/j.1467-8551.2007.00538.x
177. Salancik, G. R., and Pfeffer, J. (1978). A social information processing approach to job attitudes and task design. *Administrative science quarterly*, 224-253. https://doi.org/10.2307/2392563
178. Sampson, S. D., and Tiger, D. J. (1994). The impact of warehouse membership clubs: the wheel of retailing turns one more time. *The International Review of Retail, Distribution and Consumer Research*, *4*(1), 33-59. https://doi.org/10.1080/09593969400000003
179. Sarkar, S., and Searcy, C. (2016). Zeitgeist or chameleon? A quantitative analysis of CSR definitions. *Journal of Cleaner Production*, *135*, 1423-1435. https://doi.org/10.1016/j.jclepro.2016.06.157
180. Sawyerr, O. O., McGee, J., and Peterson, M. (2003). Perceived uncertainty and firm performance in SMEs: The role of personal networking activities. *International Small Business Journal*, *21*(3), 269-290. https://doi.org/10.1177/02662426030213002
181. Schneider, M., and Somers, M. (2006). Organizations as complex adaptive systems: Implications of complexity theory for leadership research. *The Leadership Quarterly*, *17*(4), 351-365. https://doi.org/10.1016/j.leaqua.2006.04.006
182. Shannon, C. E., and Weaver, W. (1949). *The mathematical theory of communication.* University of Illinois Press.
183. Sharfman, M. P., and Dean Jr, J. W. (1991). Conceptualizing and measuring the organizational environment: A multidimensional approach. *Journal of management*, *17*(4), 681-700. https://doi.org/10.1177/014920639101700403
184. Sillanpää, V., and Liesiö, J. (2018). Forecasting replenishment orders in retail: value of modelling low and intermittent consumer demand with distributions. *International Journal of Production Research*, *56*(12), 4168-4185. https://doi.org/10.1080/00207543.2018.1431413
185. Simon, H. (1957). Abehavioral model of rational choice. *Models of Man, Social and Rational: Mathematical Essays on Rational Human Behavior in a Social Setting*.
186. Simon, H. A. (1973). Applying information technology to organization design. *Public administration review*, *33*(3), 268-278. https://doi.org/10.2307/974804
187. Sivadasan, S., Efstathiou, J., Frizelle, G., Shirazi, R., and Calinescu, A. (2002). An information-theoretic methodology for measuring the operational complexity of supplier-customer systems. *International Journal of Operations and Production Management*. https://doi.org/10.1108/01443570210412088

188. Sivadasan, S., Efstathiou, J., Calinescu, A., and Huatuco, L. H. (2006). Advances on measuring the operational complexity of supplier–customer systems. *European Journal of Operational Research*, *171*(1), 208-226. https://doi.org/10.1016/j.ejor.2004.08.032
189. Smart, C., and Vertinsky, I. (1984). Strategy and the environment: A study of corporate responses to crises. *Strategic management journal*, *5*(3), 199-213. https://doi.org/10.1002/smj.4250050302
190. Smith, W., (1983). *Organizing as power process, the creation of a conceptual framework and its application to the design of development projects*. (Doctoral Dissertation). Philadelphia: University of Pennsylavania.
191. Snowden, D. J. (2000). New wine in old wineskins: From organic to complex knowledge management through the use of story. *Emergence, A Journal of Complexity Issues in Organizations and Management*, *2*(4), 50-64. https://doi.org/10.1207/S15327000EM0204_07
192. Snowden, D. J. (2002). Archetypes as an instrument of narrative patterning. *Info Today*, 209-216.
193. Snowden, D. J., and Boone, M. E. (2007). A leader's framework for decision making. *Harvard business review*, *85*(11), 68.
194. Spiegler, V. L. (2013). *Designing supply chains resilient to nonlinear system dynamics* (Doctoral dissertation, Cardiff University).
195. Stacey, R. D. (1992). *Managing the unknowable: Strategic boundaries between order and chaos in organizations*. John Wiley and Sons.
196. Stacey, R. (1993). Strategy as order emerging from chaos. *Long range planning*, *26*(1), 10-17. https://doi.org/10.1016/0024-6301(93)90228-8
197. Stacey, R. D. (2011). Strategic Management and Organization Dynamics: The Challenge of Complexity. 6 p. Harlow. *Financial Times/Prentice Hall*.
198. Stern, L. W., and Reve, T. (1980). Distribution channels as political economies: a framework for comparative analysis. *Journal of marketing*, *44*(3), 52-64. https://doi.org/10.1177/002224298004400306
199. Stewart Jr, W. H., May, R. C., and Kalia, A. (2008). Environmental perceptions and scanning in the United States and India: convergence in entrepreneurial information seeking?. *Entrepreneurship Theory and Practice*, *32*(1), 83-106. https://doi.org/10.1111/j.1540-6520.2007.00217.x
200. Surana, A., Kumara*, S., Greaves, M., and Raghavan, U. N. (2005). Supply-chain networks: a complex adaptive systems perspective. *International Journal of Production Research*, *43*(20), 4235-4265. https://doi.org/10.1080/00207540500142274
201. Sutcliffe, K. M., and Huber, G. P. (1998). Firm and industry as determinants of executive perceptions of the environment. *Strategic management journal*, *19*(8), 793-807. https://doi.org/10.1002/(SICI)1097-0266(199808)19:8<793::AID-SMJ980>3.0.CO;2-Y
202. Taylor, T. A., and Xiao, W. (2009). Incentives for retailer forecasting: Rebates vs. returns. *Management Science*, *55*(10), 1654-1669. https://doi.org/10.1287/mnsc.1090.1045
203. Thietart, R.A. and Forgues, B. (1995). Chaos theory and organization. *Organization Science*, 6(1), 19-31. https://doi.org/10.1287/orsc.6.1.19
204. Thompson, J. D. (1967). *Organizations in action: Social science bases of administrative theory*. New York: McGraw-Hill.
205. Towers, N., & Ashford, R. (2001). The supply chain management of production planning and sustainable customer relationships. *Management research news*. https://doi.org/10.1108/01409170110782423
206. Tsai, B. H. (2013). Modeling diffusion of multi-generational LCD TVs while considering generation-specific price effects and consumer behaviors. *Technovation*, *33*(10-11), 345-354. https://doi.org/10.1016/j.technovation.2013.05.002
207. Tung, R. L. (1979). Dimensions of organizational environments: An exploratory study of their impact on organization structure. *Academy of Management Journal*, *22*(4), 672-693. https://doi.org/10.5465/255808
208. Uncles, M. D. (2010). Retail change in China: retrospect and prospects. *The International Review of Retail, Distribution and Consumer Research*, *20*(1), 69-84. https://doi.org/10.1080/09593961003594501
209. Utterback, J. M., and Uhlmann, L. (1979). Product and process innovation in a changing competitive environment. In *Industrial Innovation* (pp. 135-150). Palgrave Macmillan, London. https://doi.org/10.1007/978-1-349-03822-0_7
210. Van Donselaar, K. H., Peters, J., de Jong, A., and Broekmeulen, R. A. (2016). Analysis and forecasting of demand during promotions for perishable items. *International Journal of Production Economics*, *172*, 65-75. https://doi.org/10.1016/j.ijpe.2015.10.022

211. Van Eck, N. J., and Waltman, L. (2010). Software survey: VOSviewer, a computer program for bibliometric mapping. *scientometrics*, *84*(2), 523-538. https://doi.org/10.1007/s11192-009-0146-3
212. Van Eck, N. J., and Waltman, L. (2014). Visualizing bibliometric networks. In *Measuring scholarly impact* (pp. 285-320). Springer, Cham. https://doi.org/10.1007/978-3-319-10377-8_13
213. Van Egeren, M., and O'Connor, S. (1998). Drivers of market orientation and performance in service firms. *Journal of Services Marketing*. https://doi.org/10.1108/08876049810202357
214. Van Rensburg, A. C. J. (2012). Can complexity analysis support business performance insight?. *South African Journal of Industrial Engineering*, *23*(2), 16-28. http://www.scielo.org.za/scielo.php?script=sci_arttextandpid=S2224-78902012000200004andlng=enandtlng=en
215. Vasconcelos, F. C., and Ramirez, R. (2011). Complexity in business environments. *Journal of Business Research*, *64*(3), 236-241. https://doi.org/10.1016/j.jbusres.2009.11.007
216. Verstraete, G., Aghezzaf, E. H., and Desmet, B. (2019). A data-driven framework for predicting weather impact on high-volume low-margin retail products. *Journal of Retailing and Consumer Services*, *48*, 169-177. https://doi.org/10.1016/j.jretconser.2019.02.019
217. Vicari, S. (1998). *La creatività dell'impresa. Tra caso e necessità*. Etas, Milano.
218. Vida, I., and Fairhurst, A. (1998). International expansion of retail firms: A theoretical approach for future investigations. *Journal of Retailing and Consumer Services*, *5*(3), 143-151. https://doi.org/10.1016/S0969-6989(96)00056-2
219. Volberda, H. W. (1997). Building flexible organizations for fast-moving markets. *Long range planning*, *30*(2), 169-148. https://doi.org/10.1016/S0024-6301(96)00110-0
220. Wang, P. and Chan, P.S. (1995). Top management perception of strategic information processing in a turbulent environment. *Leadership and Organization Development Journal*, Vol. 16 No. 7, pp. 33-43. https://doi.org/10.1108/01437739510100937
221. Ward, P. T., Duray, R., Leong, G. K., and Sum, C. C. (1995). Business environment, operations strategy, and performance: an empirical study of Singapore manufacturers. *Journal of operations management*, *13*(2), 99-115. https://doi.org/10.1016/0272-6963(95)00021-J
222. Waterhouse, J. H., and Tiessen, P. (1978). A contingency framework for management accounting systems research. *Accounting, Organizations and Society*, *3*(1), 65-76. https://doi.org/10.1016/0361-3682(78)90007-7
223. White, D. W. (1998). The impact of environmental uncertainty on strategy creation style in a franchise channel setting. *Journal of Strategic Marketing*, *6*(4), 273-304. https://doi.org/10.1080/096525498346568
224. Winsor, R. D. (1995). Marketing under conditions of chaos: percolation metaphors and models. *Journal of Business Research*, *34*(3), 181-189. https://doi.org/10.1016/0148-2963(94)00115-U
225. Winsor, R. D. (1995). Marketing under conditions of chaos: percolation metaphors and models. *Journal of Business Research*, *34*(3), 181-189. https://doi.org/10.1016/0148-2963(94)00115-U
226. Wollmann, D., and Steiner, M. T. A. (2017). The strategic decision-making as a complex adaptive system: a conceptual scientific model. *Complexity*, *2017*. https://doi.org/10.1155/2017/7954289
227. Yasai-Ardekani, M., and Haug, R. S. (1997). Contextual determinants of strategic planning processes. *Journal of Management Studies*, *34*(5), 729-767.
228. Yu, W. (2011). *Operations strategy, business environment, operations resources and performance: an empirical study of retail firms in China* (Doctoral dissertation, University of Nottingham).
229. Yu, W., and Ramanathan, R. (2011). Effects of firm characteristics on the link between business environment and operations strategy: evidence from China's retail sector. *International Journal of Services and Operations Management*, *9*(3), 330-364. https://doi.org/10.1504/IJSOM.2011.041103
230. Yu, W., and Ramanathan, R. (2012). Effects of business environment on international retail operations: case study evidence from China. *International Journal of Retail and Distribution Management*. pp. 218-234. https://doi.org/10.1108/09590551211207175
231. Zhang, X. (2009). Retailers' multichannel and price advertising strategies. *Marketing Science*, *28*(6), 1080-1094.
232. Zhao, K., Zuo, Z., and Blackhurst, J. V. (2019). Modelling supply chain adaptation for disruptions: An empirically grounded complex adaptive systems approach. *Journal of Operations Management*, *65*(2), 190-212. https://doi.org/10.1002/joom.1009

Chapter II

Influence of the internal and external drivers on consumers' visits and sales productivity

2.1 Introduction

During recent years, in the era of diversity and variety of consumer demands, the retail sector has been involved in a period of continuous transformation, where the appearance of retail formats coming into direct competition with the so-called traditional commercial formats, their medium-sized outlets, lacking range of choice and low prices and with products of medium-high quality. On one side, there is the condition of the saturated retail markets on the other, the retailers setting the balanced combination of positioning tactics over a time for the matter of survival. One of the arising issues in the retail's positioning is the role of the size and type of retail format and the assortment size to define consumer's choice of a retailer becoming a more often decisive factor of perceived positioning. The actual competitive retail environment is characterized by a growing heterogeneity of demand and by the expansion of new retail formats, determining cross-shopping behaviors by consumers - consumers change from a store to another according to their purchase needs and the attributes of each store (Bustos-Reyes and González-Benito, 2008; Russo et al.,2019). The trend is to focus on location, format, and the spatial distribution of consumers in modeling the competitive environment (Ellickson et al., 2020). The choice of the retail format prevails the preference of the customer for the choice of the assortment type. In a marketing literature prior decision research contain opposite findings: By Broniarczyk et al. (1998) when choosing among assortments, consumers prefer the variety offered by larger assortments, opposite, consumers are less confident in larger assortment rather prefer smaller assortments (Iyengar and Lepper, 2000; Chernev 2003a, 2003b; 2006; Goodman, and Malkoc, 2012; Chernev et al., 2015; Kim et al., 2015; Aurier and Mejía, 2020). However lessening category assortments show less or no negative impact on sales in the affected product categories, it decreases the customer's visits (Fox and Sethuraman, 2010). Choice overload largely depends on the subjective state of the customer during the decision-making process, and it can be categorized as process-based indicators and outcome-based indicators (Chernev et al., 2015). Aurier and Mejía (2020) state that the choice of the small assortment is time-saving and cost–saving in the decision-making process. Even the smaller assortment can benefit the future disinterest of customers it gives the possibility to retailers to quickly improve the quality of the products.

As existing literature has already provided the relevant research on how the different assortment sizes and variety affect consumer decisions (Borin et al., 1994; Borin and Farris, 1995; Iyengar and Lepper, 2000; Chernev 2003a, 2003b; 2006, Kim et al., 2015) albeit, this work aims to investigate the relationships between assortment sizes, type of retail stores, number of customers and sales productivity. Using the data of retail chain drug stores, this work investigates how assortment size and retail format type influence consumers' visits and what external factors - as competition distance and school holidays - or internal ones - such as promotional strategies and the size of the assortment- influence sales productivity.

Providing large assortments in drug stores could be the visual appeal for a consumer, at the same time it does not satisfy the consumer's need and expectations. This is a relevant issue for marketing managers that have to manage the portfolio of the assortments and decide how many options to include in the category, the question arising for retailers who need to care about the shelf display and how many products can be enough to satisfy consumer preference for the multiple attributes of a product. This confusion is reflected in the trend of large retailers using multiple retail formats with mixed sizes of the product assortment to satisfy the different consumers' preferences (Reutterer and Teller, 2009; Ruiz-Real et al., 2016).

The work is organized as follows: at first, the literature background regarding the internal drivers and relationships between retail format, assortment size, customers, following the link between the retail environment and sales productivity; at second, the research hypotheses have been tested through statistical analyses of the Rossmann drug retail chain data; at third, the results have been discussed proposing the related solutions and recommendations; at final, the work ends with the conclusions, implications, limitations and emerging future research directions.

2.2 Internal drivers: Retail format, assortment level and consumers' visits

2.2.1 The retail format, evolution and consumers' visits

A complex environment makes the retailers shaping into new operational formats. The retailer does not appear in one mode, the so-called "traditional" format of the corner store with a limited assortment suitable to the needs of the larger consuming population. The retailers are defining into the new shapes, diverse channels adopting the multi-channel strategy (Sonneck and Ott, 2010). The choice of the retail format affecting the consumers' choice of visiting the preferred store's type.

Consumers have wide spectra of store formats and channels where to shop. They constantly making choices, unpredictably shifting the decision-making, as consumers follow their gut feeling and pattern when and where to make the purchase. The expectations become higher. There have been many studies that analyze the role of the store's format in the consumer experience (Goldman, 2001; Goldman et al., 2002; Simova et al., 2003; Gonzalez-Benito et al., 2005; Reutterer and Teller, 2009; Brown, 2010; Teller et al., 2012; Shi et al., 2018). Brown (2010) discuss that retail format is the shape in which retailer mix the type of the merchandise, price and promotions strategies, visual aspect, approach location. The mix of these elements including the external environment creates the image for consumers in which the retail format has the social, psychological, and esthetical aspects (Goldman, 2001). When drafting a store format, retailers have to consider the local community needs and see to it that architecture and set-up blend in with the surroundings (Kumar, 2005; Russo et al., 2019).The demographic structure and habits of the place can direct the type of the retail format, for this reason, the international expansion of the retailer is difficult as it is not the standardized model that can operate in each country. The food retail industry should take consumers' external and internal shopping motivations into consideration when developing store format strategies (Jacobs et al., 2010). Retailers must provide consumers with a store that is well located, has a logical store arrangement, and offers a variety of products. The retailer has to consider the portfolio of managing the retailers' formats according to the consumer profile. French type of hypermarket is challenging in the American society and vice versa. Shi et al. (2018) show that format diversification is affecting negatively the retailer's performance while geographic diversification benefits the retailers. Today, the new way of format so-called the "digital" format taking the place in the contact with the consumer. Digitalization is allowing a retailer to create new business concepts using the available information that is available by new technological devices. The mobile device creates a unique hedonistic consumer experience (Weill and Woerner 2015; Rigby, 2011); also the plenty source of information that the company can use and filter to build future strategies. Purchasing using a mobile device can discover more than enough details about a consumer, where, when, what he did purchase, the quantity, the choice of the products. These details create the big databases that become the source for the managers in predicting the consumer demand, market basket analysis, brand-choice analysis. The trend in retailing is not anymore the choosing the channel to reach the consumer, it is the combining the channels to co-create value and engage with customers. At first, the concept of multichannel implies separation and combination between channels, whereas the omnichannel concept is more focused on customers and on providing them with the ability to move between 'channels' seamlessly during one integrated purchasing process (Piotrowicz and Cuthbertson, 2014). The concept of multichannel and omnichannel retailing has emerged a the result of digitalization where the retailers can create the consumer experience regardless of the distribution channel they use (Hänninen et al., 2020; Blitz, 2016; Lemon and Verhoef, 2016). The new business concepts have a focus on maximizing the value besides the shareholder's value. The firms are forced to go digitally even the creating the digital strategy is not as simple as it seems. There is an innovation in digital-making decisions that can make even big companies struggle in keeping the market share. Although there is research that shows that incumbent firms are struggling with digital strategy innovation as they need to cope with new technologies (Lusch and Nambisan 2015; Nylén and Holmström 2015).

Hagberg (2016) observed that there are several transformations regarding the digitalization: 1) digitalization of exchanges; 2) digitalization of the actors; 3) the digitalization of the settings 4) digitalization of offerings. Focusing on the first element, the authors outline the changes in the communication between retailer and consumer. The consumer is much more concerned and sophisticated about the retailer's assortment thru the digital interface.

Consumers take under consideration however totally different types of supply (products and services) impact on the expected profit and convenience after they opt for a store sort (Kamran-Disfani et al., 2017). In step with many authors, the mix of advantages (both tangible and intangible) offered by every store influences customer's perceived utility in several ways that as their utility functions differ (Messinger and Narasimhan, 1997; Solgaard and Hansen, 2003; Cleeren et al., 2010; Russo et al, 2019). Bhatnagar and Ratchford (2004) state that buyers choose the shop that reduces the perceived prices (product value, travel, and consumer's storage cost). As steered by the service output theory (Coughlan et al., 2006) within the retail channels literature, in line with the store selection literature, customers base their call in respect of what store to buy at not completely on product attributes or value. On the contrary, they compromise between those parts and repair outputs (for example, the approach a product is purchased) after they opt for wherever to buy. The higher the service outputs level (such as customer service, reduced waiting time, information provision, spatial convenience) the more appealing to customers a store is. Customers are expected to go for the store that can potentially improve their overall experience (Kamran-Disfani et al., 2017). According to Van Waterschoot et al. (2008), the selection of a store is joined to the knowledge that customers collect concerning its product. In step with the previous literature, there is a tendency to define the subsequent analysis hypothesis (Russo et al., 2019):

H1: The type of the retail format affects the consumer's visits

2.2.2 *Assortment level and consumers' choice*

Academic literature focused on establishing a single assortment for a retailer (Mantrala et al., 2009; Hübner and Kuhn, 2012; Shin et al., 2015). Other studies showed that a retailer could have a different assortment at each store (Kök et al., 2008).

The endless debate on assortment levels often surfaces the dichotomy between "more-is-better" (Baumol and Ide, 1956) and "choice overload" (Iyengar and Lepper, 2000). The former states that large assortments benefit customers as they offer more opportunities to satisfy disparate customer preferences (Baumol and Ide, 1956). According to some studies, large assortments increase anticipated consumption utility and actual consumption (Kahn and Wansink, 2004), the likelihood of purchase (Koelemeijer and Oppewal, 1999), and the ease of comparing different options (Hutchinson, 2005). On the other hand, other studies show that large assortments are likely to decrease purchases (Iyengar and Lepper, 2000), reduce decision satisfaction (Schwartz, 2000; Haynes, 2009), and increase choice difficulty (Fasolo et al., 2009). Recent studies show that the costs of choosing from a large assortment rise faster than its benefits when the size of the assortment increases. The result is an inverted U-shaped relationship between assortment size and choice satisfaction (Shah and Wolford, 2007; Lenton et al., 2008; Reutskaja and Hogarth, 2009).

In a related line of inquiry, the authors (Chernev, 2003a; 2003b; 2006; Xu et al., 2013; Chernev et al., 2015) indicate that consumers might not be keen on making comparisons, and that could be a reason for the asymmetry between the increase in costs and benefits when assortments size increases. Meta-analysis nevertheless shows that the U-shaped relationship between assortment size and choice satisfaction does not explain all of the variance (Scheibehenne et al., 2010). Besides, studies in related pieces of literature prove that some consumers might indeed be more inclined to or interested in making comparisons (Kruglanski et al., 2013). Independently by the results obtained, the different authors agree on the influence that the assortment level exercises on consumer purchasing behavior. Consumers' perceptions of the assortment range come from the combination of multiple indicators, mostly the number of stock-keeping units proposed and the availability of the favorite brands (Amine and Cadenat, 2003). Consumers' evaluation is driven by the perceived choice within the product categories where they are highly sensitive to the assortment range. A choice among assortments is a function of consumers' decision focus: when the decision focus is on choosing the assortment then the preference for larger assortments is bigger when the focus is on choosing from the voluntarily selected assortment (Boyd and Bahn, 2009). In the research of Gázquez-Abad et al. (2015), the results show that mixed assortments, especially those that include a large number of NBs (National brands) are associated with greater store loyalty. In the category of toilet tissue, both large and small mixed assortments were associated with similar levels of store loyalty. Without going deeply into the consumer's profile and the reason behind the decisions made, the authors will try to analyze the moment of when a purchase is done and to explore the dynamics based on consumers' visits, then the second hypothesis will be tested:

H2: The assortment level influences consumer's visists

2.3 Retail environment, drivers and sales productivity

Intense competition brings growing pressure on retailers to better measure and evaluates the performance to adjust the operational strategies. The measures used in practice are the practical indicators of sales productivity, efficiency, and effectiveness (Zentes et al., 2007; 2017). Depending on what the retailer is eager to achieve, the measures can be a different evaluation of 1) productivity/effectiveness 2) inventory/supply chain 3) profit/turnover 4) financial performance. The common aim of retail managers is to increase productivity from the different aspects (Goodman, 1985; Kumar and Karande, 2000; McGoldrick, 2002; Zentes et al., 2007; 2017, Zuberi and Rajaratnam, 2020). The ratios widely used for measuring the sales productivity in retail stores are 1) sales per square meter (square foot) 2) sales per full-time employee. Indicator sales per square meter measures total sales per unit area (Kumar and Karande, 2000). It represents an indicator of the efficiency of a store's management explaining how the revenue is generated using the available amount of sales space. Although, sales productivity can be affected by several factors of the internal and external environment (Harrauer and Schnedlitz, 2016). In their research, Kumare and Karande (2000) observed that store performance is influenced by internal store environment (number of checkouts per square foot of sales area, non-grocery products sold, banking facility, openness of the store, doubles of manufacturers coupons) and external environment (total number of households and type of neighborhood).

The values of the sales productivity can give the wrong image of the store's performance, lower values do not mean lower performance at the same time, the high number of employees may link to the high service positioning strategy or the low sales per square meter that there is the more comfortable shopping space in the store. Still, these ratios are used to compare the store with benchmark competition's store inside the retail chain. The sales per square meter can benefit the controlling of the store's management inside the retail chain (Zentes et al., 2007; 2017). The retail format "shopping centers" faces a decline in productivity due to the saturated market and suffers overwhelming of the selling surface (LeHew and Fairhurst, 2000).

Intense competition reduces the sale productivity decreasing the effectiveness of the large retail formats, although the change in demographic structure and lifestyle habits added the effect on the decline in sales productivity. Today, there is the fragmentation of the consumer market leading to the micro-segmentation concerning the mass market. The cultural diversity, the increase of elderly, and the growing number of divorced people tend to change the strategy to consumers of the large retail formats. The segmentation of the market is not enough, the clustering of these segments according to the similar characteristics, habits, and preference allows large retail formats to build the mix of the promotional, price, and store layout strategies (LeHew and Fairhurst, 2000). The evidence is that many obstacles from the internal and external environment trigger sales productivity.

In the literature, the relevant factors influencing the sales productivity are the atmosphere of the place in the sense of creating an effort to build a pleasant buying environment (Kotler, 1973; Ballantine, et al. 2015), demographic changes, and socioeconomic characteristics (Ghosh and Craig, 1983; Prieto and Caemmerer, 2013), behavioral consequences (Hennigs et al., 2015), the effect of promotions (Walters and MacKenzie, 1988; Haans and Gijsbrechts, 2011), the seasonal sales (Newing, et al., 2013). Haans and Gijsbrechts (2011) investigated the effects of store size on effectiveness when promotions are applied. Hübner (2017) studies the influence of assortment size and shelf space planning to better meet consumer demand and respond to increased product variety. As the retail environment becomes more complex (Achrol,1991; Wang and Chan,1995; Hwang, 2005; Moretta Tartaglione et al., 2018; 2019) additional variables from the macro and micro-retail environment must be considered when evaluating store performance there is the need for new metrics in evaluating store performance (Zuberi and Rajaratnam, 2020). Research focus on external drivers as the financial-economic crisis (Harrauer and Schnedlitz 2016), consumers changing needs, climate change and fuel price (Moretta Tartaglione et al., 2018; 2019), the level of competition (Ghosh, 1984), and competition distance (Mintz and Currim, 2013), internal drivers as promotions (Christen et al., 1997; Haupt and Kagerer, 2012; Smith et al., 2019), store assortment, store size and store location (Keh and Chu, 2003; Korhonen and Syrjänen, 2004; Vaz et al., 2010; Haans and Gijsbrechts, 2011), multichannel impact (Lapoule and Colla, 2016).

Since the store productivity is related to the sales and consumers purchasing behavior that is hypothesized to be influenced by the store type, the following hypothesis is:

H3: The size of the retail format influences the effects of internal and external factors on sales productivity

2.4 Method

The research hypotheses are tested using multiple methods based on a mix of classification and regression tools to analyze secondary data of a drug retailer.

Particularly, the classification tool Linear Discriminant Analysis (LDA) (Fisher, 1940; Loog et al., 2001) is used to test hypotheses H1 and H2. Classification aims to examine the belonging of data to specific groups regarding a set of attributes. The LDA classifier is the linear function of observation x, where $\hat{u}_k$ is the average of all training observations belonging to the kth class, $\hat{\sigma}^2$ weighted average of the sample variance for each K classes and $\hat{\pi}_k$ is the proportion of the training observations belonging to the kth class (James et al., 2013).

$$\hat{\sigma}_k(x) = x * \frac{\hat{u}_k}{\hat{\sigma}^2} - \frac{\hat{u}_k{}^2}{2\hat{\sigma}^2} + \log(\hat{\pi}_k)$$

Equation 2.1

LDA is generally used when populations are multivariate normal and all different groups have equal covariance matrices (Fisher, 1940; Loog et al., 2001; James et al., 2013). In this work, this tool is used to classify the high-dimensional dataset by clustering customers concerning store type and assortment they prefer, according to the variables indicated in Table 2-1. LDA is compared to QDA (quadratic discriminant analysis) that is recommended if the training set is very large, the decision boundary is quadratic (non-linear) so that the variance of the classifier is not a major concern, or if the assumption of a common covariance matrix for the K classes is untenable (Fisher, 1940; Loog et al., 2001; James et al., 2013); and multinomial logistic regression(MLogR) where the assumption is the same as LDA, linearity and Gaussian distribution.

Multiple regression analysis (Anderson, 1984; Mason and Perreault, 1991) considers finding the linear combination of a set of predictors (Equation 2) that provides the best point estimates of the dependent variable across a set of observations.

$$y_i = \beta_0 + \beta_1 x_i + \beta_2 x_2 + \beta_3 x_3 + \cdots + \beta_d x_i \ + e_i$$

Equation 2.2

Predictive accuracy is defined by the level of the R--squared and the statistical significance of the completed model. The purpose of the multiple regression model is to interpret the effect of the individual predictor variable on the response. In that case, the center is the (standardized) regression coefficients, their estimated standard errors, and the associated t-test probabilities. These statistics are used to test hypotheses about the effect of predictors on the dependent variable or to evaluate their relative weight. Multiple Linear Regression (MLinR) is used to detect the influence of internal and external factors on the sales productivity (Equation 2) for each cluster detected with LDA to test hypothesis H3. Running the regression on different aggregates of large data provides a clearer connection between the predictors and the outcome, thus increasing the statistical significance (Cook et al., 2002; Russo et al., 2019)

Table 2-1. Relevant categorical variables

Assortment	Basic- articles of personal hygiene, cleanings products, baby and family care etc.
	Extra - Basic assortment with additional fresh products
	Extended – larger basic assortment with more toys and extended household products
Store type	Neighborhood market
	Train station store
	Shopping Centre
	Retail park
Net sales area (square meters)	< 200 - <900
	<300 - <800
	<300 - <1000
	<300 - <900

Source: Relevant categorical variables (Russo et al., 2019; p 138.)

According to the Table 2-1. three store formats are presented in data: *Neighborhood market*- a store adapted to satisfy the needs and wants of local surroundings with net sales area 200-800 sqm.; *Train station store*-a store located in nearby train stations; *Shopping centre*- large retail format of interconnected shops that enables the customer to walk from one to another with net sales area: 500 - 1000 sqm; *Retail park*: unenclosed complex retail format of big and small retailers with a surface area from 20,000-55,000 sqm (in the paper the drug store located in a retail park with net sales area 500-1000 sqm).

Table 2-2. Relevant internal and external variables

Internal factors	
Sales productivity	Daily sales turnover for sales area
Promo	Ongoing promotions on the day of the observation
Promo_2	Continuing and consecutive promotion for some stores
Assortment	Assortment level of the store: basic, extra, extended
External Factors	
Competition distance	Distance in meters to the nearest competitor store
School holiday	Closure of public schools on the day of the observation

Source: Relevant continuous and dichotomous variables (Russo et al., 2019; p 138.)

2.4.1 Results from descriptive analysis

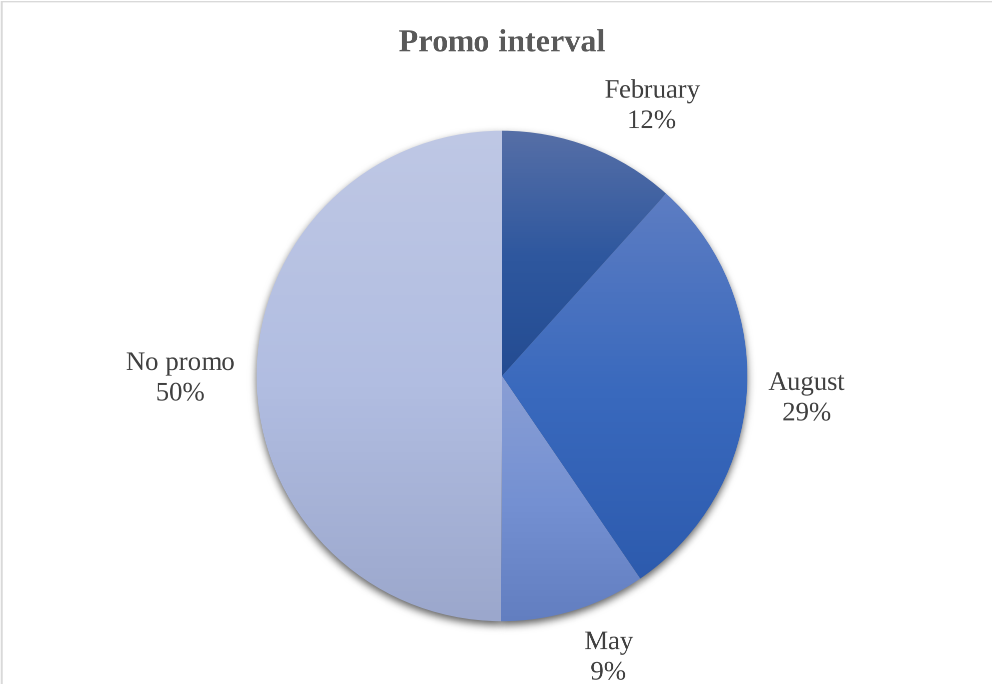
Retail data are from 1,115 drug stores Rossmann, a retailer located in Germany, and collected online covering the period from 1st January 2013 to 31st July 2015. The permission for use of data and additional information are provided by the manager of operations of the executive board of the Rossmann retail chain. Data concerning the following variables were collected daily: Sales, Customers (the number of customers who visited the store), Store type, Assortment level, Size key (size of the store), School holidays, National holidays, Competition distance, and Promo (the store takes part in promotions) or Promo2 (the store takes a part in consecutive promotions) (Table 2-2).

Table 2-3. Descriptive statistics of data (Rossmann drug stores)

Store	**Day of the week**	**Date**	**Sales**	**Customers**	**Promo2**
Min.: 1.0	Min.: 1.000	Min.: 2013-01-01	Min.: 0	Min.: 0.0	Min.: 0.0000
1st Qu.: 280.0	1st Qu.: 2.000	1st Qu.: 2013-08-17	1st Qu.: 3727	1st Qu.: 405.0	1st Qu.: 0.0000
Median: 558.0	Median: 4.000	Median: 2014-04-02	Median: 5744	Median: 609.0	Median: 1.0000
Mean: 558.4	Mean: 3.998	Mean: 2014-04-11	Mean: 5774	Mean: 633.1	Mean: 0.5006
3rd Qu: 838.0	3rd Qu.: 6.000	3rd Qu.: 2014-12-12	3rd Qu.: 7856	3rd Qu.: 837.0	3rd Qu.: 1.0000
Max.: 1115.0	Max.: 7.000	Max.: 2015-07-31	Max.: 41551	Max.: 7388.0	Max.: 1.0000
NA	NA	NA	NA	NA	NA
National holiday	School holiday	Size key	Competition distance	Promo	
Min.: 0	Min.: 0.0000	Min.: 1.000	Min.: 20	Min.: 0.0000	
1st Qu.: 0	1st Qu.: 0.0000	1st Qu.: 4.000	1st Qu.: 710	1st Qu.: 0.0000	
Median: 0	Median: 0.0000	Median: 4.000	Median: 2330	Median: 0.0000	
Mean: 0	Mean: 0.1786	Mean: 4.519	Mean: 5430	Mean: 0.3815	
3rd Qu.: 0	3rd Qu.: 0.0000	3rd Qu.: 6.000	3rd Qu.: 6890	3rd Qu.: 1.0000	
Max.: 0	Max.: 1.0000	Max.: 9.000	Max.: 75860	Max.: 1.0000	
NA's: 31050	NA	NA's: 2826	NA's: 2642	NA	

Source: Summary statistics of retail data (Rossmann drug store chain) (Russo et al., 2019; p.139.)

Figure 2-1. Consecutives promotions starting in February, May and August

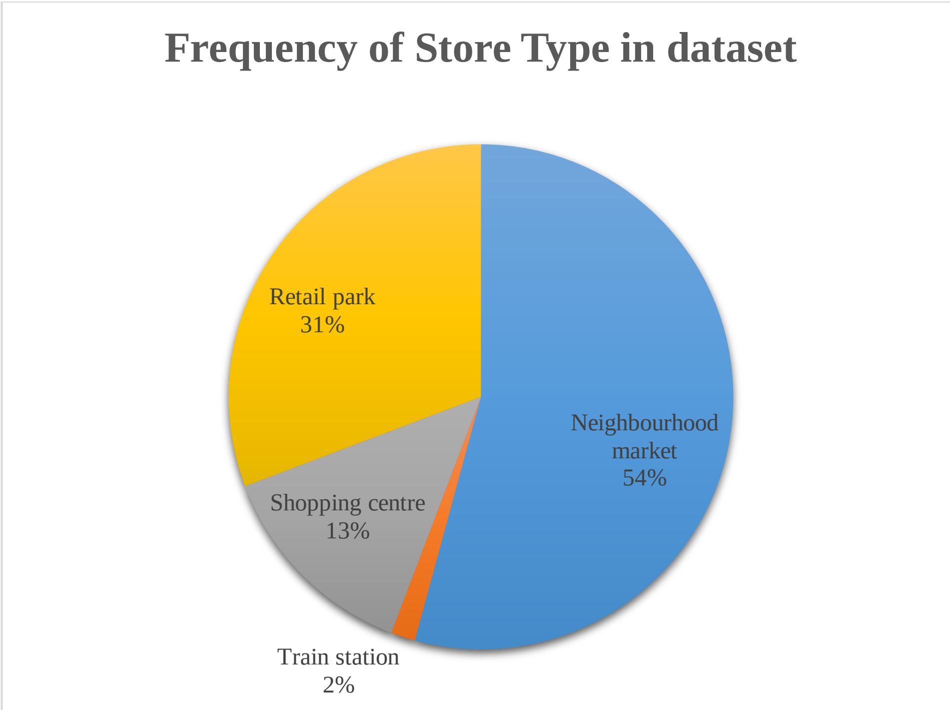


Source: Author's elaboration

Table 2-3. shows the descriptive statistics, the range of each variable in data with maximum and minimum values, mean, median, and the count of missing values. A large dispersion in variables emerges due to the size of the sample. Data appear normally distributed except for variable Competition Distance affected by extreme values, which are corrected in the further analysis. Figure 2-1. shows that the period with longer promotions starts in August (29%), then in February (12%), and in May (9%), 50% of the time the interval of promotions are shorter.

The aim of the analysis is a classification based on the store type and assortment level, time variables (Day of week and Date) are excluded from the analysis. The same occurred for the data about National Holidays and Promo_2. as there is missing information about it in the dataset.

Figure 2-2. Frequency of store format in data



Source: Author's elaboration

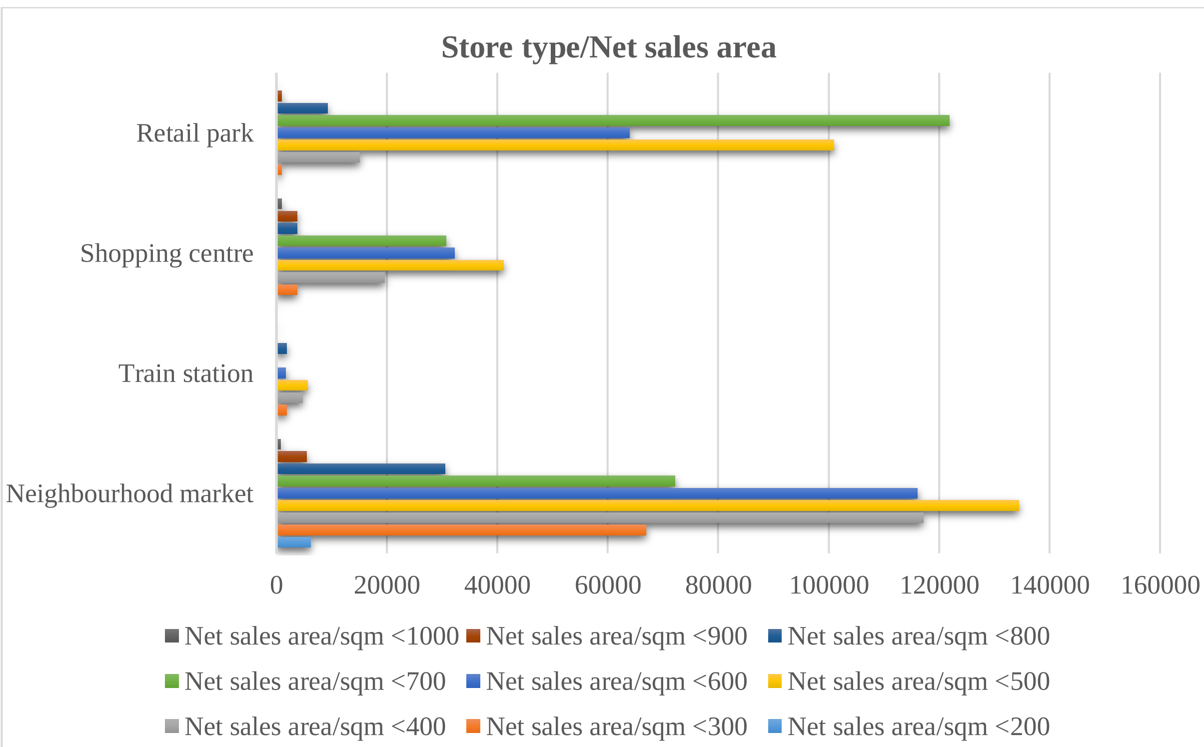
Table 2-4. Frequency of net sales area (sqm)

Net sales area (sqm)		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	*<200*	6226	.6	.6	.6
	<300	73542	7.2	7.2	7.9
	<400	156504	15.4	15.4	23.3
	<500	282127	27.7	27.8	51.1
	<600	213944	21.0	21.1	72.2
	<700	224740	22.1	22.2	94.4
	<800	45422	4.5	4.5	98.8
	<900	10178	1.0	1.0	99.8
	<1000	1700	.2	.2	100.0
	Total	1014383	99.7	100.0	
Missing	System	2826	.3		
Total		1017209	100.0		

Source: Frequency of net sales area (Size key variable) (Russo et al., 2019; p.139)

Table 2- 4. shows the frequency of data by the net sales area of the store. Net sales area represents the area indicated in square meters where products are displayed and sold. This excludes the storage, inventory, and parking surface. The most frequent sizes of the net sales area range from 500 to 700 sqm representing 70.8% of the dataset, while 25.8% of the dataset are net sales area from 200 to 400 sqm and 5.7% belongs to the range from 800 to 1000 sqm. Using cross-tabulation, it maps the frequency distribution of two categorical variables in a contingency table. The objective is to explore whether the two variables are related. By visualizing them with the type of retail store (Figure 2-3), it emerges that Neighborhood markets have the most presence in the size from 200 to 500 sqm., while Retail Park and Shopping Centre have from 500 to 1000 sqm of net sales area. The Train station store appears principally in the range from 300 to 800 sqm of net sales area (Figure 2-2).

Figure 2-3. Relationship between Net sales area and Store type

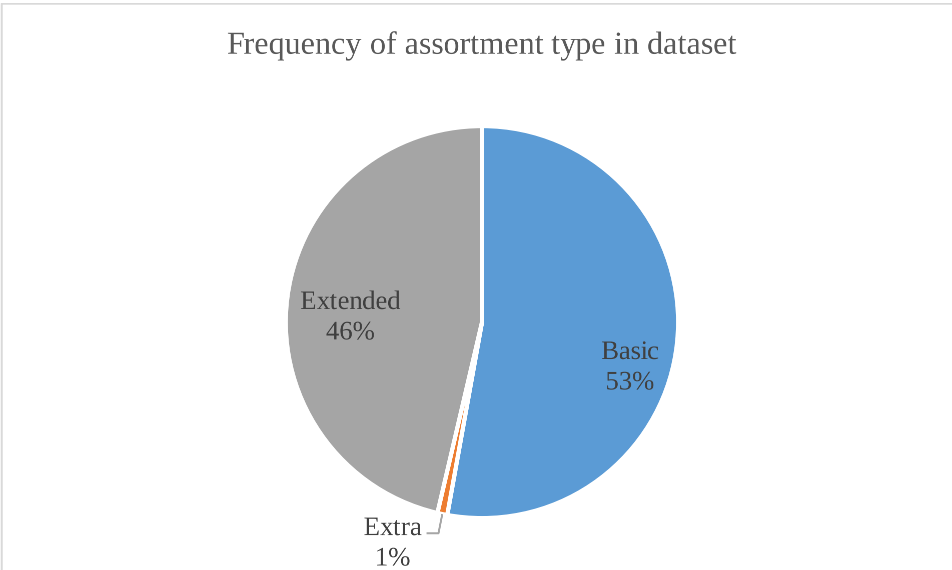


Source: Author's elaboration

Assortment represents a display of goods or services that a retailer provides to the customers.

Observing the frequency of the assortment type in the dataset (Figure 2-4), the most occurred is the Basic assortment (53%) which include the articles of personal hygiene, cleanings products, baby and family care, then Extended assortment (46%) includes larger basic assortment with more toys and extended household products and the last Extra assortment (1%) is the basis assortment with more fresh products.

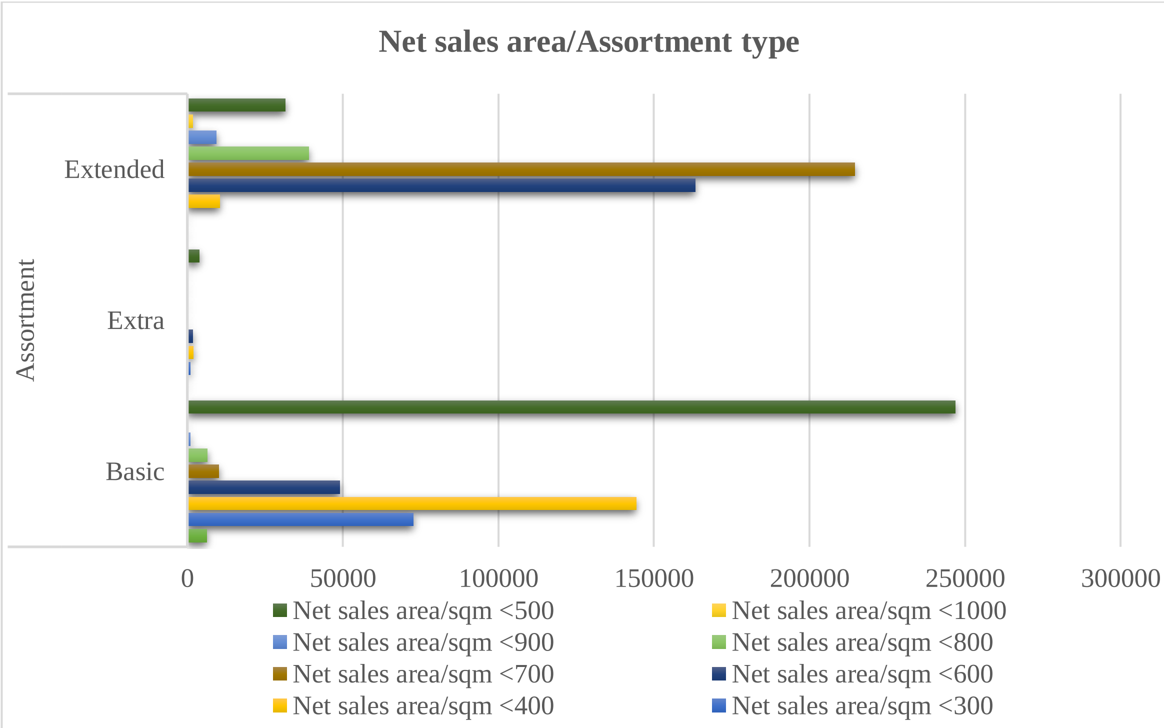
Figure 2-4. Frequency of the assortment type in dataset



Source:Author's elaboration

From the analysis of the relationships between net sales area and level of Assortment (Figure 2-5), it is possible to observe that Basic assortment occurs mostly in the range from 300 to 500 sqm, while Extended and Extra assortments occur in the range from 600 to 1000 sqm. However, the Extra assortment does not have a high incidence in the dataset, hence the focus of the analysis will be only on Basic and Extended assortments.

Figure 2-5. Relationship between Net Sales Area key and Assortment type



Source: Author's elaboration

Concerning the relationship between store type and assortment level in data, Table 2-5 shows how many times each type of Assortment occurs within each type of Store. While Basic assortment represents the major assortment level in Neighborhood markets (62.8%) and Shopping centers (52.2%), Extra assortment is the part of the Train Station stores (52.4%) while Extended assortment is the part of the Retail park (63.7%) and Shopping centre (47.8%).

Table 2-5. Type of the assortment inside in diverse retail formats

			Assortment level			Total
			Basic	Extra	Extended	
Retail format	Neighborhood market	Count	345447	0	204296c	549743
		% within Store format	62.8%	0.0%	37.2%	100.0%

	Train station store	Count	6594$_a$	8294$_b$	942$_c$	15830
		% within Store format	41.7%	52.4%	6.0%	100.0%
	Shopping Centre	Count	70878$_a$	0$_b$	65020$_c$	135898
		% within Store format	52.2%	0.0%	47.8%	100.0%
	Retail park	Count	113584$_a$	0$_b$	199328$_c$	312912
		% within Store format	36.3%	0.0%	63.7%	100.0%
Total		Count	536503	8294	469586	1014383
		% within Store format	52.9%	0.8%	46.3%	100.0%
Each subscript letter denotes a subset of Assortment categories whose column proportions do not differ significantly from each other at the .05 level.						

Source: Relationship between store type and assortment level (Russo et al., 2019; p. 141)

Having the insight in descriptive statistics of the retail dataset it helps to determine the dynamics of the particular cluster that determine the basis for the classification and regression tools.

2.4.2 *Results from the classification and regression analysis*

The classification analysis aims to cluster the variables based on to highest membership they belong to the type of the store and type of the assortment. Before the classification tools, it is estimated likelihood that the observations belong to certain clusters (Table 2- 6). Prior probabilities are used to determine which classification function to apply and adjusting them can improve the overall classification rate.

Table 2-6. Comparing means between retail format clusters

Store Type	**mean(Net sales area)***	**mean(Customers)**	**mean(Sales)**	**Prior probabilities**
Neighborhood market	4.196561	658.2962	5733.995	.542
Train station store	3.928869	1987.7208	10058.837	.016
Shopping centre	4.718348	674.9377	5745.604	.133
Retail park	5.029305	501.4349	5641.819	.308

*the measurement of net sales area variable (1-10)

Source: Author's elaboration

Results from Table 2-6. show clustering data by the store type comparing means of Net sales area, Customers and Sales. The dominant cluster "Train store station" with the highest average sales and number of customers has less than 2% of assigned observations in data, which leads to excluding that cluster in further analysis. The probabilities of belongings to clusters Neighborhood markets is 0.54, then to cluster Retail Park 0.31 and Shopping centre 0.13.

In the next step, the trends of customers visiting different store types are analyzed using the comparison of the best model fit using LDA, QDA, and MLogR. As the model is developed on the trainset (75% of data), the performance is controlled (predicted) on the rest 25% of the data or test set. The performance of the models is compared by testing the model on the test set resulting in the miss-classification rate, which shows how precise the classification is. The output of the confusion matrix (Table 2-8) that present the table of the correct and incorrect classified predicted values on the test set, and in this case the LDA1 has the least miss-classification rate of 0.339 in comparison to QDA and MLogR, that means 34% out of time the observations are misclassified and 66% of the time the classification is accurate (Table 2-7).

Table 2-7. The miss-classification rate of the LDA1, MLogR and QDA

Classification model	**The miss-classification rate**
LDA	0.339
QDA	0.343
MlogR	0.509

Source: Author's elaboration

Table 2-8. The prediction performance of LDA1 on the tests set

Confusion matrix	Neighborhood market	Train station store	Shopping centre	Retail park
Neighborhood market	35787	0	527	6381
Train station store	0	0	0	0
Shopping centre	7482	0	304	2586
Retail park	9588	0	1	15595

Source: Author's elaboration

The coefficients of LDA provide the linear combination of variables that are used to form a decision boundary distinguishing the clusters. Clustering the dataset by store type (LDA1) (Tab. 2-9), it emerges that 55% of training observations are classified to Neighborhood markets, 13% to Shopping centers, and 32% to Retail parks. Cluster's averages (Group means) reveal that average sales are higher for the type Retail Park than for Neighborhood market and Shopping Centre, while the average of customers' visits is slightly higher in the Neighborhood market than Shopping Centre and Retail Park. Competition distance is lower for Shopping Centre than for the other two clusters, while School holidays and Promotions are similar in each cluster.

Table 2-9. Linear discriminant analysis (LDA1)-classification by the type of the retail format

Store type	Prior prob	Sales	Customers	School holidays	Competition distance	Promo	Extended assortment	Net sales area
Neighborhood market	0.547	5599.643	617	0.165	3701.03	0.382	0.283	4.034
Shopping Centre	0.131	5351.198	616	0.166	2793.277	0.3811	0.502	4.812
Retail park	0.320	5651.974	480	0.163	5943.896	0.382	0.615	5.014

Source: Summary of LDA2 considering Sales productivity (sales per square meter) (Russo et al.,2019; p.141)

In retail practice, it is common to use the sales per square meter to express how much sales revenue it is possible to obtain for each square meter of retail space provided. The following calculations are used:

$$\textit{Sales per square meter} = \textit{Total Net Sales} \div \textit{Total Net Sales Area}$$

(Equation 2.3)

The analysis has been applied to check the sales productivity by store type (LDA2) by replacing the variable Sales with Sales per square meter (Tab. 2-10). The results show that the higher sales productivity is for the cluster Neighborhood market type and the lowest for the Retail Park.

Table 2-10. Linear discriminant analysis (LDA2) including Sales per square meter)

Store type	Prior prob	Sales productivity (sales per sqm)	Customers	School holidays	Competition distance	Promo	Extended assortment
Neighborhood market	0.547	1513.813	617	0.165	3701.03	0.382	0.283
Shopping Centre	0.131	1184.443	616	0.166	2793.277	0.3811	0.502
Retail park	0.320	1170.535	480	0.163	5943.896	0.382	0.615

Source: Summary of LDA2 considering Sales productivity (sales per square meter) (Russo et al.,2019; p.142)

Further classification tools have been applied to examine which clusters are defining based on the type of the assortment.

Clustering the dataset according to the level of assortment (LDA3) (Tab. 2-11), it shows that 58% of training observations belong to the typology “basic assortment”, while 42% belong to the “extended assortment”. The model excludes the “extra assortment” due to missing information in data. Average Sales productivity and average Customers are higher in cluster Basic assortment than for Extended assortment, while average Competition distance is higher for Extended assortment. Promotions and

School holidays are constant in both clusters.

Table 2-11. Linear discriminant analysis LDA3 –classification by the assortment level

Assortment type	Prior probabilities	Sales productivity	Customers	Competition distance	School holiday	Promo
Basic	0.581	1568.466	575	3826.702	0.164445	0.381471
Extended	0.418	1071.584	570	4958.157	0.166373	0.383167

Source: Summary of LDA3 considering Sales productivity (sales per square meter) (Russo et al.,2019; p.142)

In the final analysis, to test Hypothesis H3 whether the internal or external variables affect the sales productivity in the three store's clusters: Neighborhood market, Shopping Centre and Retail Park, Multiple linear regression (MLinR) has been performed (Russo et al.,2019). In Table 2-12. the output of the three MLinR models is illustrated outlining parameter estimates (β) and (t) - test statistics. There is the absence of multicollinearity in every model given by low VIF (variance inflation factor). For Neighborhood markets, adjusted *R-squared* shows that 65% of the variance in Sales productivity is explained by the linear combination of internal variables (Extended assortment, Customers, and Promo) and external variables (Competition distance and School holiday). A significant relationship appears between Promo, School holidays, and Sales productivity, meaning that if the shop applies promotional discounts during school holidays the sales productivity increases considerably. Still positive and important, however less sturdy, is that the relationship between Customers, Competition distance and Sales productivity. A negative and important relationship emerges between extended assortment and sales productivity, showing that increasing assortment level negatively influences sales productivity in Neighborhood markets. For Shopping centers, Adjusted *R-squared* shows that 85% of the variance in outcome is explained by displayed variables. In this case, the most significant and positive relationship is between Promo and Sales productivity, while no significant relationship emerges between School holidays and Sales productivity. Like for Neighborhood markets, also the relationship between Customers, Competition distance and Sales productivity is significant and positive and the relationship between Extended assortment and Sales productivity is negative. This similar scenario is conferred in Retail parks, Adjusted *R-squared* shows that 84% of the variance in Sales productivity is explained by the displayed variables.

Table 2-12. Multiple linear regression (MlinR) for the retail format clusters (Response variable-Sales productivity)

Store type	**Neighborhood market**		**Shopping center**		**Retail park**	
Parameters	*β*	*t*	*β*	*t*	*β*	*t*
Intercept (β_0)	176.35	43.31***	55.34	13.23***	173.53	60.88***
Extended assortment	-683.96	-206.72***	-270.20	-85.73***	-296.51	-155.08***
Customers	1.22	226.23***	1.70	267.18***	2.47	423.92***
Promo	483.63	144.45***	227.61	70.514***	192.29	91.62***
Competition distance	0.00023	8.32***	0.03	49.17***	0.0014	7.79***
School holidays	36.02	8.45***	0.67	0.169	7.58	2.99**
Adjusted R^2	**0.646**		**0.853**		**0.837**	
VIF (range)	**(1.02-2)**		**(1.03-2.51)**		**(1.01-2.67)**	
*** p-value ≤0.001 significance of parametric (linear) relationship ** p-value ≤0.01 * p-value ≤0.05 . p-value ≤0.1						

Source: Parameter estimates for the clusters by store format (Outcome-Sales productivity) (Russo et al.,2019; p.142)

In the second turn, the results of internal and external factors on sales productivity have been calculated supported the assortment levels. Table 2-13. displays the set of parameters of the MlinR models relating to the two clusters: basic and extended assortment. With relation to Basic assortment, all predictors, except Competition distance, show a positive and significant relationship with sales productivity. Adjusted *R-squared* shows that 63% of the variance in sales productivity is explained by Customers, Promo, and faculty Holidays, following the absence of multiple correlations given by low VIF (variance inflation factor). For the Extended assortment cluster, all predictors have a positive impact on the sales productivity of stores. During this case, the sales productivity is plagued by competition distance showing that the increase in competition distance will increase the sales productivity for all sorts of stores with extended assortment

Table 2-13. Multiple linear regression (MlinR) for the clusters by assortment level (Response variable-Sales productivity)

Assortment type	**Basic assortment**		**Extended assortment**	
Parameters	*β*	*T*	*β*	*t*
Intercept (β_0)	-2.71	-0.728	-66.24	-27.67***
Customers	1.44	259.02***	1.34	366.43***
Promo	472.79	145.45***	245.8	128.58***
Competition Distance	No significant		129.9	75.15***
School Holidays	28.48	7.23***	20.71	8.91***
Adjusted R^2	**0.642**		**0.776**	
VIF (range)	**(1.02-1.92)**		**(1.02-2.12)**	

*** p-value ≤0.001 significance of parametric (linear) relationship

** p-value ≤0.01

* p-value ≤0.05

. p-value ≤0.1

Source: Parameter estimates for the clusters by assortment level (Outcome-Sales productivity) (Russo et al.,2019; p.143)

2.5 Discussion

Findings obtained by the classification and regression tools determine the outcome of testing the hypothesis**.** Linear discriminant analysis, as the type of cluster analysis, has been used as a data reduction technique to easy and better approach the individual observations (Punj and Stewart, 1983). Obtaining homogenous clusters can give us a better insight into the data structure and dynamics of observations inside the cluster.

By comparison customers' visits within the three clusters of store varieties (Table 2-9), it emerges a transparent preference of consumers for Neighborhood markets and Shopping centers. However, considering the sales productivity (Table 2-10), it results considerably greater in Neighborhood markets. This implies that, at equal visits, customers tend to shop for additional in these stores. Given the fact that the sales per square meter are the effective measurement of how effective the store layout and sales staff are at selling the products, the result of the LDA2 gives a clearer frame of where the customer's visits increase the profitability of the sales and in this case that is in the Neighborhood markets. The large retail chain will use a similar design concept for all of their stores and will ensure uniformity in each shop's location. Therefore, the sales per square meter can be used as well to identify when cultural differences are affecting the sales and warrant further inspection.

This result may rely on the very fact that whereas Shopping centers square measure places of leisure during which people go conjointly only for walks, playtime and take a glance at the merchandise within the stores, in Neighborhood markets people go solely and solely once they have to be compelled to get one thing. Therefore, it is additional frequent than in the stores the visits change into purchases. Since it is attainable to spot a transparent correlation between the kind of store and the purchase trends of shoppers, we can state that the primary hypothesis may be accepted:

H1: The type of the retail format influences consumer's visits

In the conflict of research, whether customers prefer the large or smaller assortment, the recent research shows that customers may be less attracted to the larger assortment (Goodman and Majkoc, 2012). The consumer is finding it difficult to choose and in front of the large assortment and experiences the less joyful moments and is frustrated about time spending (Kim and Ellinger, 2015). In this analysis, there are two approaches to see if the customer prefers the small or large assortment. The first, using the classification analysis to see in which cluster of the assortment type the customers are assigned. The results of the model LDA3 (Table 2-11.) show that stores with basic assortment have a greater number of visits and major sales per sqm than stores with extended assortment. One reason supporting these results could be that basic assortment takes the greatest part in Neighborhood markets and Shopping centers, which, as shown above, are those with the greatest number of visits and higher sales productivity. The second approach is examining the influence of the extended assortment in the regression analysis by the store's cluster (Table 2-12). In each store cluster, it is evident that extended assortment is negatively affecting the sales productivity, which means increasing the level of assortment size decreasing the sales per square meter in each cluster's store. The study can give us small evidence of the consumer choice but cannot give us the state of the consumer's mind and behavior that should investigate further of what makes the consumer make the final choice based on the assortment preference. The study of Kim and Ellinger (2015) shows that consumers felt overwhelmed, stressed, confused, and regretful when dealing with a purchase in the stores with the enlarged assortment. The benefit of the small assortment Neighborhood markets can have, which permit to manage and improve the quality of the smaller assortment that become popular to the customers. The fact that customers visit the convenience stores to purchase with the clear aim and fast decision-making, the smaller and known assortment can benefit the percentage rise of the finalized purchase. Any addition to the assortment helps to amplify the assortment evaluation, in the terms of additional benefit of the different items (and not the extent to which the items add to attribute-based variety) that leads to amplified assortment evaluations. Effects of item presence in an assortment on store evaluation depend on the customer's interest in the particular item (Oppewal and Koelemeijer, 2005). On the opposite aspect, the literature supports these results with studies on decision-making in keeping with that consumers have some difficulties in creating a selection once an outsized assortment is provided (Chernev, 2006). Selection within the giant assortments creates uncertainty in customers' decision-making decreasing the need to buy (Broniarczyk and Hoyer, 1998). Identification of ideal point availability as a key factor moderating the impact of assortment on choice and showing that contrary to the common wisdom that more choice is always better, selections made from large assortments can lead to weaker preferences (Chernev, 2015). The negative effects of enormous assortments stem from consumers' aversion to creating comparisons (Shugan 1980; Chernev 2003; Chernev et al., 2015; Xu et al., 2013). Despite the causes or the sign of the relation between assortment and customers' selections, the results clearly show a correlation between assortment level and buying trends of consumers, so supportive the second hypothesis:

H2: The consumers' visits are affected by the assortment level in the store

Once the factors influencing consumer-purchasing behaviors are understood, it is necessary to spot the most external and internal factors that affect the productivity of sales (Tab. 2-12).

Among the internal factors touching sales productivity, adoption of promotional ways is the most relevant because it will increase productivity for every type of store and shows larger positive effects than different positive factors like several customers' visits. Whereas the foremost important negative internal issue is that the presence of an extended assortment that looks to decrease the productivity of sales for every format of the store. These results conjointly support the second hypothesis. Customers are also intended by extended assortment to go to the drug store in looking centers and retail parks however the choice to get might not be created for an equivalent reason. Iyengar and Lepper (2000) mentioned that customers are at first attracted by the number of merchandise on the shells however because of decision-making problems consumers felt demotivated to create an alternative, exploit the shop while not buying. They are doubtless to get during a store with a basic assortment, ideally placed within the neighborhood. The explanations are also the daily habits and past looking experiences, they need in drug stores.

Concerning external factors, competition distance shows to be associated with sales productivity in each case however, the relation is not therefore sturdy. However, the competition distance is an important element that can affect the profitability of the sales. The complexity of decision-making is increasing when the retailer wants to open a new shop, taking into account the geo-competition and distance to the competitors' stores (Roig-Tierno et al., 2013). Understanding the competitor's location can help to understand the better the shift and direction in a particular industry and to help to deeper understand the actions of the competitors and based on that better project future moves. In this particular example, there is a limit of information on which kind of store format is the competitor's store. The things are changing if the distance is from the store of the same or different stores. In the first case, the cannibalization effect may be the case in which the stores of the same chain are close to each other in the aim of the economy of the scope (Aguirregabiria and Suzuki, 2016). In the second case, the distance from the competitor's store from the different retail chain can have either positive or negative effects. The competition distance in this example has a positive impact on the profitability of the sales; it can be hypothesized that supermarket is the observed competitor and drug retail chain is competitive due to the reduction of price. Completely different relations will be discovered once examining the influence of gap throughout School holidays on Sales productivity. Throughout School holidays, sales productivity will increase for the Neighborhood markets and, and slightly however still positive, for Retail parks. School holidays have, instead, no effects on sales productivity of Shopping centers.

The outline of the internal and external effects on sales productivity of the three completely different store types is conferred in Table 2-14.

Table 2-14. Internal and external effects on the sales productivity by the retail format

Retail format / ***Factors***		**Neigbourhood market**	**Shopping center**	**Retail park**
Internal factors	Extended assortment	Decrease	Decrease	Decrease
	Customers visits	Increase	Increase	Increase
	Promo	Increase	Increase	Increase
External factors	Competition distance	Increase	Increase	Increase
	School holiday	Increase	No effect	Increase

Source: Author's elaboration

Since for each store type the effects produced by these factors appear homogeneous within the cluster but, in some cases, dissimilar among the different store types, the third hypothesis can be accepted:

H3: The retail format influences the effects of internal and external factors on sales productivity

2.6 Conclusions

Retailers are focusing on delivering the product or service to diverse multichannel forms. Being an innovative retailer keeps customers active and fresh in choosing where to do the purchase, as their choice is based on the occasion, situation, or even individual decision-making pattern (Sonneck and Ott, 2010). The retailers' choice of the format affect also the consumer's choice. Understanding retail customers in terms of the store type and assortment level they choose, as well as their behavior in different situations, is the critical issue for retail managers to build strategies at many levels. In this work, classification and regression tools permit the analysis of the multivariate data creating clusters and facilitating the interpretation of data. The results of this study show that consumers prefer rather purchase in neighborhood markets that offer a limited assortment. These so-called convenience stores show the highest level of sales profitability during promotions and school holidays periods. However, retailers cannot disregard other types of stores, stores in shopping centers and retail parks, which are those that consumers prefer to visit. These stores thus become a significant showcase to increase the visibility of a retailer and make products known to consumers. Therefore, the best solution for retailers is to adopt multi-channel formats. Rolling out more store types by a big retailer there are two important benefits: (1) the convenience of Neighborhood markets allows attracting consumers who generally buy at shopping centers for occasional and rapid shopping; (2) stores in shopping centers allow exploiting the great merchandise and sales area (Russo et al., 2019).

When puzzling the different types of stores, a retailer should focus on the assortment level and the display of the articles. The results of this study show that basic assortments stimulate sales per square meter and they are convenient for the smaller "neighbor's" store where the goal is to increase sales profitability. The extended levels of assortment can be suitable for shopping centers where the main objective is to increase visits and display products independently of sales. However, too many choices in one assortment have negative effects in any type of store. Regardless of the type of store, retailer's main key tactics are promotions, as they become the most important element for boosting and predicting sales. Nevertheless, the retailer needs to think about the geolocation of the store and its position to the competitors, the distance appears to have a significant effect on the sales in any type of store.

The main limitation of this work is the lack of information about the consumer's profile and preferences as well the analysis carried out on a single retailer and in a single country. The validity of the results cannot, therefore, be generalized. Future research can be aimed at repeating the investigation for another retail chain especially comparing the cases in different countries. To the increasing market volatility, the future collection of data will be complex and will create the necessity to approach the structures with a variety of tools. In this analysis, we show that with the combination of the classification and regression tools it is possible to manage a large sample and deliver a considerable conclusion about dynamics in the retail context.

2.7 References

1. Achrol, R.S. (1991). Evolution of the marketing organization: new forms for turbulent environments. *Journal of Marketing*, 55(October), 77-93. https://doi.org/10.1177%2F002224299105500406
2. Aguirregabiria, V., and Suzuki, J. (2016). Empirical games of market entry and spatial competition in retail industries. In Handbook on the Economics of Retailing and Distribution. Edward Elgar Publishing https://doi.org/10.4337/9781783477388.00019
3. Amine, A., and Cadenat, S. (2003). Efficient retailer assortment: a consumer choice evaluation perspective. *International Journal of Retail and Distribution Management*, *31*(10), 486-497. https://doi.org/10.1108/09590550310497012
4. Anderson, T. W. (1984). An introduction to multivariate statistical analysis. New York: Wiley
5. Aurier, P., and Mejía, V. D. (2020). The interplay of brand-line assortment size and alignability in the sales of brand-lines and line-extensions of frequently purchased products. Journal of Business Research, 117, 163-175. https://doi.org/10.1016/j.jbusres.2020.05.024
6. Baumol W. J., Ide E. A. (1956). Variety in retailing. Manag Sci 3:93–101. https://doi.org/10.1287/mnsc.3.1.93
7. Ballantine, P. W., Parsons, A., and Comeskey, K. (2015). A conceptual model of the holistic effects of atmospheric cues in fashion retailing. *International Journal of Retail and Distribution Management*. https://doi.org/10.1108/IJRDM-02-2014-0015
8. Bhatnagar, A., and Ratchford, B. T. (2004). A model of retail format competition for nondurable goods. International Journal of Research in Marketing, 21(1), 39–59. https://doi.org/10.1016/j.ijresmar.2003.05.002
9. Blitz, A. (2016). Beset by the digital revolution successful retailers embrace technology that enhances customer value. *Strategy and Leadership*. https://doi.org/10.1108/SL-09-2016-0073
10. Boyd, D. E., and Bahn, K. D. (2009). When do large product assortments benefit consumers? An information-processing perspective. *Journal of Retailing*, *85*(3), 288-297. https://doi.org/10.1016/j.jretai.2009.05.008
11. Borin N, Farris PW (1995) A sensitivity analysis of retailer shelf management models. J Retail 71:153–171 https://doi.org/10.1016/0022-4359(95)90005-5
12. Borin N, Farris PW, Freeland R (1994) A model for determining retail product category assortment and shelf space allocation. Decision Science 25, 359–384. https://doi.org/10.1111/j.1540-5915.1994.tb00809.x
13. Broniarczyk, S. M., Hoyer, W. D., and McAlister, L. (1998). Consumers' perceptions of the assortment offered in a grocery category: The impact of item reduction. *Journal of marketing research*, 166-176. https://doi.org/10.1177%2F002224379803500203
14. Brown, J. R. (2010). Managing the retail format portfolio: An application of modern portfolio theory. *Journal of Retailing and Consumer Services*, *17*(1), 19-2 https://doi.org/10.1016/j.jretconser.2009.09.001
15. Bustos-Reyes, C.A., González-Benito, Ó. (2008). Store and store format loyalty measures based on budget allocation. Journal of Business Research, 61 (9), 1015–1025. https://doi.org/10.1016/j.jbusres.2007.03.008
16. Chernev, A (2003a). Product Assortment and Individual Decision Processes. *Journal of Personality and Social Psychology*, 85 (July), 151–62. https://psycnet.apa.org/doi/10.1037/0022-3514.85.1.151
17. Chernev, A (2003b).When More Is Less and Less Is More: The Role of Ideal Point Availability and Assortment in Consumer Choice. *Journal of Consumer Research*, 30 (September), 170–83. https://doi.org/10.1086/376808
18. Chernev, A. (2006). Decision focus and consumer choice among assortments. *Journal of Consumer Research*, *33*(1), 50-59. https://doi.org/10.1086/504135
19. Chernev, A., Böckenholt, U., and Goodman, J. (2015). Corrigendum to "Choice overload: A conceptual review and meta-analysis". Journal of Consum Psychol 22 (2015) 333-358. https://doi.org/10.1016/j.jcps.2015.07.001
20. Christen, M. Gupta, S. Porter, J.C., Staelin, R. and Wittink, D.R. (1997). Using market-level data to understand promotion effects in a nonlinear model. *Journal of Marketing Research*, pp.322-334.

https://doi.org/10.1177%2F002224379703400302
21. Cleeren, K., Verboven, F., Dekimpe, M. G., and Gielens, K. (2010). Intra-and interformat competition among discounters and supermarkets. Marketing Science, 29(3), 456–473. https://doi.org/10.1287/mksc.1090.0529
22. Cook, T. D., Campbell, D. T., and Shadish, W. (2002). *Experimental and quasi-experimental designs for generalized causal inference*. Boston: Houghton Mifflin.
23. Coughlan, A. T., Anderson, E., Stern, L. W., and El-Ansary, A. I. (2006). Marketing channels (7th ed.). Upper Saddle River, NJ: Prentice Hall.
24. Ellickson, P. B., Grieco, P. L., and Khvastunov, O. (2020). Measuring competition in spatial retail. The RAND Journal of Economics, 51(1), 189-232. https://doi.org/10.1111/1756-2171.12310
25. Fasolo B., Carmeci F. A., and Misuraca R. (2009). The effect of choice complexity on perception of time spent choosing: when choice takes longer but feels shorter. Psychology and Marketing, 26 (3), 213-228. https://doi.org/10.1002/mar.20270
26. Fisher, R. A. (1940). The precision of discriminant functions. *Annals of Eugenics*, *10*(1), 422-429. https://doi.org/10.1111/j.1469-1809.1940.tb02264.x
27. Fox, E. J., and Sethuraman, R. (2010). Retail competition. In *Retailing in the 21st Century* (pp. 239-254). Springer, Berlin, Heidelberg.
28. Gázquez-Abad, J. C., Martínez-López, F. J., Mondéjar-Jiménez, J. A., and Esteban-Millat, I. (2015). Mixed assortments vs. store brand-only assortments: The impact of assortment composition and consumer characteristics on store loyalty. *Revista Española de Investigación de Marketing ESIC*, 19(1), 24-45. https://doi.org/10.1016/j.reimke.2014.11.001
29. Ghosh, A. (1984). Parameter nonstationarity in retail choice models. *Journal of Business Research*, *12*(4), 425-436. https://doi.org/10.1016/0148-2963(84)90023-7
30. Ghosh, A., and Craig, C. S. (1983). Formulating retail location strategy in a changing environment. *The Journal of Marketing*, 56-68. https://doi.org/10.1177%2F002224298304700307
31. Goldman, A. (2001). The transfer of retail formats into developing economies: the example of China. *Journal of retailing*, *77*(2), 221-242. https://doi.org/10.1016/S0022-4359(01)00044-6
32. Goldman, A., Ramaswami, S., & Krider, R. E. (2002). Barriers to the advancement of modern food retail formats: theory and measurement. *Journal of Retailing*, *78*(4), 281-295. https://doi.org/10.1016/S0022-4359(02)00098-2
33. González-Benito, O., Munoz-Gallego, P. A., and Kopalle, P. K. (2005). Asymmetric competition in retail store formats: Evaluating inter-and intra-format spatial effects. Journal of Retailing, 81(1), 59–73.
34. Goodman, C.S. (1985), Comment: on output measures of retail performance. *Journal of Retailing*, Vol. 61, pp. 77-82.
35. Goodman, J. K., and Malkoc, S. A. (2012). Choosing here and now versus there and later: The moderating role of psychological distance on assortment size preferences. Journal of Consumer Research, 39(4), 751-768. https://doi.org/10.1086/665047
36. Hand, D. J., and Henley, W. E. (1997). Statistical classification methods in consumer credit scoring: a review. *Journal of the Royal Statistical Society: Series A (Statistics in Society)*, *160*(3), 523-541. https://doi.org/10.1111/j.1467-985X.1997.00078.x
37. Haans, H., and Gijsbrechts, E. (2011). One-deal-fits-all? On category sales promotion effectiveness in smaller versus larger supermarkets. *Journal of Retailing*, *87*(4), 427-443. https://doi.org/10.1016/j.jretai.2011.05.001
38. Hänninen, M. (2020). Review of studies on digital transaction platforms in marketing journals. *The International Review of Retail, Distribution and Consumer Research*, *30*(2), 164-192. https://doi.org/10.1080/09593969.2019.1651380
39. Harrauer, V. and Schnedlitz, P. (2016).Impact of environment on performance measurement design and processing in retailing, *International Journal of Retail and Distribution Management*, 44(3), pp.320-335. https://doi.org/10.1108/IJRDM-03-2015-0048
40. Hagberg, J., Sundstrom, M., and Egels-Zandén, N. (2016). The digitalization of retailing: an exploratory framework. *International Journal of Retail and Distribution Management*. https://doi.org/10.1108/IJRDM-09-2015-0140

41. Haupt, H. and Kagerer, K. (2012). Beyond mean estimates of price and promotional effects in scanner-panel sales–response regression. *Journal of Retailing and Consumer Services*, 19(5), pp.470-483. https://doi.org/10.1016/j.jretconser.2012.06.002
42. Haynes, G. A. (2009). Testing the boundaries of the choice overload phenomenon: The effect of number of options and time pressure on decision difficulty and satisfaction. Psychology and Marketing, 26, 204–212. https://doi.org/10.1002/mar.20269
43. Hennigs, N., Wiedmann, K.-P., Klarmann, C. and Behrens, S. (2015). The complexity of value in the luxury industry: From consumers' individual value perception to luxury consumption. *International Journal of Retail and Distribution Management*, Vol. 43 No. 10/11, pp. 922-939. https://doi.org/10.1108/IJRDM-07-2014-0087
44. Hübner A. H., Kuhn H. (2012). Retail categorymanagement: state-of-the-art reviewof quantitative research and software applications in assortment and shelf space management. Omega Int J Manag Sci 40:199–209. https://doi.org/10.1016/j.omega.2011.05.008
45. Hübner, A., and Schaal, K. (2017). An integrated assortment and shelf-space optimization model with demand substitution and space-elasticity effects. *European Journal of Operational Research*, *261*(1), 302-316. https://doi.org/10.1016/j.ejor.2017.01.039
46. Hutchinson, J. M. (2005). Is more choice always desirable? Evidence and arguments from leks, food selection, and environmental enrichment. *Biological reviews*, *80*(1), 73-92. https://doi.org/10.1017/S1464793104006554
47. Hwang, E. J. (2005). *Strategic management and financial performance in South Korean apparel retail stores* (Doctoral dissertation, Virginia Tech). http://hdl.handle.net/10919/29855
48. Iyengar, S. S., and Lepper, M. R. (2000). When choice is demotivating: Can one desire too much of a good thing? *Journal of personality and social psychology*, *79*(6), 995. https://psycnet.apa.org/doi/10.1037/0022-3514.79.6.995
49. Jacobs, S., Van Der Merwe, D., Lombard, E., & Kruger, N. (2010). Exploring consumers' preferences with regard to department and specialist food stores. *International Journal of Consumer Studies*, *34*(2), 169-178. https://doi.org/10.1111/j.1470-6431.2009.00844.x
50. James, G., Witten, D., Hastie, T., and Tibshirani, R. (2013). *An introduction to statistical learning* (Vol. 112). New York: springer. https:/doi.org/10.1007/978-1-4614-7138-7
51. Kahn B. E., Wansink B. (2004). The Influence of Assortment Structure on Perceived Variety and Consumption Qualities, Journal of Consumer Research; 30 (4), 519-533. https://doi.org/10.1086/380286
52. Kamran-Disfania, O., Mantralab, M. K., Izquierdo-Yustac, A., Martínez-Ruiz, M. P. (2017). The impact of retail store format on the satisfaction-loyalty link: An empirical investigation, Journal of Business Research 77, 14–22. https://doi.org/10.1016/j.jbusres.2017.04.004
53. Keh, H. T., and Chu, S. (2003). Retail productivity and scale economies at the firm level: a DEA approach. *Omega*, *31*(2), 75-82. https://doi.org/10.1016/S0305-0483(02)00097-X
54. Kahn, B. E., and Wansink, B. (2004). The influence of assortment structure on perceived variety and consumption quantities. *Journal of consumer research*, *30*(4), 519-533. https://doi.org/10.1086/380286
55. Kim K., Allaway A.W., Ellinger A.E. (2015) Retail Assortment Size and Customer Choice Overload: The Influence of Shopping Enjoyment and Time Pressure. In: Robinson, Jr. L. (eds) Marketing Dynamism and Sustainability: Things Change, Things Stay the Same…. Developments in Marketing Science: Proceedings of the Academy of Marketing Science. Springer, Cham. https://doi.org/10.1007/978-3-319-10912-1_45
56. Koelemeijer K., Oppewal H. (1999). Assessing the effects of assortment and ambience: a choice experimental approach, Journal of Retailing 75 (3), 319-345. https://doi.org/10.1016/S0022-4359(99)00011-1
57. Kök A. G., Fisher M. L.,Vaidyanathan R. (2008). Assortment planning: reviewof literature and industry practice. In: Agrawal N, Smith SA (eds) Retail Supply Chain Manag, vol 223. Springer, USA, pp 99–153. https://doi.org/10.1007/978-0-387-78902-6_6
58. Korhonen, P., and Syrjänen, M. J. (2004). Resource allocation based on efficiency analysis. *Management Science*, 50(8), 1134–1144. https://doi.org/10.1287/mnsc.1040.0244
59. Kotler, P. (1973). Atmospherics as a marketing tool. *Journal of retailing*, *49*(4), 48-64.
60. Kruglanski A. E., Pierro A., Mannetti L., and Higgins T. E. (2013). The Distinct Psychologies of "Looking" and

"Leaping": Assessment and Locomotion as the Springs of Action, Social and Personality Psychology Compass 7/2 (2013): 79–92. https://doi.org/10.1111/spc3.12015
61. Kumar, P., (2005). The competitive impact of service process improvement: Examining customers' waiting experiences in retail markets. J. Retail. 81 (3), 171–180. https://doi.org/10.1016/j.jretai.2005.07.002
62. Kumar, V., and Karande, K. (2000). The effect of retail store environment on retailer performance. *Journal of business research*, *49*(2), 167-181. https://doi.org/10.1016/S0148-2963(99)00005-3
63. Lapoule, P., Colla, E. B., Towers, N., and Kotzab, H. (2016). The multi-channel impact on the sales forces management. *International Journal of Retail and Distribution Management*. https://doi.org/10.1108/IJRDM-11-2014-0159
64. LeHew, M.L.A. and Fairhurst, A.E. (2000), "US shopping mall attributes: an exploratory investigation of their relationship to retail productivity", International Journal of Retail and Distribution Management, Vol. 28 No. 6, pp. 261-279. https://doi.org/10.1108/09590550010328535
65. Lenton, A. P., Fasolo, B. and Todd, P. M. (2008). 'Shopping' for a mate: expected vs. experienced preferences in online mate choice. IEEE Transactions on Professional Communication (Special Section: Darwinian Perspectives on Electronic Communication), 51, 169–182. https://doi.org/10.1109/TPC.2008.2000342
66. Lemon, K. N., and Verhoef, P. C. (2016). Understanding customer experience throughout the customer journey. *Journal of marketing*, *80*(6), 69-96. https://doi.org/10.1509%2Fjm.15.0420
67. Loog, M., Duin, R. P. W., and Haeb-Umbach, R. (2001). Multiclass linear dimension reduction by weighted pairwise Fisher criteria. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, (7), 762-766. https://doi.org/10.1109/34.935849
68. Lusch, R. F., and Nambisan, S. (2015). Service innovation. *MIS quarterly*, *39*(1), 155-176. https://www.jstor.org/stable/26628345
69. Mantrala M. K., Levy M., Kahn B. E., Fox E. J., Gaidarev P., Dankworth B., Shah D. (2009). Why is assortment planning so difficult for retailers? A framework and research agenda. J Retail 85:71–83. https://doi.org/10.1016/j.jretai.2008.11.006
70. Mason, C. H., and Perreault Jr, W. D. (1991). Collinearity, power, and interpretation of multiple regression analysis. *Journal of marketing research*, *28*(3), 268-280. https://doi.org/10.1177%2F002224379102800302
71. McGoldrick, P. J. (2002). Retail marketing. McGraw-Hill.
72. Messinger, P. R., and Narasimhan, C. (1997). A model of retail formats based on consumers' economizing on shopping time. Marketing Science, 16(1), 1–23. https://doi.org/10.1287/mksc.16.1.1
73. Mintz, O. and Currim, I.S. (2013). What drives managerial use of marketing and financial metrics and does metric use impact performance of marketing mix activities? *Journal of Marketing*, 77(2), 17-40. https://doi.org/10.1509%2Fjm.11.0463
74. Moretta Tartaglione, A., Bruni, R., Bozic, M., and Cavacece,Y. (2018). How the environmental complexity affects forecasting in retail? Actual research trends, *Journal of Marketing Trends,* 5(2), 71.
75. Moretta Tartaglione, A., Bruni, R., and Bozic, M. (2019). Exploring the retail industry environment using nonlinear analysis. *International Journal of Retail and Distribution Management*, Vol. 47 No. 4, pp. 453-47. https://doi.org/10.1108/IJRDM-09-2017-0224
76. Newing, A., Clarke, G., & Clarke, M. (2013). Identifying seasonal variations in store-level visitor grocery demand. *International Journal of Retail & Distribution Management.*
77. Nylén, D., and Holmström, J. (2015). Digital innovation strategy: A framework for diagnosing and improving digital product and service innovation. *Business Horizons*, *58*(1), 57-67. https://doi.org/10.1016/j.bushor.2014.09.001
78. Oppewal, H., and Koelemeijer, K. (2005). More choice is better: Effects of assortment size and composition on assortment evaluation. *International Journal of Research in Marketing*, *22*(1), 45-60. https://doi.org/10.1016/j.ijresmar.2004.03.002
79. Pal, K. (2018). Building High Quality Big Data-Based Applications in Supply Chains. In *Supply Chain Management Strategies and Risk Assessment in Retail Environments* (pp. 1-24). IGI Global. https://doi.org/10.4018/978-1-5225-3056-5.ch001

80. Piotrowicz, W., and Cuthbertson, R. (2014). Introduction to the special issue information technology in retail: Toward omnichannel retailing. *International Journal of Electronic Commerce*, *18*(4), 5-16. https://doi.org/10.2753/JEC1086-4415180400
81. Prieto, M., and Caemmerer, B. (2013). An exploration of factors influencing car purchasing decisions. *International Journal of Retail and Distribution Management*. https://doi.org/10.1108/IJRDM-02-2012-0017
82. Punj, G., and Stewart, D. W. (1983). Cluster analysis in marketing research: Review and suggestions for application. *Journal of marketing research*, *20*(2), 134-148. https://doi.org/10.1177%2F002224378302000204
83. Reutskaja, E., and Hogarth, R. M. (2009). Satisfaction in Choice as a Function of the Number of Alternatives: When "Goods Satiate" but Bads Escalate. Psychology and Marketing, 26 (3), 197-203. https://doi.org/10.1002/mar.20268
84. Reutterer, T., and Teller, C. (2009). Store format choice and shopping trip types. *International Journal of Retail and Distribution Management*, *37*(8), 695-710. https://doi.org/10.1108/09590550910966196
85. Rigby, D. (2011). The future of shopping. *Harvard business review*, *89*(12), 65-76.
86. Roig-Tierno, N., Baviera-Puig, A., Buitrago-Vera, J., and Mas-Verdu, F. (2013). The retail site location decision process using GIS and the analytical hierarchy process. *Applied Geography*, *40*, 191-198. https://doi.org/10.1016/j.apgeog.2013.03.005
87. Ruiz-Real, J. L., Gázquez-Abad, J. C., and Martínez-López, F. J. (2016). Assortment Size and PL Penetration in Grocery Retailers' Portfolios during Economic Crisis: Empirical Evidence from Spain. In *Handbook of Research on Strategic Retailing of Private Label Products in a Recovering Economy* (pp. 244-274). IGI Global. https://doi.org/10.4018/978-1-5225-0220-3.ch01
88. Russo, Giuseppe, Maja Bozic, Ylenia Cavacece, and Giuseppe Granata. "The Influence of Retailer Choices on Consumer Behaviors and Sales Productivity." In *Predicting Trends and Building Strategies for Consumer Engagement in Retail Environments*, pp. 133-151. IGI Global, 2019. https:/doi.org/10.4018/978-1-5225-7856-7.ch007
89. Scheibehenne, B., Greifeneder R. and Todd P. M. (2010). Can There Ever Be Too Many Options? A MetaAnalytic Review of Choice Overload, Journal of Consumer Research, 37 (3), 409–25. https://doi.org/10.1086/651235
90. Schwartz, B. (2000). Self determination: The tyranny of freedom. American Psychologist, 55, 79–88.
91. Shah, A. M., and Wolford, G. (2007). Buying behavior as a function of parametric variation of number of choices. Psychological Science, 18, 369–370.
92. Shi, Y., Lim, J. M., Weitz, B. A., and France, S. L. (2018). The impact of retail format diversification on retailers' financial performance. *Journal of the Academy of Marketing Science*, *46*(1), 147-167. https://doi.org/10.1007/s11747-017-0559-0
93. Shin H., Park S., Lee E., BentonW. C. (2015). A classification of the literature on the planning of substitutable products. Eur J Oper Res 246:686–699. https://doi.org/10.1016/j.ejor.2015.04.013
94. Shugan, S. M. (1980). The cost of thinking. *Journal of consumer Research*, *7*(2), 99-111. https://doi.org/10.1086/208799
95. Simova, J., Clarke-Hill, C. M., & Robinson, T. (2003). A longitudinal study of changes in retail formats and merchandise assortment in clothing retailing in the Czech Republic in the period 1994-1999. *International Journal of Retail & Distribution Management*. https://doi.org/10.1108/09590550310483314
96. Smith, A. N., Rossi, P. E., and Allenby, G. M. (2019). Inference for product competition and separable demand. *Marketing Science*, *38*(4), 690-710. https:/doi.org/10.1287/mksc.2019.1159
97. Solgaard, H. S., and Hansen, T. (2003). A hierarchical Bayes model of choice between supermarket formats. Journal of Retailing and Consumer Services, 10(3), 169–180. https://doi.org/10.1016/S0969-6989(03)00008-0
98. Sonneck, P., and Ott, C. S. (2010). Future trends in multi-channel retailing. In Retailing in the 21st Century (pp. 221-238). Springer, Berlin, Heidelberg. https://doi.org/10.1007/978-3-540-72003-4_14
99. Teller, C., Kotzab, H., & Grant, D. B. (2012). The relevance of shopper logistics for consumers of store-based retail formats. *Journal of Retailing and Consumer Services*, *19*(1), 59-66. https://doi.org/10.1016/j.jretconser.2011.09.001

100. Van Waterschoot, W., Sinha, P.K., Van Kenhove, P., De Wulf, K., (2008). Consumer learning and its impact on store format selection. J. Retail. Consum. Serv. 15 (3), 194–210. https://doi.org/10.1016/j.jretconser.2007.03.005
101. Vaz, C. B., Camanho, A. S., and Guimarães, R. C. (2010). The assessment of retailing efficiency using network data envelopment analysis. *Annals of Operations Research*, *173*(1), 5-24. https://doi.org/10.1007/s10479-008-0397-z
102. Walters, R. G., and MacKenzie, S. B. (1988). A structural equations analysis of the impact of price promotions on store performance. *Journal of marketing research*, 51-63. https://doi.org/10.1177%2F002224378802500105
103. Weill, P., and Woerner, S. L. (2015). Thriving in an increasingly digital ecosystem. *MIT Sloan Management Review*, *56*(4), 27.
104. Wang, P. and Chan, P.S. (1995). Top management perception of strategic information processing in a turbulent environment. *Leadership and Organization Development Journal*, Vol. 16 No. 7, pp. 33-43. https://doi.org/10.1108/01437739510100937
105. Xu J., Jiang Z., and Dhar R. (2013). Mental Representation and Perceived Similarity: How Abstract Mindset Aids Choice from Large Assortments. Journal of Marketing Research, L(August 2013), 548-559. https://doi.org/10.1509%2Fjmr.10.0390
106. Zentes, J., Morschett, D., and Schramm-Klein, H. (2007). *Strategic retail management*. Springer.
107. Zentes, J., Morschett, D., and Schramm-Klein, H. (2017). In-store Marketing. In *Strategic Retail Management* (pp. 327-350). Springer Gabler, Wiesbaden. https://doi.org/10.1007/978-3-658-10183-1_15
108. Zuberi, M. F., and Rajaratnam, D. (2020). Measuring retail performance in an omni-channel world. Journal of Marketing Channels, 26(2), 120-126. https://doi.org/10.1080/1046669X.2020.1741992

Chapter III

The causal effect of the environmental and promotional variables on the sales forecasting

3.1 Introduction

The environment is characterized by uncertainty, heterogeneity, and hostility (Crozier, 1964; Lawrence and Lorsch, 1967; Thompson, 1967; Hickson et al., 1971; Duncan,1972; Khandwalla, 1972; Miller and Friesen, 1983; Milliken,1987; Daft et al., 1988; Miller,1993; Buchko, 1994; Hwang, 2005; Rueda-Manzanares et al, 2008; Yu and Ramathan, 2011; Moretta Tartaglione et al., 2018).

That affects the relationship between suppliers and retailers, manufacturer power, retailer power, and competition (Dawson, 2000). The necessity to explain a deterministic flow process through a stochastic medium is suggested as useful for the study of consumer adoption processes and marketing management (Winsor, 1995). Description of an integrative model that develops interrelationships of retail and store environment, self-congruity, and retail patronage. (Sirgy, et al., 2000). Literature focus on the environmental factors, main decisions, and management changes that resulted from an ultimate demise of a successful retailer to help retail and other sales managers to increase their potential in the supply chain. (Lightfoot, 2003). QR (Quick Response) is recognized as an operational strategy that can provide demand-information in the last possible moment and quickly respond to a dynamic and volatile marketplace. Rueda - Manzanares (2008) discuss that (1) an organizational capability of stakeholder integration is related to a firm's assumption of a dynamic environmental strategy; (2) a complex business environment has a negative impact and an uncertain business environment has a positive impact on a firm's environmental strategy; and (3) complexity has a negative moderating impact on the alliance between a firm's stakeholder integration capability and its environmental strategy. Studying environment is a prior strategy when a retailer wants to expand its business internationally (Yu and Ramathan, 2011; 2012) suggesting appropriate strategies for multinational retailers willing to expand their business in emerging markets. Presentation of new practice theory on management control and performance measurement concentrating on the interface between information communication and interpretation at the retail sales floor including the dimensions of diversity, hostility, complexity, and turbulence (Harrauer and Schnedlitz, 2016). Complexity, its perception of the supply chain, and its relationship with performance help understanding the sources of the complexity and is it useful for the system or not (Ekinci and Baykasoğlu, 2019). There are also authors that state that environmental conditions don't have an important effect on stores' performance, even though stores' strategic management has been influenced by the perception of environmental uncertainty (Hwang, 2005).

The retail industry shows the market complexity associated with the dynamic consumption (Hristov and Reynolds, 2015; Moretta Tartaglione et al., 2019) (e.g., customers' preferance, technological orientation, increase in the role of communication, search for more variety in assortments, switch of purchases) and changes in retail offerings (Hackney et al., 2006; Alexander et al., 2008; Fornari, 2009; Mai et al., 2011) (e.g., technology evolution, increase in efficiency of partner networks, integration of on-line and off-line retail activities, new retail formats, customization of retail offerings); omni-channel retailing (Piotrowicz and Cuthbertson, 2014; Verhoef *et al.,* 2015); connectivity in the supply chain (Choi, 2001; Sivadasan et al., 2002; 2006; Surana et al., 2005; Pathak et al., 2007; Battini et al., 2007; Rensburg, 2012; Cheng et al., 2013; Nair et al., 2016; Nair and Reed-Tsochas, 2019; Zhao et al., 2019); an increasing variety of products and services (Mai et al., 2011) – e.g., virtual salespersons, interactive kiosks and displays, RFID (Radio Frequency Identification) systems connecting offerings with customers (Hackney et al., 2006; Pantano and Naccarato, 2010), and delivery choices; and transnational competition between different retailers in the global marketplace (Fernie and Arnold, 2002; Chase, 2013; 2016).

The retailer managers face difficulty in the decision-making due to the increase of the variety and variability of factors that have to be considered in the sales forecasting. In this work, the aim is to explore the type of relationships between the external, internal environmental drivers and sales. Nonlinear methods can support the analysis and the management of these kinds of scenario, functioning as flexible tools for gaining a clear picture of data variation (particularly important when linear methods are not representative), as opposed to the types of insights revealed through linear methods (Hofmann and Gavin, 1998; Weisberg, 2005; Dawson, 2014). Using the General Additive Model – GAM – (Hastie and Tibshirani*,* 1987;1990; James *et al.,* 2013) it is possible to generate both linear and nonlinear relationships between response and predictors. In general, traditional tools in retail management are based on historical analysis of data (Heshmaty and Kandel, 1985; Chu and Zhang, 2003; Hyndman and Athanasopoulos, 2014) using linear approaches. The purpose is to support retail management decisions with nonlinear methods.

Environmental drivers mixed with internal drivers are at the future base for managing the relationships between company and environment. To identify the external drivers it is necessary to exploit and select the irregular patterns in the supply chain. This research aims to reveal the patterns from the retail environment that can make the variation in effects on the result, taking the first step in the identification of the external variables and the following steps in recognizing the type of the relationships between the variables. It is precisely the functional and intentional concept of interaction that clarifies this comprehension of the significance of "non-linearity". The relationship between the organization (a company, for example) and its environment can be interpreted first under the interaction between its components, both internal and external.

This work contributes to the literature suggesting the use of additive models in which the linear and nonlinear function can apply together in the management and definition of forecasting strategies in retail chains. Specifically, from the management side, the model presented contributes to the interpretation of relationships between internal and external variables in their dynamics in the retail chain environment and benefits the determinants of the forecasting.

This research is organized as follows. First, a literature background is provided. Second, the methodology used in the explorative analysis is outlined and managerial considerations of the results are provided. In the final part, this research provides a discussion of the results, implications, and conclusions.

3.2 Literature background

3.2.1 Sales forecasting and environmental drivers

Sales forecasting is the important input for the planning process in the supply chain (Sagaert et al., 2018). The planning of demand for the company's products can be uncertain due to the uncertainty of the sales forecasting process. The environment can affect the availability of data about orders, shipments, and demand, the economic condition of the country and the level of competition, and the competitor's actions are factors that can influence the sales forecasting process (Mentzer et al., 2005; p. 260). Several managerial indicators (e.g., specific historical situations of the retailer, strategic decisions concerning expansion, consolidation of markets, crises) and environmental variables (e.g., geographical position of the retailer, the dynamism of consumption, regional activity, consumer sensitivity to technology, political influence, socio-demographic conditions) could affect the decision of which models to use (Moretta Tartaglione et al., 2019). For retail companies, the switch in use of traditional models to new approaches is challenging especially for retailers focused on traditional models of managing – retailers whose strategies are owner-managed, retailers limited in offering added services, and retailers with basic offerings and limited assortments. In stable markets and for small retail chains, the use of linear models for analyzing and managing data are preferred, particularly when detecting patterns between sales and internal and external variables (Pantano et al., 2017).

External drivers that are studied in the literature are varied according to the aim of the forecasting activities, but the common thing is to find the optimal combinations of added external patterns to get accurate predictions for the short or long term. In the literature, the common external drivers are related to the competition and industry concentration (Mintz and Currim, 2013), collaboration and relations inside the supply chain (Hubner et al., 2013; Albarune and Habib, 2015; Chase, 2013; 2016), competitive information (Huang et al., 2014), competition among the multilevel generations of the technological products (Tsai, 2013); shared information the financial crisis and market volatility (Kung and Chen, 2014; Everingham et al., 2016; Harrauer and Schnedlitz, 2016); the location of decoupling point, the socio-demographic profile of consumer and irregular demand and (Hubner et al., 2013; Prieto and Caemmerer, 2013; Sillanpää and Liesiö, 2018); the influence from the internet sources and online channels (Carrière-Swallow and Lebbe, 2010; Papanagnou and Matthews-Amune, 2018; Fildes et al., 2019); the big amount of sales data and high dimensionality of data (Ma et al., 2016; Ren et al., 2018); climate change and weather conditions (Arunraj and Ahrens, 2016; Everingham et al., 2016; Bertrand and Parnaudeau, 2017; Camacho and Conover, 2019; Verstraete et al., 2019).

Future sales are the critical point for the retail store together with minimizing the inventory stock. The traditional approach is to use the historical data based on the internal sales flows and promotional bumps. However, the inclusion of the multi-factors that can better reflect the internal and external environment is becoming the trend in the forecasting process. The retailer has emerged in the supply chain and collaborative forecasting is essential for the efficiency of satisfying the demand. The forecasting model including the information from the supply chain (Aviv, 2001, Trapero et al., 2013; Williams et al., 2014), then the inclusion of the promotional data from the diverse locations (Gur Ali, 2013; Aguilar-Palacios et al., 2019); numerous promotional data (Ma et al., 2016), including competitor's price and promotions (Huang et al., 2014); thus the forecasting needs the expert adjustments (Fildes et al., 2009; Baecke et al., 2017; Broeke et al., 2019). Human experts can benefit from pre-identifying relevant wide groups of indicators, and then a smaller set is forward identified using a statistical model.

3.2.2 Macroeconomic variables and sales

Sales forecasting as the management function can differ for the level, time horizon, and interval depending on the goals. The level of detail of the information refers to the design detail required. If forecasting is requested annually sales by a line of merchandise for a marketing plan, data at the identical level and time horizon are enough. However, if it is necessary to do weekly, annual line of products data are going to be of little support. Sales forecasts are needed for a variety of various functional plans, data at the extent of detail as if each of those planning needs is critical. This level of detail is called the forecasting hierarchy and is defined as all the planning levels and time horizons/intervals at which forecasts are needed (Mentzer et al., 2005; p.32). The level of the forecasting is depending on which functional unit is essential. For the marketing and finance function, long-term forecasting will be requested, thus for the sales function the operational levels will be required quarterly, weekly, or even daily. When there is planning for a great retail chain that has a huge portfolio of retail formats over the country and different regions the influence of the macro-environment can be essential for strategic or tactical decisions. Macroeconomic indicators can contain influential information, such as changing economic conditions, taxes, and rates according to the regional or country policies (Sagaert et al., 2018). Usually, companies are following the behavior of the national markets and price policies to adjust the forecasting of the price and demand. In this segment, the horizon of the forecast can have a crucial role as it is for the long or medium-term horizons. In this context, a piece of macroeconomic information is relevant.

In the literature, there are studies related to the influence of the fuel price on car sales (Pindyck and Rotemberg, 1983; Hamilton, 1988; Barber et al.,1999; Hamilton, 2011); on the several industries (Lee and Ni, 2002). The increase in the fuel price affects the supply and demand of the consumers. In the car industry, the increase in fuel price disturbs the supply chain and switches consumer preferences to smaller vehicles. In the research of Shahabuddin (2009) besides the fuel price, the effect of income level, interest rate, financial aggregate, and the unemployment rate is measured.

Inflation can increase the cost of living and redirect the resources from investment to consumption, the inflation can affect the chain of destabilization from the companies' profits to the dividends and stock, which can affect the macro variable CPI (consumer price index). Inflation has been seen thru the nominal interest rate and an increase in nominal interest rates increases the discount rate which results in a reduction of value of today and cash flows (Ray, 2012). High inflation affects corporate profits and dividends resulting in lower stock prices. CPI provides aggregate price changes of the sample of products and services taken from various retail outlets (housing, food and beverages, clothes, transportation, medical care, entertainments, and others). CPI could give information about competitors' prices. The transactions are still performed offline in the stores, the online retail price reflects the retail price and provides the accurate daily CPI along with the consumer price inflation in real-time (Cavallo, 2017). The changes in income affect consumption as well the unemployment as the reflection of the labor force as the socio-economic factor. This is associated with the personal habit of individual consumers as the use of products in the present context of the everyday. Research on the influence of the indicators on sales forecasting is limited (Sagaert et al., 2018). The indicator Unemployment is a reflection of the labor force –socio-economic factor and Fuel Price as a reflection of financial markets, Temperature as an atmospheric driver. However, the regression model including the macroeconomic indicators can suffer from heteroscedasticity that can affect the significance of the variables' effects of the model. The problem of heteroscedasticity can decrease the efficiency of the model as a whole.

3.2.3 External data and exogenous variables

The retail market is fluctuated due to many factors that are controllable and non-controllable. The controllable factors are internal variables as promotions that can significant affect the sales and which peaks occur in the predictable way (Thomassey, 2014). The effect of external variables can be seen in the sudden switch of the sales volume and those are hard to control. Macroeconomic variables can have the impact that influence globally on the sales fluctuations and besides the effect of the market regulations; weather conditions that are not easy to quantify. Trend is that companies are starting to use the external indicators in the forecasting models, seeing the benefit of adding them thus there is the lack of methods capable of measuring the external data (Sanders and Manrodt, 2003; Weller and

Crone, 2012). The capturing the external data has been still done by the expert judgment that comprise the forecasting accuracy. In the research of Thomassey (2014) author illustrates the practical guidelines of including the external indicators in the forecasting system: 1) there is no exact list of the external indicators that should be included, 2) the impact of these variables is challenging to estimate, as there is no constancy over time. 3) the external data can be correlated and that can affect modelling strongly 4) some external data as (weather conditions) are limited and not available and thus cannot be integrated in the forecasting system. The aim to include the exogenous variable is to demonstrate better the effect of endogenous variables. There are the events in which some economic variables show the deviation in certain intervals, and nonlinear behavior or the sharp exponential trend due to it becomes a problem in the practice to define what is the optimal number of the exogenous indicators (Gutiérrez et al., 1999).

To assess the exogenously it is necessary to include the instrumental variables that will be independent of the latent error (Guevara, 2018). The identification of the exogenous variables can be over-identified. In the case of the linear models, the test for the exogenous variables is developed by Sargan (1958) where the author uses the residuals of the instrumental variables to test the level of the exogenoity, while in nonlinear models or semi-parametric models e.g. general additive models (GAM) and principal component analysis can be used to clear up endogeneity by fitting smoothing splines (in case of o GAM) in which regularization parameter in spline functions help in optimizing the coefficients (to not overfit). Adding the different functions every time to predictors in the GAM back fitting algorithm applies to allow to fit independently with new predictor and putting the residuals as a response from previous fitted functions. The function will automatically yield smooth function estimates that take into account the non-linearity already present in the error term, and recover the residual amount of non-linearity needed to clear up the endogeneity of the endogenous variables in the model. Recently some authors have worked on a flexible instrumental variable approach (Coakley et al., 2002; Marra and Radice, 2011; Bai and Ng, 2013).

Causal relationships in the regression model can be simultaneous the case when causality runs not only from the explanatory variables () to the dependent but also vice versa (). The spillover of the mutual causality can be found in real events where the increase of demand can affect economic factors and the change of the economic factors can affect the demand. Cao and Song (2017) explained in the context of the influence of the tourism demand on the economic situation of the countries. The causality between the economic factors and tourist demand can be simultaneously and the variation in one variable can cause the variation in another variable so vice versa. The incoming tourists can affect the economy and press the demand for the destination's country, also the effect can prevail on other things as the external costs for the residents for the public transportation and thus the tourist's spending can increase the income for the transport, the spillover effect of the increased demand will be the increased inflation. Interdependence is forcing that some variables that should be treated as exogenous are treated as endogenous and that is causing the over-identification. There is the challenge to evaluate the independence between the instrumental variable and the latent error to measure the exogenoity (Guevara, 2018). The author suggests the minimum number of instrument variables be used to correct the endogeneity, like the one instrument variable per endogenous variable. When there is the over-identification with instrumental variables the residuals of the linear model can be used to test for the exogeneoity of the model. The test is known as the Sargan test developed by Sargan (1958). Sargan tests check the constraint overidentification. The principle is that the model is overidentified if you have more than one instrument per endogenous variable, and you have some excess details. For the inferences to be right, all of the instruments must be accurate. Therefore, it checks that all exogenous tools are truly exogenous, and are not associated with the residuals of the model. The tests have a limit and it shows the inconsistency related to the sample size (Newey, 1985; Guevara, 2018). Additional improvement in testing is that it should be chosen the subset of the instrumental variables as exogenous and then test the model. The leading indicators are often used in industry for the long-term forecasting horizons (Stock and Watson, 2001).

The correct setup of the variables for the modeling is essential in forecasting. In the selection of variables the two facts matter 1) not to omit the significant economic indicators 2) and to reduce the interdependence between the economic variables. There are a couple of approaches to select variables; it depends on the size of the dataset and the numbers of the variables. The most common techniques are stepwise regression, ridge, and lasso regression. These techniques are used in the research depending on the sample size, the stepwise regression is mostly used thus it can overfit and retain the unnecessary variable inside the model (Sargerat et al., 2018; Huang et al., 2014), unlike the ridge and lasso can shrink the coefficients close or exact to zero, that can in the case of the larger number of predictors then the observations can help retain the relevant predictors (Tibshirani, 1996). Ordinary least regression does not perform well in terms of the prediction accuracy producing unstable regression coefficients especially when the number of observations is less than the number of predictor variables. The instability of the parameter estimates is increasing when the predictors are highly correlated and that is the main issue in the relevance of the OLS models. Regard tests of multicollinearity: the presence of multicollinearity can result that the variable effects cannot be separated and extrapolation is likely to be seriously erroneous (Meloun, et al. 2002, p. 443). Change in data sets can produce multicollinearity that can influence the validation of the model on the test (as correlation may serve as an indicator of collinearity- naturally considering that macroeconomic variables might be highly correlated). Test for the multicollinearity in the linear model can be done by Variance Inflation Factor (VIF) diagnostics (Thompson et al., 2017). If there is a high correlation between variables VIF test will show high values in particular if VIF > 10, the results indicate multicollinearity. Besides VIF diagnostics advanced methods as principal component analysis (PCA) can be performed in a way to remove correlations in a variable set and to reduce collinearity (Stock and Watson, 1998; 2002; 2006). PCA produces orthogonal (i.e. perfectly uncorrelated) axes as output, so, the PC-axes may be used directly in subsequent analyses in place of the original variables' (Dormann et al., 2013). Using PCA is a strategy to deal with multicollinearity, not the only possible solution. If there are a high number of variables, the selection of the relevant variables can be challenging in that case PCA can reduce the high dimensionality of data and provide fewer inputs that are not collinear with each other also. The latent variables might be processed further to better interpret the effect of the original variables.

Table 3-1. Literature background of drivers affecting sales and demand forecasting

Research topics	Authors
Economic conditions, level of competition, data about orders shipments	Mentzerr et al., 2005; Mintz and Currim, 2013; Sagaert et al., 2018
Causality of economic conditions and tourism demand	Cao and Song, 2017
Strategic decisions, environmental drivers	Moretta Tartaglione et al., 2019
Relations inside the supply chain	Aviv, 2001; Hubner et al., 2013; Trapero et al., 2013; Williams et al., 2014 ; Albarune and Habib, 2015; Chase, 2013; 2016
Market volatility, financial crisis	Kung and Chen, 2014; Everingham et al., 2016; Harrauer and Schnedlitz, 2016
Socio-demographic profile of consumer	Hubner et al., 2013; Prieto and Caemmerer, 2013; Sillanpää and Liesiö, 2018
Climate change and weather conditions	Arunraj and Ahrens, 2016; Everingham et al., 2016; Bertrand and Parnaudeau, 2017; Camacho and Conover, 2019; Verstraete et al., 2019
The influence of the fuel price on car sales	Pindyck and Rotemberg, 1983; Hamilton, 1988; Barber et al.,1999; Shahabuddin, 2009; Hamilton, 2011
Inflation, Consumer Price Index (CPI), Unemployment rate	Ray, 2012; Cavallo, 2017; Moretta Tartaglione et al., 2019
Competition's price and promotional information	Huang et al., 2014
Lack of methods capable of measuring the external data	Sanders and Manrodt, 2003; Weller and Crone, 2012; Sagaert et al., 2018
High dimensionality of sales data	Ma et al., 2016; Ren et al., 2018
Promotions	Gur Ali, 2013; Thomassey, 2014; Ma et al., 2016 ; Aguilar-Palacios et al., 2019
Online channels	Carrière-Swallow and Lebbe, 2010; Papanagnou and Matthews-Amune, 2018; Fildes et al., 2019
Expert adjustments	Fildes et al., 2009; Baecke et al., 2017; Broeke et al., 2019

Source: Author's elaboration

3.3 Methodology

3.3.1 Linear, nonlinear and additive models

Given the variability of relationships (Moretta Tartaglione et al., 2019) within the retail sector (Dawson, 2000), the use of nonlinear methods of analysis is becoming more common in the fields of economics (Guo, 2017) and retail management, e.g., studies focused on the relationships between promotion, price and sales (Christen *et al.,* 1997; Haupt and Kagerer, 2012; Moretta Tartaglione et al., 2019;), and weather and sales (Parsons, 2001; Murray *et al.,* 2010; Bahng and Kincade, 2011; Arunraj and Ahrens, 2016; Bertrand and Parnaudeau, 2017).

To address gaps in the literature, Grewal and Levy (2007) suggest more and deeper research on the role of multi-layered approaches, including multi-scalar, multi-method, and combining quantitative and qualitative approaches. The need for sustainability in the measurement of complexity influenced by the environment in retail is suggested by Harrauer and Schedlitz (2016); the case analysis by Yu and Ramanathan (2012) has further suggested that business environment has a substantial impact on retail operations strategy (such as low cost, quality, flexibility, and delivery); and Alon *et al.* (2001) point to the necessity for advanced methods that can capture dynamic nonlinear trends and seasonal patterns in retail sales. Nonlinear analysis is now a part of the theory found in sales literature (Johnson, 2014).

Linear methods, used primarily in the estimation of sales, prices, and promotions, can provide only limited information about market mechanisms for multiple reasons. In particular, linearly-derived estimations can be affected by the inherent entanglement of different strategies applied to each lever – sales, price, promotions. Although linear methods are used by different retailers to predict sales and study market dynamics, the linear nature of these models can limit the analysis and managerial considerations (Haupt and Kagerer, 2012).

Nonlinear analytical approaches assume that relationships between variables can take many forms, and these approaches do not constrain the estimate by requiring precisely linear associations. Compared to linear, nonlinear analysis permits heterogeneity in the rate of change in a dependent variable for various levels of an independent variable. In fact, increases in an independent variable may have a positive effect on a dependent variable at one level and a negative effect at a different level (Johnson, 2014). Outlining the nonlinear theory (Wecker and Ansley, 1983; Hastie and

Tibshirani, 1987;1990; Green and Silverman, 1993; James et al., 2013; Watts and Bates, 2014; Hastie, 2017), it is stated that non-parametric regression typically assumes little about the shape of the regression function (beyond some degree of smoothness). Such tools lead to more durable inferences and seem to be relevant to every managerial activity.

GAM represents advanced nonlinear tools (Hastie and Tibshirani, 1987; 1990; Guisan *et al*., 2002; James *et al*., 2013; Hastie, 2017; Wood, 2017) for fitting smooth curves on data. GAM uses the function to create the relationship between the mean of the response variable and a "smoothed" of the explanatory variable(s) (Guisan *et al*., 2002), having the power to deal with highly nonlinear and non-monotonic relationships between the response and explanatory variables. GAM estimation technique of 'backfitting' is classy and has the advantage that it allows the component functions of an additive model to be expressed using any smoothing or modeling technique (Wood, 2017; p.318). The GAM method reflects a non-parametric perspective that says, "let the data determine the shape" of the response curve.

To identify a function in GAM (James *et al.,* 2013), it is possible to replace linear component $\beta_j x_{ij}$ in the multiple linear regression model

$$y_i = \beta_0 + \beta_1 x_{i1} + \beta_2 x_{i2} + \beta_3 x_{i3} + \cdots + \beta_p x_{ip} + e_i$$

Equation 3.1

with a (smooth) nonlinear function $f_j(x_{ij})$ obtaining the model formula below

$$\begin{aligned} y_i &= \beta_0 + \sum_{j=1}^{p} f_j(x_{ij}) + e_i \\ &= \beta_0 + f_1(x_{i1}) + f_2(x_{i2}) + f_3(x_{i3}) + \cdots + f_p(x_{ip}) + e_i \end{aligned}$$

Equation 3.2

GAM permits the calculation to a separate for each and then add together all the contributions. This method is relevant to the use of building burdens to fit an additive model. The smoothing parameter, λ, controls the trade-off between the smoothness of the estimated and trust to the data. $\lambda \to \infty$ leads to a straight line estimate, while $\lambda = 0$ results in a non-penalized piecewise linear regression estimate (Wood, 2017; p. 167).

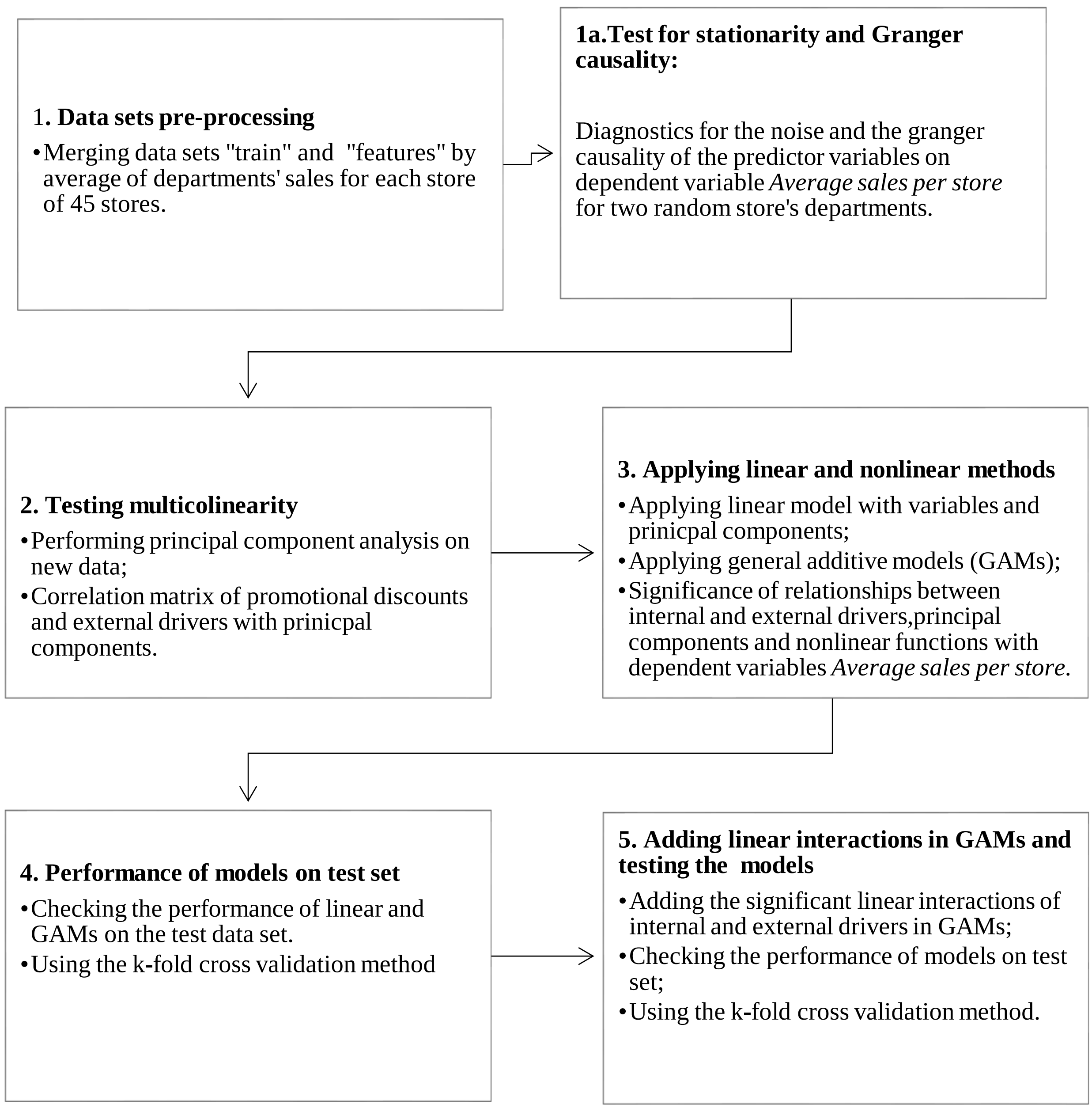
The relevance of GAM application has been shown in different research domains from medical research (Hastie and Tibshirani, 1995) over research in ecology (Yee and Mitchel, 1991; Guisan *et al*., 2002) and very few marketing-related studies (Kalyanam, 1998; Kumar *et al*., 1998; Shively, 2000; Coussement *et al*., 2010). In the research of Kalyanam (1998) stochastic spline regression is used for estimation of irregular price effects in certain product categories, while in research of Kumar *et al.* (1998), using spline regression the authors find the asymmetry of car dealers' punitive actions toward their key suppliers. Coussement *et al*. (2010) have shown that GAM improves marketing decision-making by identifying risky customers and increases the interpretability of the customer churn model by visualizing the nonlinear relationships. This work presents the research and results in two parts: (1) the collection and explorative analysis of the data and (2) discussion of results and implications for the retail sector.

The exploratory analysis starts with data interpretation, testing the stationarity and causality, modeling, and, at final, a validation of linear and nonlinear models.

3.2 Data collection and analysis

The data which include the two data sets come from 45 retail stores located in different regions of the US; one data set represents historical sales data (train) with 421,570 observations and five variables (Store, Department, Date, Weekly Sales, IsHoliday), and the other data set (features) contains 8,190 observations and twelve variables (Store, Date, MarkDown1, MarkDown2, MarkDown3, MarkDown4, MarkDown5, Consumer Price Index, Unemployment, Fuel Price, Temperature, IsHoliday).

Figure 3-1. The process of data analysis



Source: Author’s elaboration

In the first part of analysis it has been performed the test for the stationarity and Granger causality.

Stationarity means the consistency of data over time and the Dickey-Fuller test can measure the presence of unit roots (Dickey and Fuller, 1979:1981). Granger test (Granger, 1988) has done used on the, randomly, two departments and their historical data. The concept of Granger causality relates to the fact the variable *X* is said to Granger-cause another variable, *Y*, if the current value of $Y\,(y_t)$ is dependent on the past values of $X\,(x_{t-1}, x_{t-2}, \ldots, x_0)$ then the history of X is likely to predict *Y*. Therefore, the results from Granger causality show that some variables influence sales (linearly-Granger) in one period but not in another. Granger-causality test Granger-causality test can decide whether the relationships among variables are causal, and it can also distinguish exogenous variables from all variables. In the case of a no-cointegration relationship among variables in the model, the Granger theory can be tested for a judgment of endogenous and exogenous variables in the VAR and VECM. F-Statistic is used to do Granger-causality test in this paper showing tests if the *p-value* is less than 5%, the null hypothesis must be accepted, that is, there is a causal relationship among tested variables. Otherwise, the null hypothesis is rejected.

The point of this work is not to study the change, causality, and relationships over time as there is missing information from internal drivers (markdowns) in 2010 and 2011 so uncoordinated data by the time data can unravel the unreliable effects of the external drivers. In this sense, the causality test lost its applicability. The dependent variable *Y* after the reshaping data is the averaged sum of sales by the store.

In the data, pre-processing *Date* (time-variable) is omitted by averaging and summing all departments' sales by store. The point was to arrange/sum/reduce data taking the upper-level perspective of the retail chain for 45 stores and the influence of internal and external factors on the chain's sales.

In the first phase, the data sets to train and features are adapted and merged. Department sales for each of the 45 stores are averaged and looked at in terms of the available information on internal and environmental drivers, to explore the nature of the relationships outside of time. For this reason, Principal Component Analysis (PCA) is applied to eliminate the moderate correlations (Dormann *et al.,* 2013) between the promotional discounts (MarkDown (1-5)) and environmental drivers (CPI, Unemployment, Fuel Price, Temperature, IsHoliday).

To compare linear and nonlinear models, the linear approach is applied to data; backward fitting (Derksen and Keselman, 1992; Baumann *et al.,* 2012) is used to reveal the main effects of key variables on the dependent variable (Average sales per store). The same set of key variables (internal and external drivers) are used in nonlinear modeling with the application of GAM, advanced nonlinear tools for fitting smooth curves onto data. In addition to analysis, the important linear interactions between internal and environmental drivers are added in GAM. In the final phase, the performance of four models is tested with k-fold cross-validation.

The data used – from the period 2010-2013 – come from a confidential research project that was published on a US retailer's website in 2014. Specific rules on the privacy and integrity of the project mean that the source of the data must be kept confidential.

The historical data set (train) covers weekly sales from 5th February 2010 to 1st November 2012 and has the following fields: Store (the store number), Dept (the department number), Date (the week), Weekly_Sales (sales for the given department in the given store), and IsHoliday (whether the week was a special holiday week).

The data set features contain additional information related to the store, department, and regional activity for the given dates. It contains the following fields: Store (the store number), Date (the week), Temperature (average temperature in the region), Fuel_Price (cost of fuel in the region), MarkDown1 to 5 (anonymized data related to promotional markdowns – promotional discounts – offered by the retail store), CPI (the consumer price index), Unemployment (the unemployment rate), and IsHoliday (whether was the public holiday or no on that day).

3.3.2 Test for stationarity and Granger causality

On the randomly chosen two stores departments' sales data, it has been presented the descriptive statistics tabular and graphically (Table 3-2 and Figure 3-2), then the Dickey-Fuller test for stationarity and granger causality (Table 3-3 and Table 3-4). The goal is to provide a brief insight into raw data before modifying data sets for the paper. Looking at the dynamics of the variables at Figure 3-2 it shows that Weekly Sales are stable over time with seasonal effects in the first and final quarter and line with the applied promotional discounts Markdown 1-5, Fuel Price characterizes the increasing dynamics with up (I and II quarters) and down (III and IV quarters) fluctuations, while CPI is increasing linearly with fluctuations in the first year. Temperature is changing following the seasons, with lower temperatures in the first and fourth quarters and higher temperatures in the second and third quarters.

After applying the test for stationarity and granger causality it emerges results for *Department 1/Store 1 (*Table 3-3): time series of *Weekly Sales, MarkDown1, MarkDown3, MarkdDown4, MarkDown5,* and *Temperature* is stationary at level, *CPI*, *Fuel Price*, and *Unemployment* are stationary at the order of integration 1, and *MarkDown2* is stationary at the order of integration 2. In transformed variables the correlogram in the case of variable *CPI* ACF (autocorrelation) dies out after the 4th lag but not totally, and PAC (partial correlation) has a large spike in lag 18 and 35 and dies out oscillating, for *Fuel Price* ACF dies out after the 2nd lag, and PAC has a spike in lag 32nd, *Unemployment* has large spikes in 13th and 27th lag, and in the case of *MarkDown2* AC dies out after the 2nd lag, and PAC has large spikes in lags 2, 3,4, 5, 6 and 7.

Regard the granger causality test (143 weeks) (Table 3-4): *MarkDown1, d(Fuel_Price), MarkDown4 and Temperature* granger cause *Weekly Sales*, and *Weekly Sales* granger caused (*MarkDown2*, 2)-(prob 0.08).

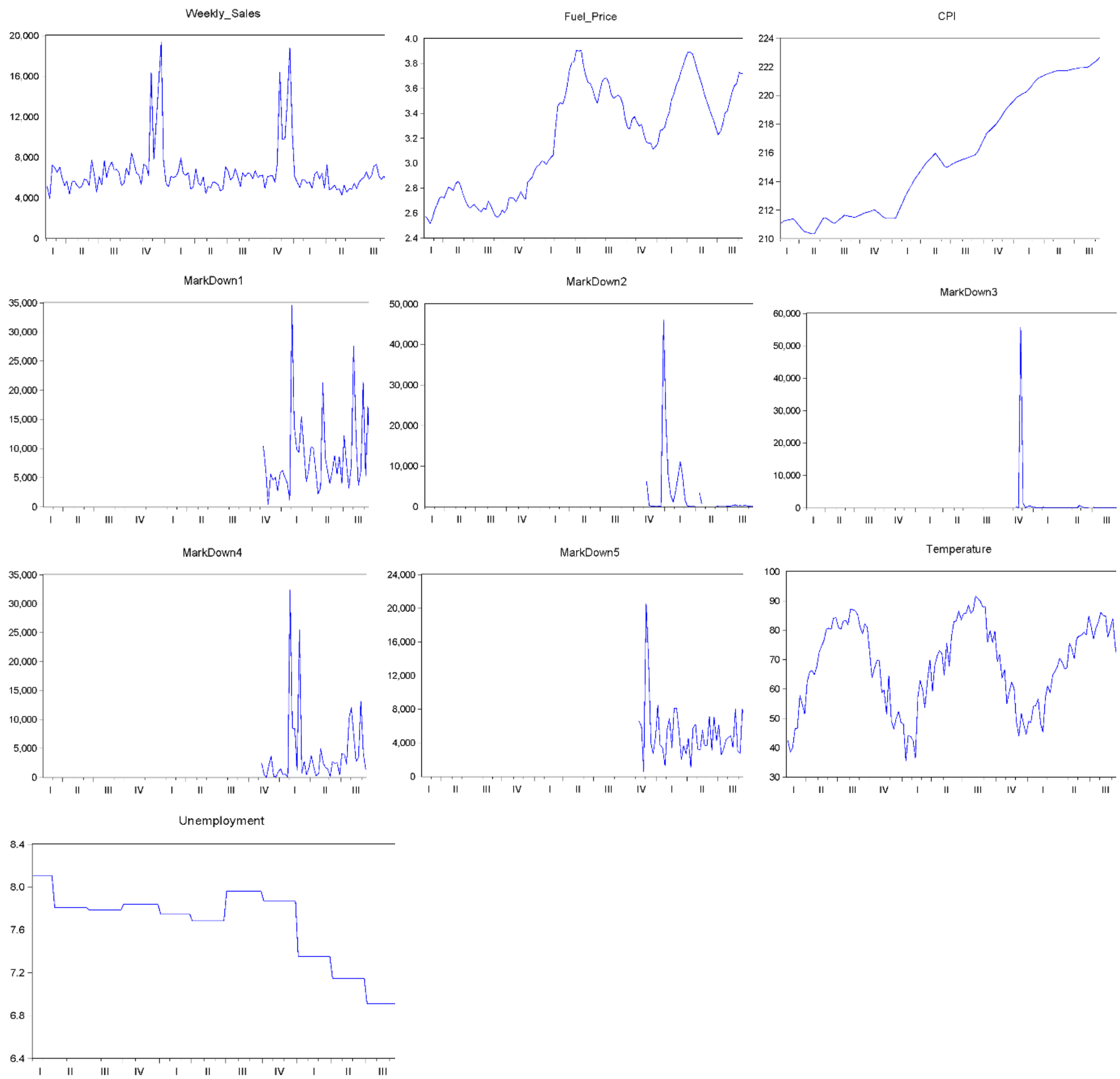
Table 3-2. Descriptive statistics for department sales (1) at the store (1)

Department 1 at Store 1/ Descriptive statistics

	CPI	FUEL PRICE	MARKDOWN1	MARKDOWN2	MARKDOWN3	MARKDOWN4	MARKDOWN5	TEMPERA...	UNEMPLO...	WEEKLY ...
Mean	215.9969	3.219699	8090.766	2941.315	1225.400	3746.085	5018.655	68.30678	7.610420	22513.32
Median	215.4599	3.290000	6154.140	144.8700	25.96500	1822.550	4325.190	69.64000	7.787000	18535.48
Maximum	223.4443	3.907000	34577.06	46011.38	55805.51	32403.87	20475.32	91.65000	8.106000	57592.12
Minimum	210.3374	2.514000	410.3100	0.500000	0.250000	8.000000	554.9200	35.40000	6.573000	14537.37
Std. Dev.	4.350890	0.427313	6606.896	7955.538	7879.964	5999.641	3281.845	14.25049	0.383749	9854.349
Skewness	0.265983	-0.151813	2.060335	4.255270	6.847344	3.257523	2.463007	-0.402643	-1.046172	1.992864
Kurtosis	1.547967	1.600071	7.776051	22.30328	47.93170	14.35968	11.77756	2.114668	3.043197	6.202758
Jarque-Bera	14.24867	12.22644	84.55501	778.8301	4596.671	364.4123	215.2862	8.534116	26.09612	155.7728
Probability	0.000805	0.002213	0.000000	0.000000	0.000000	0.000000	0.000000	0.014023	0.000002	0.000000
Sum	30887.56	460.4170	412629.1	123535.3	61270.02	191050.3	255951.4	9767.870	1088.290	3219405.
Sum Sq. Dev.	2688.095	25.92863	2.18E+09	2.59E+09	3.04E+09	1.80E+09	5.39E+08	28836.84	20.91137	1.38E+10
Observations	143	143	51	42	50	51	51	143	143	143

Source: Author’s elaboration

Figure 3-2. Weekly time series for department sales (Department 1/Store 1)



Source: Author's elaboration

Table 3-3. Dickey-Fuller test for stationarity (Department 1/Store 1)

Variable	Order of integrity	Lag length	Dickey- Fuller test/t statistics	Prob.
Weekly sales	Unit root	1	-7.619	0.0000
CPI	2	1	-4.365	0.0005
Fuel price	2	0	-6.364	0.0000
MarkDown1	Unit root	0	-6.478	0.0000
MarkDown2	2	9	-5.737	0.0004
MarkDown3	Unit root	0	-6.772	0.0000
MarkDown4	Unit root	0	-6.344	0.0000
MarkDown5	Unit root	0	-6.009	0.0000
Temperature	Unit root	10	-4.837	0.0001
Unemployment	2	0	-12.094	0.0000

Source:Author's elaboration

Table 3-4. Granger causality test (Department 1/Store 1)

Pairwise Granger Causality Tests Sample: 2/05/2010 10/26/2012 Lags:4			
Null Hypothesis:	**Obs**	**F-Statistic**	**Prob.**
D(CPI) does not Granger Cause WEEKLY SALES WEEKLY SALES does not Granger Cause D(CPI)	138	1.61928 0.46875	0.1733 0.7586
D(FUEL PRICE) does not Granger Cause WEEKLY SALES WEEKLY SALES does not Granger Cause D(FUEL PRICE)	138	2.07076 1.03448	0.0883 0.3921
MARKDOWN1 does not Granger Cause WEEKLY SALES WEEKLY SALES does not Granger Cause MARKDOWN1	47	2.39102 0.28317	0.0677 0.8871

D(MARKDOWN2,2) does not Granger Cause WEEKLY SALES WEEKLY SALES does not Granger Cause D(MARKDOWN2,2)	27	0.34979 2.42318	0.8407 0.0859
MARKDOWN3 does not Granger Cause WEEKLY SALES WEEKLY SALES does not Granger Cause MARKDOWN3	42	1.82378 0.20686	0.1477 0.9328
MARKDOWN4 does not Granger Cause WEEKLY SALES WEEKLY SALES does not Granger Cause MARKDOWN4	47	2.59944 0.76540	0.0513 0.5544
MARKDOWN5 does not Granger Cause WEEKLY SALES WEEKLY SALES does not Granger Cause MARKDOWN5	47	1.20413 1.52816	0.3250 0.2135
D(UNEMPLOYMENT) does not Granger Cause WEEKLY SALES WEEKLY SALES does not Granger Cause D(UNEMPLOYMENT)	138	0.24897 1.41060	0.9098 0.2341
TEMPERATURE does not Granger Cause WEEKLY SALES WEEKLY SALES does not Granger Cause TEMPERATURE	139	8.54678 0.84912	4.E-06 0.4966

Source: Author's elaboration

In the case of *Department 22/Store 25* results emerge (Table 3-5): time series *Weekly Sales*, *MarkDown1, MarkDown3, MarkdDown4, MarkDown5, and Unemployment* are stationary at level, *CPI, Fuel Price, MarkDown2 and Temperature* are stationary at order of integration 1, in transformed variables based on the correlograms, in the case of *CPI* ACF dies out after the 3rd lag partially, and PAC with large spike in lag 18^{th} and dies out oscillating, for *MarkDown2* ACF dies out after 2^{nd} lag and PAC has a spike in 6^{th} lag, Fuel Price ACF dies out after the 2^{nd} lag., and in the case of *Temperature* ACF is oscillating in norms.

Regard the granger causality test (139 weeks) (Table 3-6): *MarkDown3, MarkDown5 granger cause Weekly Sales* and *Weekly Sales* granger cause *Unemployment* and there is a bi-directional causality between *Weekly Sales* and d(*MarkDown2).*

Table 3-5. Dickey-Fuller test for stationarity (Department 1/Store 1)

Department 22/Store 25

Variable	Order of integrity	lag length	Dickey- Fuller test/t statistics	Prob.
Weekly sales	Unit root	0	-6.324	0.0000
CPI	2	1	-4.389	0.0005
Fuel price	2	0	-5.042	0.0000
MarkDown1	Unit root	0	-4.950	0.0002
MarkDown2	2	9	-7.158	0.0001
MarkDown3	Unit root	3	-19.718	0.0001
MarkDown4	Unit root	0	-4.851	0.0002
MarkDown5	Unit root	0	-5.590	0.0000
Temperature	2	0	-16.295	0.0000
Unemployment	1	0	-2.883	0.0498

Source: Author's elaboration

Table 3-6. Granger causality (Department 25/Store 22)

Pairwise Granger Causality Tests Sample: 2/05/2010 10/26/2012 Lags:4			
Null Hypothesis:	**Obs**	**F-Statistic**	**Prob.**
D(CPI) does not Granger Cause WEEKLY SALES WEEKLY SALES does not Granger Cause D(CPI)	138	0.10704 1.27449	0.9799 0.2834
MARKDOWN1 does not Granger Cause WEEKLY SALES WEEKLY SALES does not Granger Cause MARKDOWN1	47	0.11508 0.59638	0.9764 0.6674

D(MARKDOWN2) does not Granger Cause WEEKLY SALES WEEKLY SALES does not Granger Cause D(MARKDOWN2)	26	3.38829 33.5103	0.0327 7.E-08
MARKDOWN3 does not Granger Cause WEEKLY SALES WEEKLY SALES does not Granger Cause MARKDOWN3	32	41.4059 0.79222	3.E-10 0.5423
MARKDOWN4 does not Granger Cause WEEKLY SALES WEEKLY SALES does not Granger Cause MARKDOWN4	47	0.24348 0.22879	0.9118 0.9205
MARKDOWN5 does not Granger Cause WEEKLY SALES WEEKLY SALES does not Granger Cause MARKDOWN5	47	7.70645 0.64217	0.0001 0.6357
D(TEMPERATURE)does not Granger Cause WEEKLY SALES WEEKLY SALES does not Granger Cause D(TEMPERATURE)	138	1.54179 0.78579	0.1939 0.5364
UNEMPLOYMENT does not Granger Cause WEEKLY SALES WEEKLY SALES does not Granger Cause UNEMPLOYMENT	139	0.53733 2.50155	0.7085 0.0455
D(FUEL PRICE) does not Granger Cause WEEKLY SALES WEEKLY SALES does not Granger Cause D(FUEL PRICE)	138	0.93845 1.40283	0.4440 0.2367

Source: Author's elaboration

3.4 Results from the linear and general additive models (GAMs)

In this section, the results from the exploratory analysis are presented with a focus on the complete representations of applied models and their comparison. The practical analysis is presented within the next steps:

1. Applying the linear method to the data using the full set of promotional discounts (MarkDown1-5), external drivers (CPI, Unemployment, Fuel Price), and dummy variable (IsHoliday);
2. Applying the principal component analysis to reduce the multicollinearity;
3. Applying the non-linear methods (spline functions) to the promotional discounts in a GAM
4. According to the results from two previous steps, combining the linear and non-linear methods using the full set of promotional discounts and external drivers in a GAM model;

5. Including the linear interaction between internal drivers and external drivers and promotional discounts in GAM model;
6. Comparing the performances of all models on the test set by the k fold cross-validation method.

3.4.1 Linear model

Using data mining, department sales are averaged for each store in the historical data set (train) and merged with the features data set, where information concerning promotional discounts, CPI, Temperature, Fuel_Price, and Unemployment are randomly missing. After a backward fitting meant to reveal the most significant set of variables (removing the insignificant variable MarkDown4), the linear model (Table 3-7) is applied, relating the dependent variable Average sales per store to internal drivers MarkDown1, MarkDown2, MarkDown3, MarkDown5 and environmental drivers Temperature, Fuel Price, CPI, Unemployment, and IsHoliday. To assess multicollinearity among the variables, the Variance Inflation Factor (VIF) is computed; a VIF value of approximately 1 indicates the complete absence of collinearity. The relationships between the dependent variables and internal and environmental drivers show strong significance ($p \leq 0.01$), but the figure for R^2 (R-squared) is 0.1591, which shows a very low variability. This means 16% of Average sales per store can be explained by the chosen variables.

Table 3-7. Results of the linear model (LM)

Dependent variable: Average sales per store					
Coefficients:	*Estimate*	*Standard error*	*t-value*	*p-value*[3] *(>|t|)*	*Variance Inflation Factor (VIF)*
(Intercept)	1.391e+04	2.784e+03	4.997	6.47e-07 ***	
MarkDown1	1.630e-01	1.453e-02	11.215	< 2e-16 ***	1.12
MarkDown2	1.909e-01	1.777e-02	10.744	< 2e-16 ***	1.33

[3] The *p*-value is used in the context of testing in order to quantify the idea of the statistical significance of the evidence.

MarkDown3	1.229e-01	1.369e-02	8.979	< 2e-16 ***	1.36
MarkDown5	2.202e-02	7.806e-03	2.821	0.004853 **	1.00
Temperature	6.766e+01	9.558e+00	7.080	2.18e-12 ***	1.26
Fuel_Price	1.490e+03	6.867e+02	2.003	0.030220 *	1.49
CPI	2.202e-02	4.643e+00	-5.190	2.62e-11 ***	1.39
Unemployment	-6.879e+02	1.024e+02	-6.715	1.98e-15 ***	1.19
IsHolidayTRUE	-4.206e+03	5.908e+02	-7.120	1.64e-12 ***	1.54
Residuals: *Min 1Q Median 3Q Max* *-14192.5 -4889.7 -489.9 4172.8 16519.9*					
*Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1* *Residual standard error: 6202 on 1563 degrees of freedom* *Multiple R-squared: 0.1639, Adjusted R-squared: 0.1591* *F-statistic: 34.05 on 9 and 1563 DF, p-value: < 2.2e-16*					

Source: Summary results of the linear model (LM) (Moretta Tartaglione et al., 2019; p. 459)

From Table 3-6 it is possible to read common behaviours in the variables' relationships. In particular, markdowns have a high impact on sales and are able to generate positive effects. It is the same with Temperature and Fuel Price; in particular, Temperature has a high positive impact on sales, probably due to the characteristics of the stores and seasonality of the purchases. Table 3-6 also shows that an increase in Unemployment creates a reduction in sales. The problem arises with the significance of the model expressed by R^2 because 16% of variations are weak – insignificance – in the interpretation of the dependent variable Average sales per store by internal and external drivers.

3.4.2 General Additive Models (GAM)

GAM provides nonlinear relationships that standard linear regressions miss. In Table 3-9, relevant GAM is presented within the applied nonlinear functions to internal and environmental drivers. Besides, PCA is applied (Table 3-8) to eliminate the moderate correlation between promotional discounts and environmental drivers (CPI, Unemployment, Fuel Price, and Temperature). Results of the PCA demonstrate the correlations (higher than 0.5) of MarkDown1, MarkDown4, Fuel_Price and CPI within the first principal component and the correlations of Fuel_Price, MarkDown4, MarkDown1, CPI and Unemployment within the second. The first four principal components (specifically PCA scores) showing an eigenvalue higher than 1 – including high correlations of promotional discounts and environmental drivers (CPI, Unemployment, Fuel Price and Temperature) – are used in the nonlinear modeling.

Table 3-8. Results from PCA – principal components with assigned variables

Correlations of variables with principal components					
Variables	*PC1*	*PC2*	*PC3*	*PC4*	*PC5*
MarkDown1	0.7579	0.5616	0.0909	0.0486	-0.0020
MarkDown2	-0.1572	0.1949	-0.7197	-0.2647	-0.1093
MarkDown3	-0.1413	-0.1388	-0.0688	0.8704	0.3465
MarkDown4	0.7342	0.5883	0.0929	0.1206	0.0170
MarkDown5	0.1033	-0.0374	-0.0084	-0.3713	0.9189
Temperature	-0.0037	-0.2636	0.7777	-0.1451	-0.0539
Fuel_Price	0.5648	-0.6312	0.1177	-0.0267	-0.1192
CPI	-0.5046	0.4369	0.4687	-0.1068	0.0177
Unemployment	0.4810	-0.5544	-0.1790	-0.0935	0.0041
Eigenvalue	*1.97*	*1.68*	*1.41*	*1.024*	*0.99*

Source: PCA of the set of promotional discounts and environmental drivers (Moretta Tartaglione et al., 2019; p. 459)

Three relevant GAMs out of the eight simulated are shown in Table 3-9 with the different variables and the nonlinear transformation – smoothing splines – of the internal and environmental drivers. The goal is to compare: the GAM1/PC with principal components (to test linear relationships with Average sales per store); GAM2 with the main set of variables (to test the nonlinear relationships with Average sales per store); and GAM3 with the main set of variables and with the significant interaction (to test combinations of linear and nonlinear relationships with Average sales per store). Table 3-9 presents a summary of the variables used, the significance of parametric and non-parametric relationships, the explanatory power (R^2) of the model and the Root Mean Squared Error (RMSE) performed on the test set.

Table 3-9. Results from General Additive Models (GAMs)

Models / ***Internal and Environmental drivers (x)***	***GAM1/PC***	***GAM2***	***GAM3***
Intercept β_0	17938.1***	1.22***	-15531.891 ***
MarkDown1		s(***)	s(***)
MarkDown2		s(***)	s(***)
MarkDown3		s(***)	s(***)
MarkDown5		s(**)	s(**)
CPI		s(***)	s(***)
Temperature		s(***)	s(***)
Fuel Price		s(***)	s(***)
Unemployment		s(***)	s(***)
IsHoliday (*dummy variable*)		-2.80	-2156.744 ***
CPI: Unemployment			25.952***
PC1	924.8***		
PC2	250.6.		
PC3	-358.9*		
PC4	369.1*		
Adjusted R^2	0.03	0.42	0.46
Root Mean Squared Error /k-fold cross validation	4997.89	5127.17	4960.23

*** p-value ≤0.001 significance of parametric (linear) relationship
** p-value ≤0.01
* p-value ≤0.05
. p-value ≤0.1

s(***) p-value ≤0.001 significance of non-parametric (nonlinear) relationship
s(**) p-value ≤0.01
s(*) p-value ≤0.05
s(.) p-value ≤0.1

Root Mean Squared Error –the standard deviation between predicted and observed value on the test set

Source: Summary results of GAM (Moretta Tartaglione et al., 2019; p. 460)

To improve the significance of coefficients in GAM, smooth parameters are introduced (Moretta Tartaglione et al., 2019). For this reason, the smoothing spline functions are fitted to eight internal and environmental drivers: MarkDown1, MarkDown2, MarkDown3, MarkDown5, CPI, Temperature, Fuel Price, and Unemployment. As seen in Table 3, coefficients of non-parametric relationships are significant – s(***) – underlining the presence of nonlinear relationships between sales and internal and environmental drivers. Compared to the coefficients of linear functions, nonlinear coefficients are numerous and diverse for each effective degree of freedom of spline function; advantage of the smoothing spline coefficients is the penalty term inside of the function that tend to make the function smooth but not too flexible to avoid the overfitting of data (Hastie and Tibshirani, 1986; 1990; Guisan *et al*., 2002; James *et al*., 2013; Hastie, 2017; Wood, 2017). Penalty term is chosen by the cross validation method and in that way controls the effective degree of freedom and roughness of the smoothing spline function (James *et al*., 2013; Wood, 2017).This makes more complex nonlinear coefficients, hence their presentation and interpretation are not easy (Zelner, 2009). For this reason, a graphical presentation seems suitable for the interpretation of the model GAM3 (Figure 1a, 1b and 1c), which has provided the best results thus far (Table 3-9).

The R^2 of the nonlinear model, in particular, GAM3 (Table 3-9), shows that 46% of variation in the dependent variable Average sales per store can be explained by internal and external drivers and linear interactions between CPI and Unemployment. The R^2 (46%) in GAM3 is more significant than R^2 GAM1/PC (3 %) and GAM2 (42%).

3.4.3 *Linear interactions in GAM model*

In an exploratory analysis, it has been tested if a linear relationship between environmental drivers changes the effect on dependent variable Average sales per store, together with the performance of GAM models (Table 3-9). The strength of GAMs is their ability to deal with high non-linear and non-monotonic relationships between the response and the set of explanatory variables. GAMs are referred to as data- rather than model-driven as data determine the nature of the relationship between the response and the set of explanatory variables rather than assuming some form of parametric relationship (Yee and Mitchell, 1991).

Interaction term explains if the relationship between the interacting variables and Average sales per store depends on the value of the other interacting variable.

In Table 3-10 the demonstration of the following general additive models with interaction terms:

a) GAM4- non - linear and linear functions of internal and environmental drivers and interaction *Fuel Price: CPI*
b) GAM5- non - linear and linear functions of internal and environmental drivers and interaction *CPI: Unemployment*
c) GAM6- non - linear and linear functions of internal and environmental drivers and interaction *Fuel Price: Unemployment*
d) GAM7- non - linear and linear functions of internal and environmental drivers and interaction *MarkDown2: CPI*
e) GAM8- non - linear and linear functions of internal and environmental drivers and interaction *MarkDown5: ISHolidayTRUE*
f) *GAM9*- non - linear and linear functions of internal and environmental drivers interactions *Fuel Price: MarkDown1, Fuel Price: MarkDown2, Fuel Price: MarkDown5*

Table 3-10. Statistical testing of general additive models with linear interactions between promotional set and environmental drivers

Intenall and Environmental drivers (x)/Models	*GAM4*	*GAM5*	*GAM6*	*GAM7*	*GAM8*	*GAM9*
Intercept β_0	5.75	5.74	1.15	1.20	1.21	1.65
Temperature	6.74 s(***)	5.4 s(***)	6.12 s(***)	6.13 s(***)	5.98 s(***)	6.28 s(***)
Fuel Price	-9.71 s(***)	3.22 s(***)	3.16 s(***)	3.14 s(***)	3.08 s(***)	1.72 s(***)
MarkDown1	7.48 s(***)	6.71 s(***)	2.95 s(***)	7.46 s(***)	8.83 s(***)	-2.18 s(***)
MarkDown2	1.22 s(***)	1.11 s(***)	1.52 s(***)	1.10 s(***)	1.34 s(***)	-2.87 s(***)
MarkDown3	9.80 s(***)	9.52 s(***)	1.28 s(***)	9.71 s(***)	1.07 s(***)	1.26 s(***)

MarkDown4	4.99 no	8.04 no	3.29 no	3.26 no	9.14 no	6.76 no
MarkDown5	2.04 s(**)	1.95 s(**)	1.80 s(**)	1.93 s(**)	-3.73 s(**)	6.10 s(*)
CPI	-2.78 s(***)	-3.36 s(***)	1.86 s(***)	-1.96 s(***)	-1.91 no	-1.88 no
Unemployment	-9.76 s(***)	-7.73 s(***)	-9.09 s(***)	-1.03 s(***)	-1.02 s(***)	-1.03 s(***)
IsHoliday						
CPI:Unemployment		4.7 ***				
Mark2:CPI				7.07 no		
Mark5:IsHolidayTRUE					9.73 ***	
FuelPrice:Mark1						6.28 .
FuelPrice:Mark2						8.77 *
FuelPrice:Mark5						-1.12 no
FuelPrice:CPI	7.23 ***					
Fuel:Price:Unemp			-2.68 *			
Root Mean Squared Error /k-fold cross validation (RMSE)	5268.69	5185.57	5236.09	5256.41	5257.8	5227.07
*(***) Test of significance for linear parameter β_1: p-value 0 ***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1* *s(***) Test of significance for non-parametric smoothing spline function: p-value<0.05: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1*						

Source: Author's elaboration

Comparing the gam models with a full set of internal and environmental drivers, it is possible to see the variation of GAM models with linear interactions between internal and external drivers (Table 3-10). GAM5 shows better performance than the rest of the model (GAM5- non - linear and linear functions of internal and environmental drivers and interaction *CPI-Unemployment)* resulting in min RMSE on the test set. Adding the interaction, the increase of one unit of CPI will change Average Sales while keeping Unemployment fixed, and the increase of one unit of Unemployment will change Average Sales per Store while keeping CPI fixed. Testing the models with interactions shows whether the existing correlations between variables in the reality are boosting the effect on the dependent variable in the model. In the following analysis, the exclusion of insignificant variable Markdown2 from GAM5 leads to the GAM3 that is in the playfield for model validation on the test set.

3.4.4 Validation and the comparison of linear and additive models

In the final phase, linear and nonlinear methods are validated, identifying the true underlying relationship on the training set and evaluating its performance on the test set. The performance of each model is tested using the k-fold cross validation method (Table 3-11). K-fold cross validation is a technique to assess how the results of linear and nonlinear models generalize to the independent data-test set often suggested for real-world data and large samples (Kohavi, 1995). It involves dividing dataset randomly into approximately equal k groups, or folds where the first fold is treated as a validation (test) set, and the method is fit on the remaining $k - 1$ folds (Li, 1987; Kohavi, 1995; Hastie *et al*., 2001; Arlot and Celisse, 2010; Rodriguez *et al.,* 2010; James *et al.,* 2013). The (root) mean squared error (RMSE) then computed on the observations in the validation (test) set. This process is repeated k times; each time, taking the different group of observations as validation (test) set. This process results in k estimates of the test error where the k-fold estimate is calculated by averaging these values. The quality of model requires the prediction capability on the independent validation (test) set minimizing the RMSE, the lower RMSE the better accuracy of the model is (Hastie *et al*., 2001; Fushiki, 2011). Of the four models, using the 5-fold cross validation, GAM3 shows the lowest RMSE of the test set and performs better than the LM which has the highest RMSE of the test set (Table 3-8).

Summing up the four models (Table 3-11), GAM3 performs better as a predictor (min RMSE) than the linear model and other GAM. It provides stronger explanatory power (46%) of the dependent variable (Average sales per store) than GAM1/PC (3%) and the linear model (16%). In general, the failure of GAM lies in the interpretation of coefficients; linear models seem better suited for precise description of relationships between predictors and dependent variables. In as much, nonlinear relationships in GAM are best presented graphically; hence, it is useful to interpret GAM3 in a graphical representation (Figure 3-3, 3-4, 3-5).

Table 3-11. The predictive accuracy of linear models and GAM on the test set

Models	*Root Mean Squared Error /k-fold cross validation (RMSE)*
GAM3	4960.23
GAM/PC	4997.89
GAM2	5127.17
LM	6167.46

Source: The performance of linear models and GAM on the test set (Moretta Tartaglione et al., 2019; p. 462)

3.5 Discussion

The analysis aims to explore the interdependencies between variables in one direction from the environmental, internal to sales in the terms of the functional relationships between them. Being aware of natural links between CPI, Unemployment, and Fuel Price, the additional tests (Granger causality) for underlying the causality and inspect the multicollinearity (PCA) are introduced. The significant interactions are included, to try to isolate the effect of the external variable itself. The flexibility of the GAM provides a solution to include the different functions (parametric, nonparametric, and interactions) and to explore each of one effect independently. The independence has not been linked to the characteristics of the variable itself but to the function between variable (predictor-environmental and internal drivers) and response (sales) in GAM. In the first part of the analysis where it is presented the time-series dynamics of raw data (Figure 3-2), the evidence that in both examples of data the noise occurs in Markdown 1, CPI, Fuel Price (Table 3-2, Table 3-4), so the nonlinearity can be related to the nonstability of these variables over the time. Change in data sets can produce multicollinearity that can influence a validation of a model on the test (as correlation may serve as an indicator of collinearity- naturally considering that CPI, Unemployment, Fuel prices, MarkDowns, and Sales might be highly correlated). In this case, we test the multicollinearity in the linear model then the principal component analysis. Using PCA is a strategy to deal with multicollinearity, not the only possible solution. Reshaping and averaging the values over the store the noise is diminished so occurred nonlinearity is not the effect of noise or unaccounted fluctuations, semi-parametric tool is able to catch both linear and nonlinear relationships. With GAMs, it is notable, at which point the function is changing its direction since GAMs produce many coefficients. Though it is difficult to interpret these coefficients, the graphical interpretation of GAMs is a more practical tool, especially for retail chain managers to 'easy' observe what is happening in the retail chain, especially in the lowest density of data where there is a nonlinearity, could be useful information for giant retailer to set/change the strategy.

The sequence of statistical methods applied in this work configures an emergent model based on a linear analysis than a nonlinear analysis and GAMs comparison. The effects of the interaction of variables are noted, with an analysis of the reactions and intensity of the influences between them.

Results from the exploratory analysis contribute to the support to managerial decisions. A comparison of the indicators (R^2 shows the weak explanatory strength of the linear model ((R^2=16%) of variation in sales. A linear model is useful for simple interpretation of coefficients and explanation of relationships; it is of low relevance for describing the interrelations among variables.

The significance of the nonlinear relationships (R^2=46%) is higher when a GAM is used that integrates both linear and nonlinear relationships inside the model. The phenomenon contributes to an explanation of the relevance of nonlinear methods. These methods can generate positive effects in data interpretation, particularly with variables influenced by the macro environment (e.g., Fuel Price, CPI, Unemployment) and generally by market dynamics, in contrast to linear analysis that can generate results with a low significance. The smooth characteristics of the nonlinear methods can describe the particularity of the behavior of specific environmental variables about the results generated by the business activity.

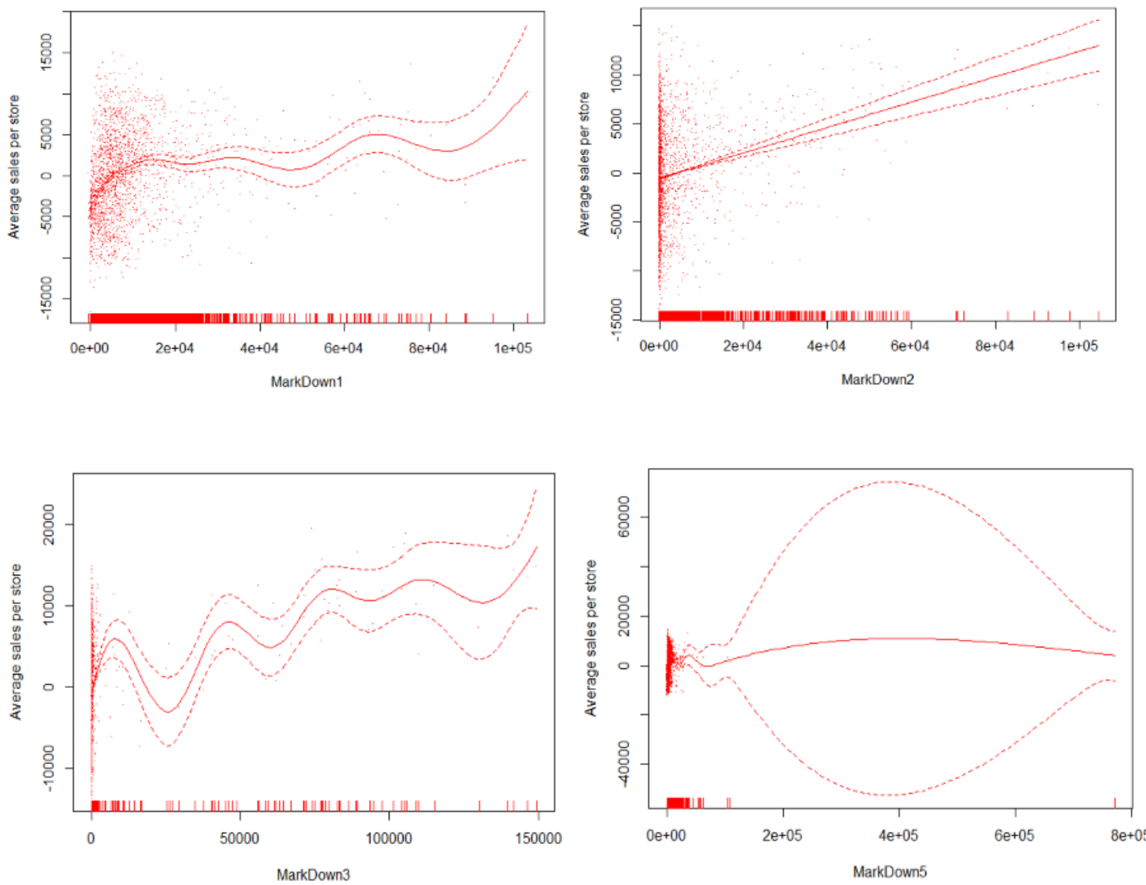
In the next two sections, discussions about the relationships between sales and internal drivers and sales and external environmental drivers are presented, using graphical presentation (Figure 3-3, 3-4, 3-5).

3.5.1 *Relationships between promotional drivers and sales*

In GAM models estimation functions are concerned with a smooth (or a random effect) set of model matrix columns and one or more corresponding penalties (Wood, 2017; p.327). Giving the complex calculations that permit highly modular smooths which are implemented via a smooth method function in the way that each smooth has a prediction matrix method function, which will create the matrix mapping the coefficients of the smooth to predictions at new estimates' values (Wood, 2017; p.327). Smooths can be presented in three or two-dimensional plots, as to better representations the two-dimensional plots are showed further.

The graphical representation of GAM3 (Figure 3-3) shows the nonlinear relationship between Average sales per store and MarkDown1(upper left); the nonlinear relationship between Average sales per store and MarkDown3(down right); the linear relationship between Average sales per store and MarkDown2 (upright); and the insignificant relationship between Average sales per store and MarkDown5. In each graph (Figure 3-3) the observed data are concentrated in the lower value range of promotional discounts, where a linear increase in sales caused by promotional discounts is noticeable. Analyzing the graphs of MarkDown1 and MarkDown3 (upper and lower left), it becomes clear that the use of nonlinear methods provides many more opportunities to interpret the relationship between promotions and sales. The smooth lines are useful when it comes to generating irregularities in the efficacy of the particular promotional politics. In particular, it is possible to perceive the irregularity of the effects of the different promotions (markdowns) on sales. The oscillations of the relationships could be crossed with other related data and taken into consideration during management decision-making processes.

Figure 3-3. *Average sales per store* as a function (GAM3) of the internal variables *Markdown 1 to 5* (promotional discounts).



Source: Figure1. Explained variable average sales per store as a function (GAM3) of the internal variables Markdown 1 to 5 (promotional discounts) by Moretta Tartaglione, A., Bruni, R., and Bozic, M. (2019). Exploring the retail industry environment using nonlinear analysis. *International Journal of Retail and Distribution Management*. Vol. 47 No. 4, pp. 453-47. Copyright (2019) Emerald Publishing Limited.

These considerations allow for interpretation of the role of each markdown concerning sales, making it possible to identify the best promotional strategy behind it, even in instances where the nature of each markdown is unknown.

The graph in the upper right (Figure 3-3) shows the positive linear behavior of MarkDown2, making it reasonable to assume MarkDown2 is the best promotional performer. On the contrary, there is no visible effect of MarkDown5 on sales (lower right), and therefore future management strategies may want to consider eliminating this kind of promotion.

3.5.2 *Relationships between macroeconomic drivers and sales*

Analyzing the four graphs, it is possible to consider the linearity (Figure 3-4. Unemployment and Temperature effects on sales) and nonlinear relations (Figure 3-5. Fuel Price and CPI effects on sales).

Figure 3-4. shows the linear relationships between unemployment and sales, and temperature and sales. For Unemployment (left graph), the generally linear relationship is negative; however, a slight inversion of this is visible in the middle of the function. For such an instance, a nonlinear illustration could permit reflection upon the exact moment other drivers – like the integrated action of different variables, e.g., promotions – can moderate a decrease in sales.

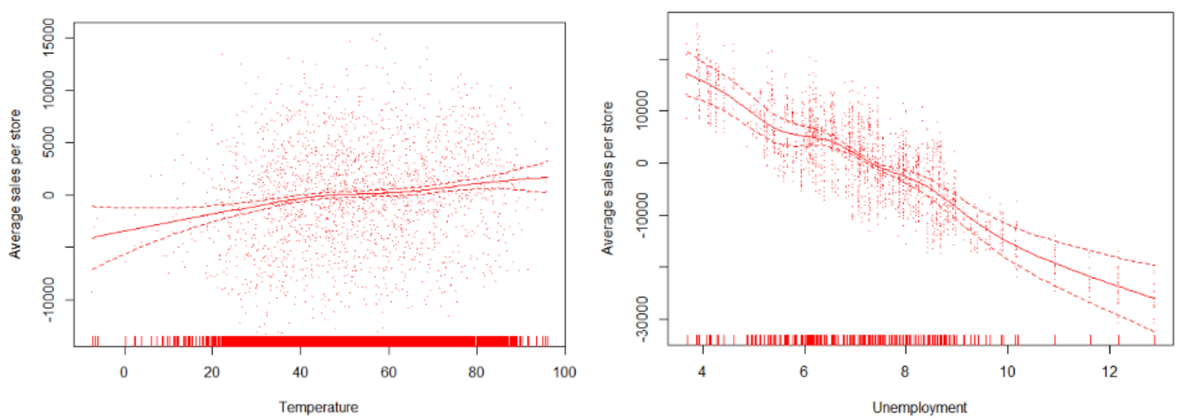
The nonlinear function represents different situations where the changes at lower temperature levels appear to have a greater effect on sales than changes occurring at higher temperature levels, where the function behaves more stably. Temperature is an unusual factor (Arunraj and Ahrens, 2016) that can have an impact on retailers: when the outside temperature is high, people may prefer to go shopping in a hypermarket where there is the comfort of air conditioning, particularly if the offering is made in a specific market with low price positioning. These observations are the lever for pricing scenarios and the retailer's positioning from the perspective of the consumer and the environment.

In each graph (Figure 3-3) the observed data are concentrated in the lower value range of promotional discounts where a linear increase in sales caused by promotional discounts is noticeable (Moretta Tartaglione et al., 2019). Analyzing the graphs of *MarkDown1* and *MarkDown3* (upper and lower left), it becomes clear that the use of nonlinear methods provides many more opportunities to interpret the relationship between promotions and sales. The smooth lines are useful when it comes to managerial considerations about the effects generated by the markdowns analyzed. In particular, it is possible to perceive the irregularity of the effects of the different promotions (markdowns) on sales. The oscillations of the relationships could be crossed with other related data and taken into consideration during management decision-making processes. Figure 3-4. shows the linear relationships between unemployment and sales, and temperature and sales. For Unemployment (right graph), the generally linear relationship is negative; however, a slight inversion of this is visible in the middle of the function. For such an instance, a nonlinear illustration could permit reflection upon the exact moment other drivers – like the integrated action of different variables, e.g., promotions –

can moderate the decrease in sales.

However, the nonlinear function represents different situations that can hypothetically divide the graph: changes at lower temperature levels appear to have a greater effect on sales than changes occurring at higher temperature levels, where the function behaves more stably. In the case of the variable *Temperature* versus *Sales*, a high value for temperature leads to a rise in sales (Figure 3-4); on the other hand, there is an inverse non-linear relationship with *Unemployment* – for a very high level of unemployment, sales are decreasing, but in the middle, the function changes, meaning that an increase in unemployment leads to an increase in sales, added the interaction of CPI and Unemployment increase the effect of both drivers on *Average sales per store.*

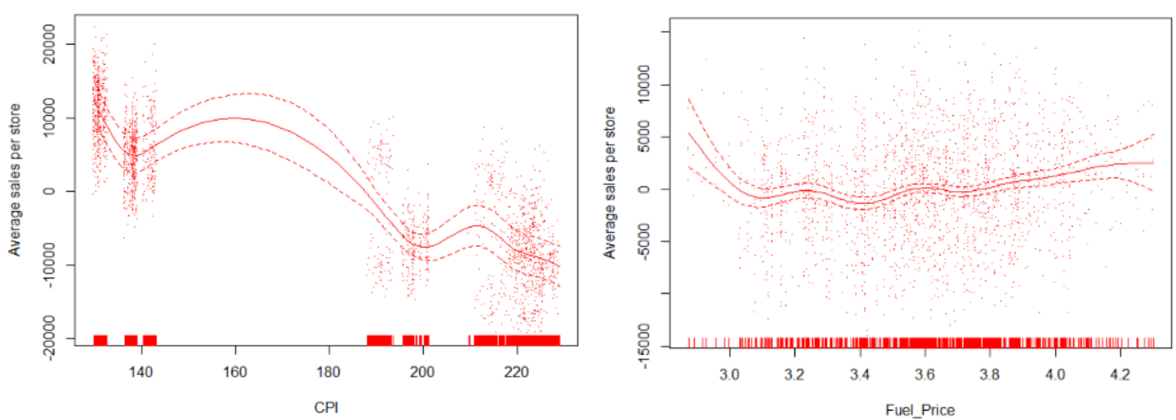
Figure 3-4. *Average sales per store* as a function (GAM3) of the environmental variables *Unemployment, Temperature.*



Source: Figure 2. Explained variable average sales per store as a function (GAM3) of the environmental variables unemployment, temperature by Moretta Tartaglione, A., Bruni, R., and Bozic, M. (2019). Exploring the retail industry environment using nonlinear analysis. *International Journal of Retail and Distribution Management*. Vol. 47 No. 4, pp. 453-47. Copyright (2019) Emerald Publishing Limited.

Figure 3-5. shows the nonlinear relationship of sales with the macro variables Fuel Price and CPI. In the case of Fuel Price (right graph), sales decrease for low values and slight increases in fuel price, while sales increase for higher fuel price values. This nonlinear behavior could be justified by the unique price politics which is recognizable as the low-price positioning strategy of the US retailer. It is possible to assume that with increasing fuel prices, consumers reduce the number of different shops they purchase from, choosing instead to shop at one store that offers consistently low prices (as with the retail chain examined here). In the case of CPI (left graph), for each variation of CPI, there is a correspondingly large reduction in sales, which has been noticed three times in the major concentrations of observed data. These phenomena could be explained by the behavior of price-sensitive consumers and the results as coherent with the fuel price effect on sales. It can be noted a non-linear relationship between the first variable *Fuel Price* and the variable *Sales* when all the other variables are fixed: for low values and a slight increase in the fuel price, sales decrease, while for higher values of fuel price the sales increase. The wiggly behavior of *the Consumer Price Index* could be interesting for analysis, since, for very low and very high values of the CPI, sales decrease tremendously.

Figure 3-5. *Average sales* per store as a function (GAM3) of the environmental variables *Fuel Price* and *CPI*



Source: Figure 3. Explained variable average sales per store as a function (GAM3) of the environmental variables fuel price and CPI by Moretta Tartaglione, A., Bruni, R., and Bozic, M. (2019). Exploring the retail industry environment using nonlinear analysis. *International Journal of Retail and Distribution Management*. Vol. 47 No. 4, pp. 453-47. Copyright (2019) Emerald Publishing Limited.

The graphics are produced using the training set for better visualization. The data are spares in graphics where data of MarkDown1 and MarkDown3 are missing. One of the opportunities using GAMs is that flexible predictor functions can uncover hidden patterns in data that the linear approach will miss, these could be useful information for tactical retail strategies. The omitted drivers have been the limitation of this analysis and data set missing drivers, from general economic development this is relatively true because also data set is missing more drivers from the retail environment as location clues, merchandise clues, price clues, competitors actions, supply chain activities necessary for providing profound strategy, Regard the macro environment the Consumer Price Index (CPI), Unemployment rate and Fuel Price are enough information for giant the retail chain.

3.6 Managerial implications and future aggregations

Effective retail business starts with the sales forecasting process (Mentzer et al, 2005; p.38). Sales forecasting is the future projection of the market demand under the given environmental context. The forecasting should have collaborated and coordinated process between the corporate and operational level and between the producer, distributor, and retailer. In the marketing and sales department, the forecasting has been focused on the yearly or quarterly level when on the logistic and retail store short-term forecasting is preferred. The results of this work should benefit the corporate level of the forecasting process in determining the strategic-decision making, especially in the retail chain that with their own business that determines the dynamics of the environment. The concept of environment in the retail industry assumes value because the involved and connected elements – the weather effects on consumption, relationships between customer and store, relationships with employees, stocks, layouts, promotions, price policies, and legal constraints – can contribute to represent particular complex contexts (Moretta Tartaglione et al., 2019). These complex contexts determine the competitive advantage of a single store and, ultimately, the competitive position of the whole retail chain. In defining the sales forecasting three facts occur 1) *level of the forecasting* 2) *accuracy* 3) *cost* (Table 3-12).

The application and comparison of the linear and nonlinear method can be the opportunity for the retail managers on how to select and isolate the effect of the external drivers. The better performance of the nonlinear model should be considered carefully as the better accuracy may lead to the higher cost of implementing the tool in the forecasting systems. In this, the scale of sales forecasting performance measurement discussed should be taken into consideration to conduct a return on investment analysis of any changes to sales forecasting management. The prices of implementing improved techniques should all be weighed against the improvements in supply chain costs, planning costs, and customer service levels. In most cases, the Return On Investments (ROI) on investments can increase drastically (Mentzer et al., 2005; p 258-259) still it should be evaluated to see what's a suitable level of sales forecasting accuracy for every business function in each level, horizon, and interval.

Table 3-12. Managerial implications and future aggregations based on determinants of forecasting

Determinants of forecasting	**Managerial implications**	**Future aggregations**
Level of forecasting	The corporate level; The additive model can isolate the functional relationships between the external drivers and the sales; Increase the opportunity for collaborative forecasting with external institutions and partners inside the supply chain; Centralized forecasting.	Including more external (competitor's data) and internal variables (display, brand loyalty, number of customers); Translate to the operational level on a weekly or quarterly basis; Compare to the simple time series models.
Accuracy	The nonlinear model with additional environmental produce better predictive accuracy than linear model (Table 3-10); Reduced heteroscedasticity.	Measuring the accuracy with lags effect in the short term forecasting.
Cost	Nonlinear models require the computational costs; Saving the efficiency in improving the forecasting models; Allow retail manage to define and to reallocate the stores or adapt to the current regional market; Reducing the costs of the inventory.	Standardizing the additive model for the forecasting products group or individual; Avoiding model to under and overfit.

Source: Author's elaboration

A lot of managerial indicators – specific historical situation of the company, strategic decisions concerning the expansion, consolidation of markets or situation of crisis and environmental variables – geographical position of retailers, the dynamism of consumption, regional activity, the sensitivity to a technology of consumer, political influence, socio-demographic conditions could affect the decision to use specific models. In any case for retail companies, the switch from the traditional model to new approaches is difficult to shift, in particular, for the change-averse retailers focused on the traditional management. They prefer to use the consolidated models to manage and analyze data in particular sales and relationships between internal and external variables.

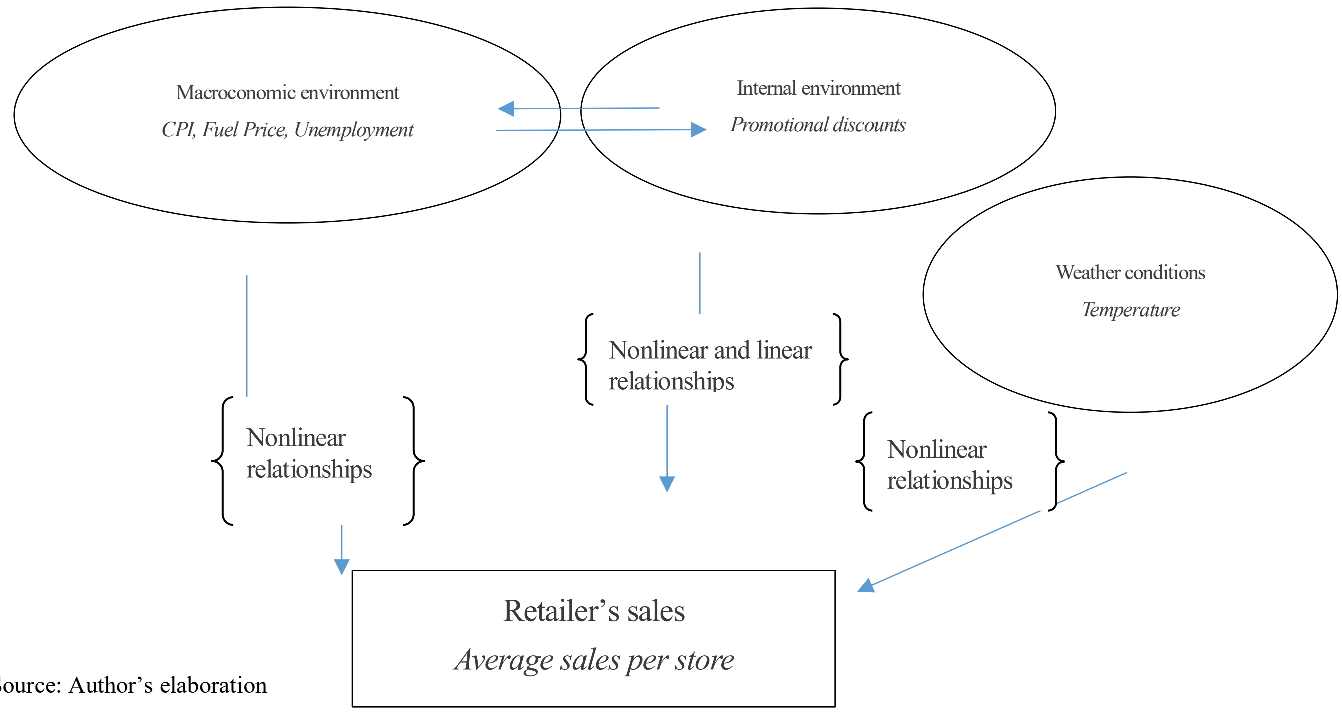
The use of the nonlinear models can find the role from the system perspective of retail companies, in the case of adaption, restructuring, and transformation. In the case of an adaptation, the effects of managerial decisions on which external or internal driver should affect the variable costs of the company. In the case of transformation and restructuring, the analysis can be shifted to future projects (investments) and the determination of the fixed costs.

Environmental factors mixed with factors internal to the company are at the base of the management of the relationship between the company and the environment in general. The simple identification of the variables is not the purpose of the analysis; the relevance of the non-linear model is the focus on the relationships between the variables. It is precisely the functional and intentional concept of interaction that clarifies our comprehension of the significance of "non-linearity". The relationship between the organization (a company, for example) and its environment can be interpreted first of all by the interaction between its components, both internal and external.

Pointing out the importance of relationships between environmental drivers *Fuel Price, CPI, Unemployment* and between environmental and internal drivers (Markdown 1-5) in GAM models (Table 3-9), leads to aggregate the drivers which represent the macroeconomic environment and retailer itself, setting the effect of *Temperature* on sales as distinct factor caused by the weather conditions. Being able to illustrate the relationship between variable

(Figure 3-6) that are elements of this research, offers the possibility for management to filter the information for the decisional schemes.

Figure 3-6. Aggregation of internal and environmental drivers for retailer's sales.



Source: Author's elaboration

In some way, in the real situation, it is challenging to quantify the effects of the environmental drivers for a retail manager (or researcher interested in additional proof of validity) and see whether any of the claimed effects are important enough to demand her or his attendance. New causal connections (which would need to be tested at the operational level), and if so the nonlinear effects need to show how exactly they are different from a linear assumption and what managerial action should change in response. In the explorative research, it is the assumption that the scenario is complex and there is widespread information. The results from such an approach are not statistical solutions but rather the development of management perspectives and determination of the role dynamism plays in the management of retail companies. On the other hand, in stable markets, and for small retail chains, the use of linear models for analyzing and managing data are preferred, particularly when examining the relationships between sales and internal variables.

Due to better interpretability and flexibility this work suggest using general additive models to explain the effect of drivers from the retail environment on dynamics of sales that can imply the reliable decision-making inside the multistore retail chain, besides for future research, the additional sales drivers as well price elasticity, competitor's actions and market drivers can be tackled and considered in gam models to illustrate the real dynamics of sales.

3.7 Conclusions

The aim of the research is to explore interdependencies based on assumption that analysis consists of interdependent and interacting variables. In the real situation, these assumptions do not work well together, as the macroeconomic indicators are entangled and have the cross-over effect, instead, using the applicable methods it is possible to isolate the effect of the external variables with additive nonlinear models and improve the significance of the statements of interdependency and interacting.

What is valuable is that nonlinearity, nonnormal errors, and heteroscedasticity are automatically resolved by nonlinear techniques (Soguero-Ruiz et al., 2014). Today's retail surroundings are characterized by complex dynamics that cannot be easily interpreted with linear models. Linear models will dismiss the patterns and the moments in which the relationships will switch. In fact, the nonlinear methods can reflect the real retail dynamics, and that retailers cannot base decisions only on the internal indicators. The result of this research is to outline the necessity for the large retail chains to consider the changes in the macroeconomic environment expressed in the indexes, rates, and prices as well as they focus on the promotional discounts which as inevitable indicators for the forecasting. At the corporate level, the decision-maker has to evaluate the long term positioning of the retailer on the market. Having the decisional framework on how to evaluate the environmental effect can benefit the efficiency of the retail strategies. Before all forms of forecasting, quantitative or judgmental, it is important to a pattern or relationship exists that can be identified and modeled (measured). Once modeled, it is possible to assume that relationship will continue into the future. Each pattern or relationship must be correctly identified and measured before it is possible to create a forecast (Chase, 2016; p. 153). In certain cases, the patterns and relationships coexist with unexplained variance and can change unpredictably over time. In this analysis, using the various modeling approaches it is isolated and measured the causal effect of the macroeconomic variables (CPI, Fuel Price, Unemployment) and the weather conditions (Temperature). While these macroeconomic indicators can be correlated and interdependent in the real scenario, the general additive model isolates the independently the relationships between external drivers and sales. Although this study focuses on a specific number of observations due to the limited variety of internal drivers (in this case, only promotions are used), the large volume of data from 45 stores permits the argument that the promotions (MarkDown 1-5) are representative of specific management strategies that affect the performance of each store. Limitations of the lack of adequate information from the link of the promotional discounts to sales are the burden of the research. Still, it is possible to affirm that nonlinear analysis shows better predictive accuracy than linear models and that included external drivers benefit to that significantly. This work with an explorative analysis presented the effects of nonlinear analysis upon the opportunities to capture the influence of the internal and external environmental drivers on the sales and different implications are generated.

Continued demand volatility and retail market saturation in which the competitive strategy leads to a more expanded product portfolio is pressing the supply chain leaders to improve the demand management performance (Chase, 2016; p. 14). Increasing the costs push the members of the supply chain to had better align the supply and demand. The necessity for centralized and collaborative forecasting is triggering the improvement of the efficient and accurate forecasting model. Partner communication and collaborative execution in the supply chain from the production to the retailer must constantly collaborate on what events are affecting the demand. (Chase, 2016; p.15).

The advanced tools have to adapt, react and change fast to the flow of the information inside the supply chain. Models represented in this work can help the retailer to contextualize the market and isolate the irregular and nonlinear patterns that simultaneously can affect the outcome. Future research on the application of additive methods in the retail industry should analyze the effects on an operational level and isolating the external effect on the lesser level as the product segment or individual article. Looking at the short-term effects weekly or daily and particularly in analyzing the drivers that managers should use as the source of competitive advantage for the particular store or the group of stores.

3.8 References

1. Aguilar-Palacios, C., Muñoz-Romero, S., and Rojo-Álvarez, J. L. (2019). Forecasting promotional sales within the neighbourhood. *IEEE Access*, *7*, 74759-74775. https://doi.org/10.1109/ACCESS.2019.2920380
2. Albarune, A. R. B., and Habib, M. M. (2015). A study of forecasting practices in supply chain management. *International Journal of Supply Chain Management*, *4*(2), 55-61.
3. Alexander, A., Cryer, D., and Wood, S. (2008). Location planning in charity retailing. *International Journal of Retail and Distribution Management*. https://doi.org/10.1108/09590550810880570
4. Alon, I., Qi, M., and Sadowski, R. J. (2001). Forecasting aggregate retail sales:: a comparison of artificial neural networks and traditional methods. *Journal of retailing and consumer services*, *8*(3), 147-156. https://doi.org/10.1016/S0969-6989(00)00011-4
5. Arlot, S., and Celisse, A. (2010). A survey of cross-validation procedures for model selection. *Statistics surveys*, *4*, 40-79.
6. Arunraj, N. S. and Ahrens, D. (2016). Estimation of non-catastrophic weather impacts for retail industry. *International Journal of Retail and Distribution Management*, *44*(7), 731-753. https://doi.org/10.1108/IJRDM-07-2015-0101
7. Aviv, Y. (2001). The effect of collaborative forecasting on supply chain performance. *Management science*, *47*(10), 1326-1343. https://doi.org/10.1287/mnsc.47.10.1326.10260
8. Baecke, P., De Baets, S., and Vanderheyden, K. (2017). Investigating the added value of integrating human judgement into statistical demand forecasting systems. *International Journal of Production Economics*, *191*, 85-96. https://doi.org/10.1016/j.ijpe.2017.05.016
9. Bahng, Y. and Kincade, D.H. (2012). The relationship between temperature and sales: sales data analysis of a retailer of branded women's business wear. *International Journal of Retail and Distribution Management*, Vol. 40 No. 6, pp. 410-426. https://doi.org/10.1108/09590551211230232
10. Bai, J., and Ng, S. (2013). Principal components estimation and identification of static factors. *Journal of Econometrics*, 176(1), 18-29. https://doi.org/10.1016/j.jeconom.2013.03.007
11. Barber, B. M., Click, R. W., and Darrough, M. N. (1999). The impact of shocks to exchange rates and oil prices on US sales of American and Japanese automakers. *Japan and the world economy*, *11*(1), 57-93. https://doi.org/10.1016/S0922-1425(98)00041-3
12. Battini, D., Persona, A., and Allesina, S. (2007). Towards a use of network analysis: quantifying the complexity of Supply Chain Networks. *International Journal of Electronic Customer Relationship Management*, *1*(1), 75-90. https://doi.org/10.1504/IJECRM.2007.014427
13. Baumann, C., Elliott, G., and Burton, S. (2012). Modeling customer satisfaction and loyalty: survey data versus data mining. *Journal of services marketing*. https://doi.org/10.1108/08876041211223951
14. Bertrand, J. L., and Parnaudeau, M. (2017). No more blaming the weather: a retailer's approach to measuring and managing weather variability. *International Journal of Retail and Distribution Management*. https://doi.org/10.1108/IJRDM-09-2016-0152
15. Buchko, A. A. (1994). Conceptualization and measurement of environmental uncertainty: An assessment of the Miles and Snow perceived environmental uncertainty scale. *Academy of Management Journal*, 37(2), 410-425. https://doi.org/10.5465/256836

16. Camacho, A., and Conover, E. (2019). The impact of receiving SMS price and weather information on small scale farmers in Colombia. *World Development*, *123*, 104596. https://doi.org/10.1016/j.worlddev.2019.06.020
17. Cao, Z., Li, G., and Song, H. (2017). Modelling the interdependence of tourism demand: The global vector autoregressive approach. *Annals of Tourism Research*, *67*, 1-13. https://doi.org/10.1016/j.annals.2017.07.019
18. Carrière-Swallow, Y., and Labbé, F. (2013). Nowcasting with Google Trends in an emerging market. *Journal of Forecasting*, *32*(4), 289-298. https://doi.org/10.1002/for.1252
19. Cavallo, A., Cruces, G., and Perez-Truglia, R. (2017). Inflation expectations, learning, and supermarket prices: Evidence from survey experiments. *American Economic Journal: Macroeconomics*, *9*(3), 1-35. https://doi.org/10.1257/mac.20150147
20. Chase Jr, C. W. (2013), *Demand-driven forecasting: a structured approach to forecasting*. John Wiley and Sons.
21. Chase, C. W. (2016). *Next generation demand management: People, process, analytics, and technology*. John Wiley and Sons.
22. Chen, J. C., Cheng, C. H., and Huang, P. B. (2013). Supply chain management with lean production and RFID application: A case study. *Expert Systems with applications*, *40*(9), 3389-3397. https://doi.org/10.1016/j.eswa.2012.12.047
23. Choi, T. Y., Dooley, K. J., and Rungtusanatham, M. (2001). Supply networks and complex adaptive systems: control versus emergence. *Journal of operations management*, *19*(3), 351-366. https://doi.org/10.1016/S0272-6963(00)00068-1
24. Christen, M. Gupta, S. Porter, J.C., Staelin, R. and Wittink, D.R. (1997). Using market-level data to understand promotion effects in a nonlinear model. *Journal of Marketing Research*, pp.322-334. https://doi.org/10.1177%2F002224379703400302
25. Chu, C. W., and Zhang, G. P. (2003). A comparative study of linear and nonlinear models for aggregate retail sales forecasting. *International Journal of production economics*, *86*(3), 217-231. https://doi.org/10.1016/S0925-5273(03)00068-9
26. Coakley, J., Fuertes, A. M., and Smith, R. (2002). *A principal components approach to cross-section dependence in panels*
27. Coussement, K., Benoit, D. F., and Van den Poel, D. (2010). Improved marketing decision making in a customer churn prediction context using generalized additive models. *Expert Systems with Applications*, *37*(3), 2132-2143. https://doi.org/10.1016/j.eswa.2009.07.029
28. Crozier, M. (1964). Pouvoir et organisation. *European Journal of Sociology/Archives Européennes de Sociologie/Europäisches Archiv für Soziologie*, 5(1), 52-64. http://www.jstor.org/stable/23998451
29. Dawson, J. (2000). Retailer power, manufacturer power, competition and some questions of economic analysis. *International Journal of Retail and Distribution Management.* 28 (1), pp. 5-8. https://doi.org/10.1108/09590550010306700
30. Dawson, J.F. (2014) Moderation in Management Research: What, Why, When, and How. *J Bus Psychol* **29,** 1–19. https://doi.org/10.1007/s10869-013-9308-7
31. Daft, R. L., Sormunen, J., and Parks, D. (1988). Chief executive scanning, environmental characteristics, and company performance: An empirical study. *Strategic management journal*, 9(2), 123-139. https://doi.org/10.1002/smj.4250090204
32. Derksen, S., and Keselman, H. J. (1992). Backward, forward and stepwise automated subset selection algorithms: Frequency of obtaining authentic and noise variables. *British Journal of Mathematical and Statistical Psychology*, *45*(2), 265-282. https://doi.org/10.1111/j.2044-8317.1992.tb00992.x
33. Dickey, D. A., and Fuller, W. A. (1979). Distribution of the estimators for autoregressive time series with a unit root. *Journal of the American statistical association*, *74*(366a), 427-431. https://doi.org/10.1080/01621459.1979.10482531
34. Dickey, D. A., and Fuller, W. A. (1981). Likelihood ratio statistics for autoregressive time series with a unit root. *Econometrica: journal of the Econometric Society*, 1057-1072. https://doi.org/10.2307/1912517
35. Dormann, C. F., Elith, J., Bacher, S., Buchmann, C., Carl, G., Carré, G., ... and Lautenbach, S. (2013). Collinearity: a review of methods to deal with it and a simulation study evaluating their performance. *Ecography*, *36*(1), 27-46. https://doi.org/10.1111/j.1600-0587.2012.07348.x
36. Duncan, R. B. (1972). Characteristics of organizational environments and perceived environmental uncertainty. *Administrative science quarterly*, 313-327. https://doi.org/10.2307/2392145
37. Ekinci, E., and Baykasoğlu, A. (2019). Complexity and performance measurement for retail supply chains. *Industrial Management and Data Systems*. 119 (4), pp. 719-742. https://doi.org/10.1108/IMDS-08-2018-0342

38. Everingham, Y., Sexton, J., Skocaj, D., and Inman-Bamber, G. (2016). Accurate prediction of sugarcane yield using a random forest algorithm. *Agronomy for sustainable development*, *36*(2), 27. https://doi.org/10.1007/s13593-016-0364-z
39. Fernie, J., and Arnold, S. J. (2002). Wal-Mart in Europe: prospects for Germany, the UK and France. *International Journal of Retail and Distribution Management*. https://doi.org/10.1108/09590550210418119
40. Fildes, R., Goodwin, P., Lawrence, M., and Nikolopoulos, K. (2009). Effective forecasting and judgmental adjustments: an empirical evaluation and strategies for improvement in supply-chain planning. *International journal of forecasting*, *25*(1), 3-23. https://doi.org/10.1016/j.ijforecast.2008.11.010
41. Fildes, R., Ma, S., and Kolassa, S. (2019). Retail forecasting: Research and practice. *International Journal of Forecasting*. https://doi.org/10.1016/j.ijforecast.2019.06.004
42. Fornari, D., Grandi, S., and Fornari, E. (2009). The role and management of product innovation in retailer assortments: evidence from the Italian FMCG market. *The International Review of Retail, Distribution and Consumer Research*, *19*(1), 29-43. https://doi.org/10.1016/j.jretconser.2013.07.004
43. Fushiki, T. (2011). Estimation of prediction error by using K-fold cross-validation. *Statistics and Computing*, *21*(2), 137-146. https://doi.org/10.1007/s11222-009-9153-8
44. Granger, C. W. (1988). Some recent development in a concept of causality. *Journal of econometrics*, *39*(1-2), 199-211. https://doi.org/10.1016/0304-4076(88)90045-0
45. Green, P. J., and Silverman, B. W. (1993). *Nonparametric regression and generalized linear models: a roughness penalty approach*. Crc Press.
46. Grewal, D. and Levy, M. (2007). Retailing research: past, present, and future. *Journal of Retailing*, Vol. 83 No. 4, pp. 447-464. https://doi.org/10.1016/j.jretai.2007.09.003
47. Guisan, A., Edwards Jr, T. C., and Hastie, T. (2002). Generalized linear and generalized additive models in studies of species distributions: setting the scene. *Ecological modelling*, *157*(2-3), 89-100. https://doi.org/10.1016/S0304-3800(02)00204-1
48. GÜR ALI, Ö. Z. D. E. N. (2013). Driver moderator method for retail sales prediction. *International Journal of Information Technology and Decision Making*, *12*(06), 1261-1286. https://doi.org/10.1142/S0219622013500363
49. Hackney, R., Grant, K., and Birtwistle, G. (2006). The UK grocery business: towards a sustainable model for virtual markets. *International Journal of Retail and Distribution Management*. https://doi.org/10.1108/09590550610660279
50. Guo, R. (2017). *An economic inquiry into the nonlinear behaviors of nations: dynamic developments and the origins of civilizations*. Springer.
51.
52. Gutiérrez, R., Román, P., and Torres, F. (1999). Inference and first-passage-times for the lognormal diffusion process with exogenous factors: application to modelling in economics. *Applied Stochastic Models in Business and Industry*, *15*(4), 325-332. https://doi.org/10.1002/(SICI)1526-4025(199910/12)15:4%3C325::AID-ASMB397%3E3.0.CO;2-F
53. Hamilton, J. D. (1983). Oil and the macroeconomy since World War II. *Journal of political economy*, *91*(2), 228-248. https://doi.org/10.1086/261140
54. Hamilton, J. (2011). NONLINEARITIES AND THE MACROECONOMIC EFFECTS OF OIL PRICES. *Macroeconomic Dynamics, 15*(S3), 364-378 https://10.1017/S1365100511000307
55. Harrauer V. and Schnedlitz P. (2016). Impact of environment on performance measurement design and processing in retailing. *International Journal of Retail and Distribution Management*, 44 (3), 320 – 335. https://doi.org/10.1108/IJRDM-03-2015-0048
56. Hastie, T., and Tibshirani, R. (1987). Generalized additive models: some applications. *Journal of the American Statistical Association*, *82*(398), 371-386.10.1080/01621459.1987.10478440
57. Hastie, T. J., and Tibshirani, R. J. (1990). *Generalized additive models* (Vol. 43). CRC press.
58. Haupt, H. and Kagerer, K. (2012). Beyond mean estimates of price and promotional effects in scanner-panel sales–response regression. *Journal of Retailing and Consumer Services*, 19(5), pp.470-483. https://doi.org/10.1016/j.jretconser.2012.06.002
59. Friedman, J., Hastie, T., and Tibshirani, R. (2001). *The elements of statistical learning* (Vol. 1, No. 10). New York: Springer series in statistics.
60. Hastie, T. J. (Ed.). (2017). *Statistical models in S*. Routledge.
61. Heshmaty, B., and Kandel, A. (1985). Fuzzy linear regression and its applications to forecasting in uncertain environment. *Fuzzy sets and systems*, *15*(2), 159-191. https://doi.org/10.1016/0165-0114(85)90044-2

62. Hickson, D. J., Hinings, C. R., Lee, C. A., Schneck, R. E., and Pennings, J. M. (1971). A strategic contingencies' theory of intraorganizational power. *Administrative science quarterly*, 216-229. https://doi.org/10.2307/2391831
63. Hofmann, D. A., and Gavin, M. B. (1998). Centering decisions in hierarchical linear models: Implications for research in organizations. *Journal of Management*, *24*(5), 623-641. https://doi.org/10.1177%2F014920639802400504
64. Hristov, L., and Reynolds, J. (2015). Perceptions and practices of innovation in retailing: Challenges of definition and measurement. *International Journal of Retail and Distribution Management*. https://doi.org/10.1108/IJRDM-09-2012-0079
65. Huang, T., Fildes, R., and Soopramanien, D. (2014). The value of competitive information in forecasting FMCG retail product sales and the variable selection problem. *European Journal of Operational Research*, *237*(2), 738-748. https://doi.org/10.1016/j.ejor.2014.02.022
66. Hübner, A. H., Kuhn, H., and Sternbeck, M. G. (2013). Demand and supply chain planning in grocery retail: an operations planning framework. *International Journal of Retail and Distribution Management*. https://doi.org/10.1108/IJRDM-05-2013-0104
67. Hwang, E. J. (2005). *Strategic management and financial performance in South Korean apparel retail stores* (Doctoral dissertation, Virginia Tech). http://hdl.handle.net/10919/29855
68. Hyndman, R. J., and Athanasopoulos, G. (2014). Optimally reconciling forecasts in a hierarchy. *Foresight: The International Journal of Applied Forecasting*, (35), 42-48.
69. Guevara, C. A. (2018). Overidentification tests for the exogeneity of instruments in discrete choice models. *Transportation Research Part B: Methodological*, *114*, 241-253. https://doi.org/10.1016/j.trb.2018.05.020
70. James, G., Witten, D., Hastie, T., and Tibshirani, R. (2013). *An introduction to statistical learning* (Vol. 112). New York: springer. https:/doi.org/10.1007/978-1-4614-7138-7
71. Johnson, J. S. (2014). Nonlinear analyses in sales research: Theoretical bases and analytical considerations for polynomial models. *Journal of Personal Selling and Sales Management*, *34*(4), 302-317. https://doi.org/10.1080/08853134.2014.903804
72. Kalyanam, K., and Shively, T. S. (1998). Estimating irregular pricing effects: A stochastic spline regression approach. *Journal of Marketing Research*, *35*(1), 16-29. https://doi.org/10.1177%2F002224379803500104
73. Khandwalla, P. (1972). Environment and its impact on the organization. *International Studies of Management and Organization*, 2 (3), pp. 297-313. https://doi.org/10.1080/00208825.1972.11656125
74. Kohavi, R. (1995, August). A study of cross-validation and bootstrap for accuracy estimation and model selection. In *Ijcai* (Vol. 14, No. 2, pp. 1137-1145).
75. Kung, L. C., and Chen, Y. J. (2014). Impact of reseller's and sales agent's forecasting accuracy in a multilayer supply chain. *Naval Research Logistics (NRL)*, *61*(3), 207-222. https://doi.org/10.1002/nav.21578
76. Kumar, N., Scheer, L. K., & Steenkamp, J. B. E. (1998). Interdependence, punitive capability, and the reciprocation of punitive actions in channel relationships. *Journal of Marketing Research*, *35*(2), 225-235. https://doi.org/10.1177%2F002224379803500208
77. Lawrence, P. R., and Lorsch, J. W. (1967). Differentiation and integration in complex organizations. *Administrative science quarterly*, 1-47. https://doi.org/10.2307/2391211
78. Lee, K., and Ni, S. (2002). On the dynamic effects of oil price shocks: a study using industry level data. *Journal of Monetary economics*, *49*(4), 823-852 https:/doi.org/10.1016/S0304-3932(02)00114-9
79. Li, K. C. (1987). Asymptotic optimality for $ C_p, C_L $, cross-validation and generalized cross-validation: Discrete index set. *The Annals of Statistics*, *15*(3), 958-975.
80. Lightfoot, W. (2003). Multi-channel mistake: the demise of a successful retailer. *International Journal of Retail and Distribution Management*. 31 (4), pp. 220-229. https://doi.org/10.1108/09590550310469211
81. Ma, S. H., Fildes, R., and Huang, T. (2016). Demand forecasting with high dimensional data: The case of SKU retail sales forecasting with intra- and inter-category promotional information. *European Journal of Operational Research, 249*(1), 245-257. https://doi.org/10.1016/j.ejor.2015.08.029
82. Mai, E. S., Yang, J., and Chen, H. (2011). Primary product network size on complementary product sales: moderating effects of customer characteristics. *International Journal of Retail and Distribution Management*. https://doi.org/10.1108/09590551111177963
83. Marra, G., and Radice, R. (2011). A flexible instrumental variable approach. *Statistical Modelling*, 11(6), 581-603. https://doi.org/10.1177%2F1471082X1001100607
84. Meloun, M., Militký, J., Hill, M., and Brereton, R. G. (2002). Crucial problems in regression modelling and their solutions. *Analyst*, *127*(4), 433-450.
85. management:

86. Mentzer, J. T., and Moon, M. A. (2005). *Sales forecasting management: a demand management approach*. 2nd ed. Sage Publications.
87. Miller, D., and Friesen, P. H. (1983). Strategy-making and environment: the third link. *Strategic management journal*, 4(3), 221-235. https://doi.org/10.1002/smj.4250040304
88. Miller, K. D. (1993). Industry and country effects on managers' perceptions of environmental uncertainties. *Journal of International Business Studies*, 24(4), 693-714. https://doi.org/10.1057/palgrave.jibs.8490251
89. Milliken, F. J. (1987). Three types of perceived uncertainty about the environment: State, effect, and response uncertainty. *Academy of Management review*, 12(1), 133-143. https://doi.org/10.5465/amr.1987.4306502
90. Mintz, O., and Currim, I. S. (2013). What drives managerial use of marketing and financial metrics and does metric use affect performance of marketing-mix activities?. *Journal of Marketing*, *77*(2), 17-40. https://doi.org/10.1509/jm.11.0463
91. Moretta Tartaglione, A., Bruni, R., Bozic, M., and Cavacece, Y. (2018). How the environmental complexity affects forecasting in retail? Actual research trends. *Journal of Marketing Trends (1961-7798)*, *5*(2).
92. Moretta Tartaglione, A., Bruni, R., and Bozic, M. (2019). Exploring the retail industry environment using nonlinear analysis. *International Journal of Retail and Distribution Management*, Vol. 47 No. 4, pp. 453-47.https://doi.org/10.1108/IJRDM-09-2017-0224
93. Murray, K.B., Di Muro, F., Finn, A. and Leszczyc, P.P. (2010), "*The effect of weather on consumer spending*", Journal of Retailing and Consumer Services, Vol. 17 No. 6, pp. 512-520. https://doi.org/10.1016/j.jretconser.2010.08.006
94. Nair, A., Yan, T., Ro, Y. K., Oke, A., Chiles, T. H., and Lee, S. Y. (2016). How environmental innovations emerge and proliferate in supply networks: A complex adaptive systems perspective. *Journal of Supply Chain Management*, *52*(2), 66-86. https://doi.org/10.1002/joom.1022
95. Nair, A., and Reed-Tsochas, F. (2019). Revisiting the complex adaptive systems paradigm: Leading perspectives for researching operations and supply chain management issues. https://doi.org/10.1002/joom.1022
96. Newey, W. K. (1985). Generalized method of moments specification testing. *Journal of econometrics*, *29*(3), 229-256. https://doi.org/10.1016/0304-4076(85)90154-X
97. Pantano, E., and Naccarato, G. (2010). Entertainment in retailing: The influences of advanced technologies. *Journal of Retailing and Consumer Services*, *17*(3), 200-204. https://doi.org/10.1016/j.jretconser.2010.03.010
98. Pantano, E., Priporas, C.V., Sorace, S. and Iazzolino, G. (2017), "*Does innovation-orientation lead to retail industry growth? Empirical evidence from patent analysis*", Journal of Retailing and Consumer Services, Vol. 34 No. 2017, pp. 88-94. https://doi.org/10.1016/j.jretconser.2016.10.001
99. Papanagnou, C. I., and Matthews-Amune, O. (2018). Coping with demand volatility in retail pharmacies with the aid of big data exploration. *Computers and Operations Research*, *98*, 343-354. https://doi.org/10.1016/j.cor.2017.08.009
100. Parsons, A. G. (2001). The association between daily weather and daily shopping patterns. *Australasian Marketing Journal (AMJ)*, *9*(2), 78-84. https://doi.org/10.1016/S1441-3582(01)70177-2
101. Pathak, S. D., Day, J. M., Nair, A., Sawaya, W. J., and Kristal, M. M. (2007). Complexity and adaptivity in supply networks: Building supply network theory using a complex adaptive systems perspective. *Decision sciences*, *38*(4), 547-580. https://doi.org/10.1111/j.1540-5915.2007.00170.x
102. Pindyck, R. S., & Rotemberg, J. J. (1983). Dynamic factor demands and the effects of energy price shocks. *The American Economic Review*, *73*(5), 1066-1079. https://www.jstor.org/stable/1814672
103. Prieto, M., and Caemmerer, B. (2013). An exploration of factors influencing car purchasing decisions. *International Journal of Retail and Distribution Management*. https://doi.org/10.1108/IJRDM-02-2012-0017
104. Piotrowicz, W., and Cuthbertson, R. (2014). Introduction to the special issue information technology in retail: Toward omnichannel retailing. *International Journal of Electronic Commerce*, *18*(4), 5-16. https://doi.org/10.2753/JEC1086-4415180400
105. Ray, S. (2012). Testing granger causal relationship between macroeconomic variables and stock price behaviour: evidence from India. *Advances in Applied Economics and Finance*, *3*(1), 470-481.
106. Ren S., Patrick Hui C., Jason Choi T. (2018) AI-Based Fashion Sales Forecasting Methods in Big Data Era. In: Thomassey S., Zeng X. (eds) Artificial Intelligence for Fashion Industry in the Big Data Era. *Springer Series in Fashion Business*. Springer, Singapore. https://doi.org/10.1007/978-981-13-0080-6_2
107. Rodriguez, J. D., Perez, A., and Lozano, J. A. (2009). Sensitivity analysis of k-fold cross validation in prediction error estimation. *IEEE transactions on pattern analysis and machine intelligence*, *32*(3), 569-575. https://doi.org/10.1109/TPAMI.2009.187

108. Rueda-Manzanares, A., Aragón-Correa, J. A., and Sharma, S. (2008). The influence of stakeholders on the environmental strategy of service firms: The moderating effects of complexity, uncertainty and munificence. *British Journal of management*, 19(2), 185-203. https://doi.org/10.1111/j.1467-8551.2007.00538.x
109. Sagaert, Y. R., Aghezzaf, E. H., Kourentzes, N., and Desmet, B. (2018). Tactical sales forecasting using a very large set of macroeconomic indicators. *European Journal of Operational Research*, *264*(2), 558-569. https://doi.org/10.1016/j.ejor.2017.06.054
110. Sanders, N. R., and Manrodt, K. B. (2003). The efficacy of using judgmental versus quantitative forecasting methods in practice. *Omega*, *31*(6), 511-522. https://doi.org/10.1016/j.omega.2003.08.007
111. Shahabuddin, S. (2009). Forecasting automobile sales. *Management Research News*. https://doi.org/10.1108/01409170910965260
112. Shively, T. S., Allenby, G. M., and Kohn, R. (2000). A nonparametric approach to identifying latent relationships in hierarchical models. *Marketing Science*, *19*(2), 149-162. https://doi.org/10.1287/mksc.19.2.149.11807
113. Sargan, J. D. (1958). The estimation of economic relationships using instrumental variables. *Econometrica: Journal of the Econometric Society*, 393-415. https://doi.org/10.2307/1907619
114. Sillanpää, V., and Liesiö, J. (2018). Forecasting replenishment orders in retail: value of modelling low and intermittent consumer demand with distributions. *International Journal of Production Research*, *56*(12), 4168-4185. https://doi.org/10.1080/00207543.2018.1431413
115. Sirgy, M. J., Grewal, D., and Mangleburg, T. (2000). Retail environment, self-congruity, and retail patronage: An integrative model and a research agenda. *Journal of Business research*, *49*(2), 127-138. https://doi.org/10.1016/S0148-2963(99)00009-0
116. Sivadasan, S., Efstathiou, J., Frizelle, G., Shirazi, R., and Calinescu, A. (2002). An information-theoretic methodology for measuring the operational complexity of supplier-customer systems. *International Journal of Operations and Production Management*. https://doi.org/10.1108/01443570210412088
117. Sivadasan, S., Efstathiou, J., Calinescu, A., and Huatuco, L. H. (2006). Advances on measuring the operational complexity of supplier–customer systems. *European Journal of Operational Research*, *171*(1), 208-226. https://doi.org/10.1016/j.ejor.2004.08.032
118. Soguero-Ruiz, Cristina, Francisco Javier Gimeno-Blanes, Inmaculada Mora-Jiménez, María del Pilar Martínez-Ruiz, and José Luis Rojo-Álvarez. "Statistical nonlinear analysis for reliable promotion decision-making." Digital Signal Processing 33 (2014): 156-168. https://doi.org/10.1016/j.dsp.2014.06.014
119. Stock, J. H., and Watson, M. W. (1998). Diffusion indexes. *NBER working paper*, (w6702). https://ssrn.com/abstract=226366
120. Stock, J. H., and Watson, M. W. (2001). Vector autoregressions. *Journal of Economic perspectives*, *15*(4), 101-115. https://doi.org/10.1257/jep.15.4.101
121. Stock, J. H., and Watson, M. W. (2002). Forecasting using principal components from a large number of predictors. *Journal of the American statistical association*, *97*(460), 1167-1179. https://doi.org/10.1198/016214502388618960
122. Stock, J. H., and Watson, M. W. (2006). Forecasting with many predictors. *Handbook of economic forecasting*, *1*, 515-554. https://doi.org/10.1016/S1574-0706(05)01010-4
123. Surana, A., Kumara*, S., Greaves, M., and Raghavan, U. N. (2005). Supply-chain networks: a complex adaptive systems perspective. *International Journal of Production Research*, *43*(20), 4235-4265. https://doi.org/10.1080/00207540500142274
124. Tibshirani, R. (1996). Regression shrinkage and selection via the lasso. *Journal of the Royal Statistical Society: Series B (Methodological)*, *58*(1), 267-288. https://doi.org/10.1111/j.2517-6161.1996.tb02080.x
125. Trapero, J. R., Pedregal, D. J., Fildes, R., and Kourentzes, N. (2013). Analysis of judgmental adjustments in the presence of promotions. *International Journal of Forecasting*, *29*(2), 234-243. https://doi.org/10.1016/j.ijforecast.2012.10.002
126. Tsai, B. H. (2013). Modeling diffusion of multi-generational LCD TVs while considering generation-specific price effects and consumer behaviors. *Technovation*, *33*(10-11), 345-354. https://doi.org/10.1016/j.technovation.2013.05.002
127. Thomassey, S. (2014). Sales forecasting in apparel and fashion industry: A review. *Intelligent fashion forecasting systems: Models and applications*, 9-27. https://doi.org/10.1007/978-3-642-39869-8_2
128. Thompson, James D. *Organizations in action: Social science bases of administrative theory*. Transaction publishers, 2003.
129. Thompson, C. G., Kim, R. S., Aloe, A. M., & Becker, B. J. (2017). Extracting the variance inflation factor and other multicollinearity diagnostics from typical regression results. *Basic and Applied Social Psychology*, *39*(2), 81-90. https://doi.org/10.1080/01973533.2016.1277529

130. Van den Broeke, M., De Baets, S., Vereecke, A., Baecke, P., and Vanderheyden, K. (2019). Judgmental forecast adjustments over different time horizons. *Omega*, *87*, 34-45. https://doi.org/10.1016/j.omega.2018.09.008
131. Van Rensburg, A. C. J. (2012). Can complexity analysis support business performance insight?. *South African Journal of Industrial Engineering*, *23*(2), 16-28. http://www.scielo.org.za/scielo.php?script=sci_arttextandpid=S2224-78902012000200004andlng=enandtlng=en
132. Verhoef, P. C., Kannan, P. K., and Inman, J. J. (2015). From multi-channel retailing to omni-channel retailing: introduction to the special issue on multi-channel retailing. *Journal of retailing*, *91*(2), 174-181. https://doi.org/10.1016/j.jretai.2015.02.005
133. Verstraete, G., Aghezzaf, E. H., and Desmet, B. (2019). A data-driven framework for predicting weather impact on high-volume low-margin retail products. *Journal of Retailing and Consumer Services*, *48*, 169-177. https://doi.org/10.1016/j.jretconser.2019.02.019
134. Watts, D. G., and Bates, D. M. (2014). Nonlinear Regression-Introduction. *Wiley StatsRef: Statistics Reference Online*. https:/doi.org/10.1002/9781118445112.stat03220
135. Wecker, W.E. and Ansley, C.F. (1983). The signal extraction approach to nonlinear regression and spline smoothing. *Journal of the American Statistical Association*, Vol. 78 No. 381, pp. 81-89
136. Weisberg, S. (2005). *Applied linear regression* (Vol. 528). John Wiley and Sons.
137. Weller, M., and Crone, S. F. (2012). Supply chain forecasting: Best practices and benchmarking study.
138. Williams, B. D., Waller, M. A., Ahire, S., and Ferrier, G. D. (2014). Predicting retailer orders with POS and order data: The inventory balance effect. *European Journal of Operational Research, 232*(3), 593-600. https://doi.org/10.1016/j.ejor.2013.07.016
139. Winsor, R. D. (1995). Marketing under conditions of chaos: percolation metaphors and models. *Journal of Business Research*, *34*(3), 181-189. https://doi.org/10.1016/0148-2963(94)00115-U
140. Wood, S. N. (2017). *Generalized additive models: an introduction with R*. CRC press.
141. Yee, T. W., and Mitchell, N. D. (1991). Generalized additive models in plant ecology. *Journal of vegetation science*, *2*(5), 587-602. https://doi.org/10.2307/3236170
142. Yu, W., and Ramanathan, R. (2011). Effects of firm characteristics on the link between business environment and operations strategy: evidence from China's retail sector. *International Journal of Services and Operations Management*, *9*(3), 330-364. https://doi.org/10.1504/IJSOM.2011.041103
143. Yu, W., and Ramanathan, R. (2012). Effects of business environment on international retail operations: case study evidence from China. *International Journal of Retail and Distribution Management*. pp. 218-234. https://doi.org/10.1108/09590551211207175
144. Verhoef, P. C., Kannan, P. K., and Inman, J. J. (2015). From multi-channel retailing to omni-channel retailing: introduction to the special issue on multi-channel retailing. *Journal of retailing*, *91*(2), 174-181. https://doi.org/10.1016/j.jretai.2015.02.005
145. Zelner, B. A. (2009). Using simulation to interpret results from logit, probit, and other nonlinear models. *Strategic Management Journal*, *30*(12), 1335-1348. https://doi.org/10.1002/smj.783
146. Zhao, K., Zuo, Z., and Blackhurst, J. V. (2019). Modelling supply chain adaptation for disruptions: An empirically grounded complex adaptive systems approach. *Journal of Operations Management*, *65*(2), 190-212. https://doi.org/10.1002/joom.1009

Chapter IV

The added value of the competition information in demand forecasting

4.1 Introduction

Factors, linked to industry consolidation, global conglomerates pressing the supply chain, making it difficult, are compelling companies to shift their attention from simple demand signals, such as trend and seasonality, to more dynamic demand signals, as price, sales promotions, competition actions, and economic indicators, to shape future demand based on sales/marketing tactics and to create more accurate demand response (Chase et al., 2013; 2016; Moretta Tartaglione al., 2018; 2019).

In the domain of demand forecasting, the common question is what and how many explanatory variables should be included in the forecasting model. In the academic research as well in the business practice the univariate model still takes the place (Raju, 1995; Hyndman et al. 2002; Taylor, 2003; Ord and Fildes, 2013; Ramos et al., 2015) where the future sales or demand is predicted on the history of past sales or demand to extract the patterns for the future predictions (Ma et al., 2016). The complexity of retail business influenced by the complexity of the supply chain system brings a lot of information that needs to be filtered, interpreted, and quantified. Nevertheless, the univariate forecasting model has been showing the functionality in the presence of many SKUs (Stock Keeping Units[4]), that has to produce a thousand models and being flexible to modify. The complexity of the process produces high dimensionality and difficulty to source valid explanatory variables for forecast accuracy (Ma et al., 2016). Recently, the dynamic environment creates the possibility for the improvements in SKU forecasting model. Authors investigate if the inclusion of the promotional events and changes of the price may improve the performance of the forecasting model (Cooper et al.,1999; Alon and Swadowski, 2001; Heerde and Leeflang, 2004; Gur Ali et al., 2009; Ramanathan and Muldermans, 2010; Trapero et al., 2013), the added competitive information (Huang et al., 2014, Ma et al., 2016), the investigation of the external factors (gas price, consumer price index) on the sales dynamics under system perspective (Moretta Tartaglione et al.,2019), the influence of the weather (Parsons, 2001; Murray et al., 2010; Bahng and Kincade, 2012; Arunraj and Ahrens, 2016; Bertrand and Parnaudeau, 2017, Moretta Tartaglione et al., 2019); incorporating the effect of the human judgment into the forecasting system (Fildes et al., 2009; Trapero et al., 2013; Baecke et al.,

[4] *Stock Keeping Unit*- an identifier for an item sold by a retailer.

2017; Broeke et al., 2018).

In the big retail chains with thousands of stores and a wide range of products, it is important to align the aggregated forecasts with the forecast on the individual SKU level. One product category may have the items that are rolling more often and leading the sales in contrast to the slow-movers where the automated forecasting models may mislead the results. For some product categories, the forecast model can be complex with the inclusion of the many underlying variables and for other categories, the simple predicting the history of sales may perform better. The tendency of the research in the domain of demand forecasting is to find the better performance of the model rather than to outline what variables are causing the demand dynamics. There is the research, in which the focus is to find the optimal balance between the best forecasting results and the optimal balance of the variables that will be less costly updating in the time.

Since the last two decades, analytical modes start to improve including other elements than a history of sales. The arising complexity in decision-making in big retail chains brings the necessity for improvements in demand forecasting methods (Fisher and Raman, 2018). Discovering which underlying factors are causing the switch in demand or the rise and fall in sales is necessary to anticipate. The research focused on the various type of promotions, the price discounts, on-shelf discounts, the way of the advertisement, what is happening when the promotions are applied to a product, and the situation when the products are not being promoted. Identifying the underlying demand variables can be a long-term process even in the recent period of advanced technology that can produce thousands of automated forecast models. Setting up the right number of explanatory variables that will be optimal for the performance and at the same time will be useful for setting up the future retail strategies is a constant challenging process. Category managers are often organizing pricing and promotional schedules across a wide assortment of goods. In these settings, it becomes necessary to estimate category demand systems that aim to flexible demand elasticities to accurately measure the impact of price promotions on category sales (Smith et al., 2019).

4.1.1 Motivation and objective of the research

This study is motivated by the research where it is investigated the value of the competitive price and promotion in forecasting the SKU level sales. Authors (Huang et al., 2014) refined the competitive information then incorporate it into the econometric forecasting model. The inclusion of the competitive information was overlooked in the previous research due to many indicators linked to the competitive information on each product (Huang et al., 2014). The proposed model shows superior performance including the competitive information in the sales forecasting. It is necessary to filter the piece of the competitive information linked to the product, where one product can have many aligned variables about the competition in that case the forecasting model can over-fitted and lack good results (Huang et al., 2014). The objective of Huang et al. (2014) is to give the retail managers the inference on how the marketing activities of the competitive and focal product can influence the sales and demand of the focal product. The result is that when the focal product is promoted the ADL model with the added competition variables is performing better and when the focal product is not promoted simple base-time lift method performs better. There are modeled the weekly time series for 122 products from 6 categories.

The paper by Ma et al. (2016) outlines some critical spots in the model of Huang et al., (2014), as manually setting the individual forecasting is not efficient way when there are thousands of items in store. Ma et al. (2016) propose the variables selection method (Lasso regression) at the multi-level process that can automatically choose the variables that can be inserted in the forecasting models. Another critical point is that Huang et al. (2014) pooled the index from the many stores and in fact when the stores are heterogeneous that procedure does not help a manager to allocate the SKU at the store level. Ma et al (2016) improve the model working with disaggregated data on the store level, although the analysis on SKU level can be complex because of the presence of noise this can be beneficial to retail managers to link directly the impact of variables. The improvement in the sales forecasting is the capturing the dynamic promotional effects over thousands of SKU individually, to overcome the high dimensionality of promotional effects-cross-category promotional information, practitioners can improve forecasts by using the additional information of intra and inter category promotional information, decreasing the cost of data storage. The forecasting model for the five top products for each category is built of a set of promotional intensity indexes for every category, including price indexes, display intensity indexes, feature advertising intensity indexes (Ma et al., 2016). As the development of its previous research, Huang et al. (2019) proposed the three-stage forecasting where the focus was to effectively use the promotional information and to include the structural change inside the model.

Research by Huang et al. (2014; 2019) and Ma et al. (2016) are underpinning points for this research which objective is to evaluate the additional value of the price and promotions of focal and competitive product on the focal demand. The belief relies on the necessity of the retail environment to better incorporate external factors in the decision-making process. It is not always necessary to include information in the forecasting model although it is becoming the crucial internal advantage of the company to ad hoc the business in the same market following what the competitors' activities are.

The objective of this analysis is to examine whether the incorporated information of competitors' price and promotion can add value to the forecasting model and what is the size and direction of the effect of competitive price and the promotion on demand of the focal product.

The data taken into considerations are coming from a large DIY (Do-it-Yourself) retailer in Europe that has an identical pricing strategy across all stores in the country. Besides, there are included competitive price and promotional information of the main DIY competitors in particular country. The goal is to investigate the direct effect of the main competitor's price and promotions on the focal demand of DIY SKUs. Further, this could improve the accuracy of forecasting models, which is beneficial for retail practitioners.

4.2 Literature background

4.2.1 Level of forecasting, promotions and seasonality

The accuracy of the forecasting depends on which level the future planning is done. The typically used are SKU level, brand level, and category level (Fildes et al., 2019). The smallest operational unit is the SKU level as the key element for the inventory planning at the store. SKU forecasts are usually implemented within and across the stores on a daily or weekly basis. The groceries, drugstores, or DIY chains have thousands of SKUs with huge combinations of Store/SKU combinations (Fildes et al., 2019). The application of the forecasting models at first depends on the level of the aggregation and the objective of the manager's plan, which is important to achieve the strategic or operational goal. The forecast on SKU level usually gives a better image of the frequency and effect of the applied promotions, and the size of the inventory stock. Moreover, the efficient short-term forecast demand on Store x SKU should lead to an increase in sales value, reduced stock, optimal distribution, and an increase in customer satisfaction (Fildes et al., 2019). The research conducted in the area of marketing and explains how the product exposure on the shelf or location of the product on the shelf can positively influence sales (Cooper et al., 1999). Cooper et al., (1999) investigate the effect of short-term promotional forecasting. The PromoCast tool based on the regression is used for predicting the demand on SKU level for the 95-store retail chain (Cooper et al., 1999). The important result is that this tool improved the forecasting performance when attributes of the SKU such as the producer, category, and sub-category. The promotions that were present longer can have mixed effects on the sales, especially for the slow-moving items. The research in the nineties was focused on how the promotional politics are influencing and changing the demand as well what is the effect on the retail strategies (Preston and Mercer, 1990; Raju, 1995; Christen et al., 1997; Cooper et al., 1999).

Recent research is more focused on advancing the forecasting tools and efficiency of using the tools in managing the overwhelming databases and comparing to the traditional linear models and time series univariate models. Additional internal and external variables are started to be inserted in the model, showing the significant impact on the demand. The more variables were taken into account the more complex the tools were necessary to be developed to catch the patterns in data. Alon and Swadowski (2001) introduced the advanced methods of neural networks to forecast the aggregate sales comparing to the traditional times series model and multivariate regression, Heerde and Leeflang (2004) investigated if the promotion bumps are creating the value from the managerial perspective. In a linear regression model, promotion effects are decomposed into cross-brand effects, cross-period effects and category expansion effects. Aburto and Weber (2007) suggest the improvement of demand forecasting merging the sarima models and neural networks where along with the past sales the maximum and minimum price on the micro-market, intermediate payment, holidays and seasonal effects. Natter et al. (2007) describe the decision support system in which the demand model for each SKU includes dynamic retail pricing and promotion planning. The tendency in the research was to develop the model the will improves at the same time the interpretability and accuracy. In the paper of Gur Ali et al. (2013) the new method for keeping the SKU inventory on the level of the demand in the presence of promotions, including also the number of competitors and price and promotional drivers. The method shows better performance increasing the sales and lowering the inventories.

Retail sales are characterized by a strong seasonality that can be cyclical, weekly, quarterly, or annually. Demand is usually higher during the weekends and low during the weekdays, high in the summer season and low in the winter (Fildes et al., 2019). Demand is also variable during the important Eastern and Christmas holidays affected by the high promotional discounts or some particularly calendar events during the year. It is necessary to catch seasonality pattern in the model, in some research with Boolean dummies for the week (Aburto and Weber, 2007; Huang et al., 2014) or the trigonometric function (Harvey et al, 2007; Huang et al., 2019); state-space model (Livera et al., 2010; Hyndman et al., 2008; Hyndman and Athanasopoulos 2015), additive Holt-Winters to linear models (Taylor, 2003).

4.2.2 Complexity and improvement of the forecasting methods

The recent trend in forecasting research is based more on improving the efficiency in manipulating the big data and producing thousands of automated models and comparing the performance to the traditional time series tools. As the complexity increases, it creates the need for the optimal balance between the explanatory variables and the accuracy of the forecasting results. The complex model may not always produce the best predictive results, but it can be useful for revealing the clear direction where the patterns are lying and changing the dynamics (Moretta Tartaglione et al., 2019). In the presence of the marketing chaos, new forecasting tools will be developed to try to maintain complexity efficiently so that the more complex model becomes the most optimal model (Moretta Tartaglione et al., 2018) in that case the increase in the complexity will tend to include the numerous internal and external factors that will tune the best predictive results for each demand of the product on the shell. Advanced statistical tools are developed to handle high-dimensional data and numerous variables in tuning the precise and accurate forecasting results. Initially, neural networks are performed for one month ahead forecast and the multiple forecasts (12 months ahead) (Alon and Swadowski, 2001); Aburto and Weber (2007) used the additive combination of the neural networks and sarima models where the residuals of the time series are trained and the promotions in the model to forecast the grocery sales. In the paper of Silva et al. (2018) the machine learning tools Artificial Neural Network (ANN) and Support Vector Machines (SVM) with traditional forecast methods (simple moving average, weighted moving average, exponential smoothing, Box-Jenkins Arima model, seasonal variation, linear regression). The models include internal factors: demand, and substitute, and complementary products; external: the considered macroeconomic variables are food inflation, GDP, employment rate, minimum wage to commercial balance, and capital flow of the nation. In terms of implementing the advanced tools. Gur Ali et al (2013) compare the accuracy of 30 SKU sales predictions, implementing the new algorithm Driver Moderator Method from the heterogonous stores from the US and Turkey. Comparing to the benchmark models: the simple exponential smoothing method (SES) with a last-like promotion lift adjustment comparing to the regression trees and support vector machines with polynomial and RBF kernels. The main challenge is univariate time series forecasting (Ma et al., 2016) and there have been introduced advancement as the hierarchical group forecasting (Hyndman et al., 2015), exponential smoothing using the state-space model (Hyndman et al., 2002; Ramos et al., 2015), these techniques are well applied to companies that handle numerous SKUs and need to produce forecasts semi-automatically (Trapero, 2013). When the company needs to face a more complex environment when at the same time external factors can influence the consumer demand, the trend is to implement multivariate techniques that are possible to extract the individual relationships of demand with external factors. In the majority of research the benchmark model is the multivariate linear regression (Alon and Swadowski 2001, Gur Ali et al., 2013; Haupt, 2014; Moretta Tartaglione et al., 2019) and some addressed the nonlinearity and employ the nonparametric regression (Dell et al., 2014; Haupt, 2014; Moretta Tartaglione et al., 2019).

4.2.3 *Competition information and forecasting model*

The awareness of the possible influence of the competitive information in forecasting models appeared at the first time in the research for the category management. The impact of the promotional activities of the competitive products can be less effective than the promotional activities of the focal product (Hoch et al., 1995). Divakar et al. (1995) included the price and promotions of the three main competitors into a Bayesian VAR model to forecast the sales of the focal brand in the canned soup category then same authors, Divakar et al. (2005), also modeled the price, display, and feature of the three main competitors (e.g. Pepsi versus Coca-Cola) to forecast the sales of the focal brand.

The retail pricing strategies is influenced by several factors (Kopalle et al., 2009) as in channel competition that consists of the cross-store price competition, the cross-channel competition the positioning of the store, link to the other marketing-mix variables , consumer behavior that lead to demand uncertainty., understanding the online and multichannel retail pricing.

At the first stage Huang et al., (2014) filter the relevant competitive information using the factor analyses to construct the small number of diffusion index and at the second stage, on disaggregate level, incorporated the variables in the Autoregressive Distributed Lag (ADL) the time series technique for forecasting. The ADL model (Huang et al., 2014) with included competition's information generate better forecast results when the focal product is on promotion and in the non – promotion period the base time lift up sales model performs better without competitive variables. The critical point in the Huang et al. (2014) model can be the that the competitive information are pooled across all the competitive explanatory variables and represented as the indexes that can miss direct individually effect on the focal product.The weekly forecasting were performed for the 1831 SKU from 48 product categories. Base-time lift up was compared to ADL model in which beside the competitive price the trigonometric variables of the capture of the week and calendar event are included. ADL shows the better accuracy than base time lift up model. Ma et al. (2016) improved the method of the pooling promotional information over the multiple categories in order to overcome multidimensionality, using the Lasso multistage regression to build the promotional indexes for every category and the level of the competitor's discount and in that way capture the dynamic promotional information. The models is supported by many moderators: brand indicators, number of competitors, SKU recent share, lags of marketing variables, season indicators and their interactions with driver variables The novel method (Ma et al, 2016) outperform the ADL by Huang et al. (2014) where the intra and inter category interactions were present. The marketing activities can affect the dynamics of SKU demand an produce the large variations, which forecast can become challenging (Huang et al., 2019). In this research Huang et al. (2019) proposed the threes stage forecasting at first selecting significant focal and competition variables, including structural change made by the marketing activities; at second, building the models and at third estimating the model in different time periods. Results shows that performance of the each model depend on whether it works in non-promotional or promotional period.

4.3 Methodology

Previous studies in the retail sector mainly use linear statistical models, although in recent years there has been a greater tendency to use sophisticated models under the nonlinear assumption.

The aim is to identify if the linear method can identify the causal relationship between the external variables the competitor's discounts and improve the forecasting activities.

The linear model assumes that the regression function is linear between the inputs and output variable, providing the simple interpretability of how inputs affect the output (Weisberg, 2005). Linear methods, used primarily in the estimation of sales, prices, and promotions, can provide the dynamics information about market mechanisms (Chu, 2011).

The methodology consists of three phases: At first, selection of SKUs and variables for modeling; at the second phase for the selected SKUs the linear model and the interpretation of the competitor's promo effect on the focal demand and final phase checking the best predictive performance and whether the additional competition information improves the predictive accuracy or the model.

4.3.1 [5]*Data collection*

Data are coming from two database sources: the daily sales demand with the promotional price of European DIY retailer and the digital retail platform that follows daily the competitors' price and promotions related to particular DIY retailer in 2018 and 2019. The products segments Construction materials and Drainpipe accessories (Table 4-1.) and are chosen based on the available external competitor data and the point of interest for the DIY retailer to follow the matching competitor's price and promotions for related SKUs. The sales are aggregated for each SKU over all stores in the country, this was possible to do because the price is defined at the national level, and applied promotions are the same for each DIY store in the country. In Figure 4-1, it is possible to see the general demand dynamics over the 84 weeks for both segments. As the higher demand occurs in the product segment of Construction materials, the peak is influenced by the seasonality and holidays typically, as Easter and Christmas, related holidays in the country. The peaks in demand are following the increasing trend from the 1st to 22nd week in the year and decreasing trend 23rd week to 52nd.

[5] Data are collected under a confidentiality agreement with Vlerick Business School, Ghent, Belgium.

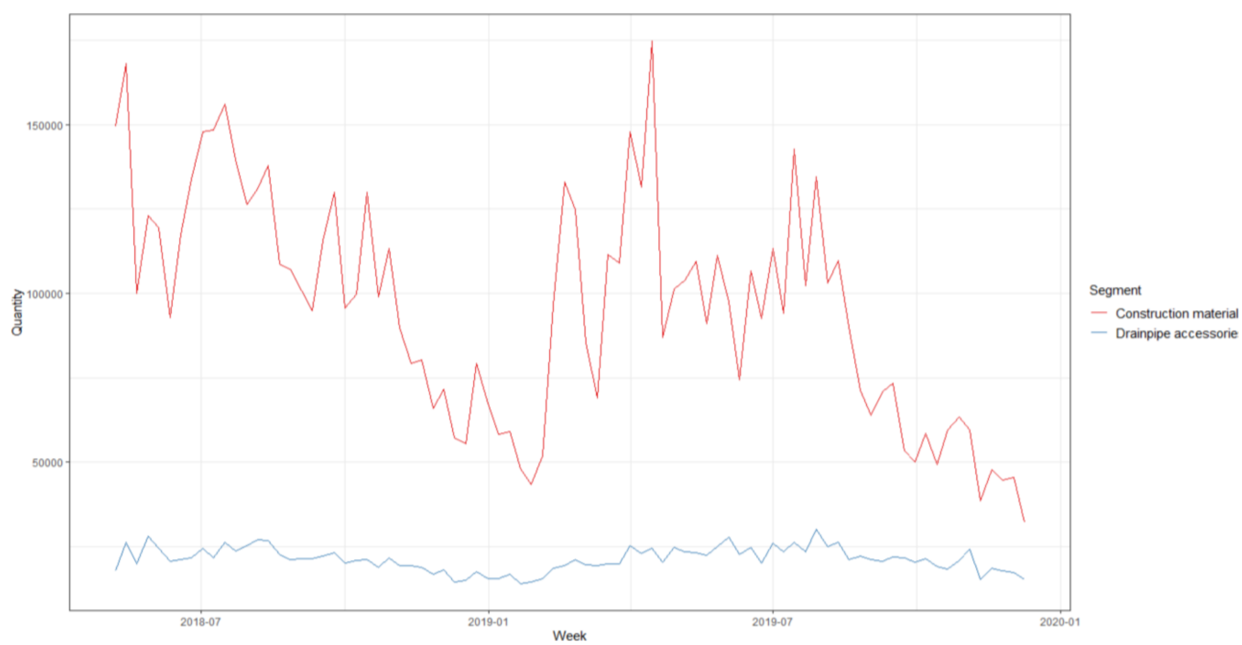
The peaks in demand can be explained by the days when the particular promotions are applied and that is one of the strategic forces of the DIY retailer.

Table 4-1. Data description of European Do-it-Yourself retailer

Type of retail stores	Number of stores	Product segment	Number of the selected SKUs	Weekly mean quantity>50	The pricing strategy	Weeks in promotion	Weeks in promotion at Competitors' store
DIY stores	178	Construction materials	18	54-26.082	Equal at national level	1-64	1-32
		Drainpipe accessories	4			2-6	1-3

Source: Author's elaboration

Figure 4-1. Dynamics of aggregated demand per SKU of products segments



Source: Author's elaboration

4.3.2 Method and results

Following the forecasting practice of the DIY retailer, daily data are summarised up to a weekly level where a total of 84 weeks are considered. Using the linear method, selected SKUs to need to meet the several conditions where it is investigated 1) the significant positive relationships between the focal promotional price and demand quantity, 2) the significant negative relationship between the competitor's promotions (%) and demand quantity, 3) weekly average sales quantity higher of 50 units, 4) SKU is in promotion more than one week in any of competitors' store 5) graphical presentation.

The linear model is developed based on available information from data about promotions and price from the dataset and linked external competition information. A dependent variable, the focal demand and independent variables, two weeks lag, three weeks lag, focal promotional price, percentage promotional discount from two main competitors, upward and downward trend, and dummy holiday (4.1).

$$
\begin{aligned}
Demand_{f(t)} = {} & Intercept + \beta_1 lag_{t-2} + \beta_2 lag_{t-3} + \beta_3 promo\ price_{f(t)} \\
& + \beta_4\ promo\ discount\ competior\ 1\ (\%)_{c(t)} \\
& + \beta_5 promo\ discount\ competior\ 2\ (\%)_{\ c\ (t)} + \beta_6 upward\ trend_{f(t)} \\
& + \beta_7 downward\ trend_{f(t)} + \beta_8 dummy\ holiday_{f(t)}
\end{aligned}
$$

(Equation 4.1)

In particular, it is studied (4.1) if the promotions of two competitors at time t (*promo discount competitor 1 (%) _c(t), promo discount competitor 2 (%)_ c (t)*) and internal promotions, actually the reduced price due to the promotions (*promo price_f(t)*) influence the weekly focal demand at the time t (*Demand_f(t)*). Besides, in the multivariate linear model, it is included two and three week's lags of demand (*lag_(t-2), lag_(t-3)*), and variables to control exogenous effect which is not explained by other variables (*upward trend f(t)), a downward trend(t)*), defined by the overall dynamics of weekly sales where the increasing trend is in the first six months and a decreasing trend in the second half of the year. The additional control variable is the dummy variable presenting whether it was the calendar holiday or not in-country (*dummy holiday_f(t)*). The first step was to select the SKUs that have sales of more than 50 weeks and available competitors' information about applied promotions at the same period. In this group of 275 SKUs from 178 stores, the linear method (4.1) is applied to get the insight of the coefficients and analyzed the sign of the relationships between the independent variables and dependent variable the focal demand.

In Table 4-2. the outcome of the linear model results that the major significant cases show expected behavior as 1) positive relationships between the two and three weeks lagged demand and focal demand at the time t, 2) negative significant relationship between the discounted price and focal demand, 3) the negative relationship between the competitor's promotional discounts (percentage) and focal demand. The expected behavior is the one where the positive influence on the demand is affected by the delayed effect of either price of promotions. During the promotion time, the demand is expected to increase due to the good offer that boosts the purchase amount per customer of the promoted items (Wang et al., 2015). Usually, the effect of the promotion today is delayed effects on the demand up to two-three weeks further. It is expected that the delayed demand of two and three weeks influences the demand of today. Sales reduction can last one or more weeks before the sales go back to baseline (Wang et al., 2015). It is expected the negative effect of the increased price or reduced/removed promotions on the focal demand. The competitor's reaction to imposed promotion is setting the promotion, as the ability to react on the promo-price intensity and consumer sensitivity to the price (Steenkamp et al., 2005). It is expected behavior if the competitor increases the promotions at the time *t* the focal demand will decrease at the time *t*.

At the second level, the top 22 SKUs are selected from 178 DIY stores in one country, two product segments the construction materials and drainpipe accessories. The rules for selecting the articles are that weekly average demand is higher than 50 units, that SKU is, at least, one week in promotion at one or another competitors' store, and that there is, at least, one significant negative relationship between of the competitors' promotional variables and focal demand (Table 4-2. and Table 4-3.). A graphical representation (Figure 4-2) it is possible to see the coherent behavior of the demand and the applied internal and external promotions. From Figure 4-2 it is possible to see the example of the individual SKU (281656) from the Construction materials' segment and demand dynamics (Sold quantity) and promotional price (Focal promo price) of the DIY retailer together with the competitor's promotional discount (Competitor's promotions). The overall graph shows that demand

(Sold quantity) is affected by the seasonality, it is higher during the summer period and lowers during the winter period. The demand is also affected by the change of the promotional price (Focal promo price), when the price is lower the demand tends to increase and reach the peak. The point of interest is to observe the dynamics of the competitor's promotional discount concerning the focal demand. Looking at the interval between 07-2018 and 08-2018 when there is a high presence of promotions, the introduction of the competitor’s promotional discount line affects the slight decrease in focal demand in the weeks forward producing the decreasing trend in demand. The decreased effect appears while the internal promotions are still active. The graphical presentation supports the outcome of the linear method in which occurs the negative coefficient between the demand and promotional price as well as demand and the competitor's promotional discount.

Table 4-2. Results of linear model - positive and negative significant coefficients β

Coefficients β	Negative significant β	Positive significant β	Complete cases	Left out variables
Intercept	1	132	275	0
lag(t-2)	5	132	275	0
lag(t-3)	1	94	275	0
promo price f(t)*	74	9	275	0
promo discount competitor 1 (%) c(t)*	12	6	233	42
promo discount competitor 2 (%) c(t)	18	3	66	209
upward trend f(t)	13	26	275	0
downward trend f(t)	59	6	275	0
dummy holiday f(t)	19	3	275	0

*f(t) – focal at the time t

*c(t)-competitor’s at the time t

Source: Author’s elaboration

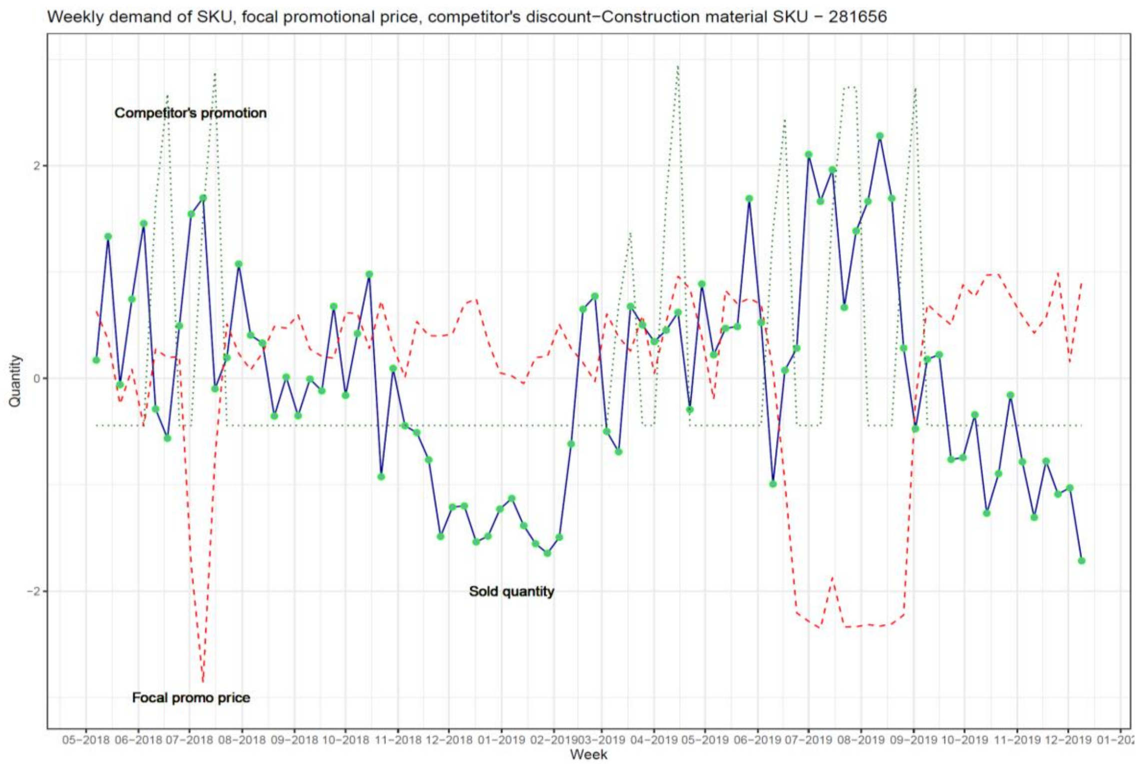
Table 4-3. Selected SKUs, the average weekly demand, number of weeks when SKU is on promotion at the DIY retailer and its main competitors

Product segment	SKU_nr	Weakly quantity (mean)	*N of weeks/focal promotion	N/weeks/promotion/ Competitor _1	N/weeks/promotion/ Competitor _2
Construction materials	219950	26.082	6	2	25
Construction materials	397138	8.679	5	0	32
Construction materials	221073	6.355	5	2	19
Construction materials	339665	5.444	1	7	17
Construction materials	300490	4.740	2	0	10
Construction materials	281656	3.166	3	11	14
Construction materials	132446	1.298	20	0	29
Construction materials	132447	769	19	0	29
Construction materials	346592	763	4	0	3
Construction materials	339644	723	3	7	37
Construction materials	609235	477	2	0	9
Construction materials	102476	467	3	0	2
Drainpipe accessories	463141	372	2	3	0
Construction materials	258428	331	6	0	3
Drainpipe accessories	889020	154	17	1	0
Construction materials	339643	123	49	2	27
Drainpipe accessories	389674	90	6	2	0
Construction materials	174189	86	64	0	23
Construction materials	134494	85	51	0	2
Construction materials	493304	77	35	2	5
Drainpipe accessories	19208	65	6	2	0
Construction materials	102905	54	34	0	2

*N-number

Source: Author's elaboration

Figure 4-2. Weekly demand of individual SKU, promotional price, competitor's promotion*



*standardized: x-mean(x)/standard deviation(x)

Source: Author's elaboration

4.4 Discussion

Regression analysis is not always the demonstration of the causality; it is the result of research design as well the identification strategy (Ziglari, 2017; Thompson, 1992). Multiple regression can give us better insight into the dynamics of the phenomenon by interpretation of beta coefficients and structure coefficients. (Ziglari, 2017).

The observed association between the predictor and the response leads to intuitively the result of causal relationships, in other words, if *X* causes the changes in *Y*, if and only if there is at least one unit in that population for which intervening in the world to change *Y,* that produces the correlation between the *X* and *Y* (Keele et al., 2019).

However, the challenge can be if the causation is moving in another direction, then the association is non-directional. The relationship between the *X* and *Y* can have the combination of causal and no causal components that need to identify at the first, the techniques supporting the identification of the matter of the causality are non-parametric function that can isolate the causal effect without forcing the functional form, and on the other way the graphical representation (Keele et al., 2019). Graphical representation can be used when there is doubt of the potential causal interpretations of estimates that usually depend on the causal assumptions (Keele et al.,2019). The problem with the causal interpretation is that can be based on untestable assumptions and then the report can be doubtful, so it is necessary to identify how many estimates can be given the causal interpretation (Keele et al.,2019).

In the particularity of this research, the causal assumption is that the competitor's promotions influence the weekly change in demand, besides the internal predictor variables as the two and three week's lags, the promo price, upward and downward trend, and dummy holiday. The influence of the competitor's promotions has been studied from several aspects: 1) the retailers are in channel competitors and the main threats for the retailer. 2) the graphical causal interpretation between demand and competitor's promotions 3) the expected negative linear relationship between the demand and the competitor's promotions, the parameter coefficient from the linear regression. The following aspects are based on the previous research in exploring the competitor's effects on the retail pricing. As the retail price affects the demand directly it is necessary to reflect on how the competitor's activity is taking the part in the retail price strategy. In literature, there is significant research on how the competitor affects the retail pricing (Kumar and Leone, 1988; Shankar and Bolton, 2004; Richards and Hamilton, 2006; Nijs et al., 2007; Kopalle et al., 2009). Each research demonstrates that competition is the key element for the price strategy in retail stores. In particular, Shankar and Bolton (2004) show that the competition activity is determining the retail price in four dimensions: consistency, price-promotion intensity, price-promotion coordination, and relative brand price. To set up the retail strategy the monitoring the in-channel competitor can help to create the optimal price that can adapt to the changes in the market. There are a lot of similarities among in-channel retailers such as the question of the location, the price and the promotion levels, private label strategies that can be the reason why the retailer follow and lean on each other to make variations in pricing strategy (Richards and Hamilton, 2006, Kopalle et al., 2009). The most of research used weekly data to follow the effect of the variation in competitive prices and promotions (Kumar and

Leone, 1988; Shankar and Bolton, 2004; Nijs et al., 2007) and to demonstrate that supermarket retailers tend to compete for the market share using at first price and variety and that there is evidence of the higher elasticity of substation of the products among retailers (Richards and Hamilton, 2006).

4.4.1 Retail price mixed strategies and competitors

In the research of Nijs et al. (2007) the study shows based on 123 weeks that the competitive price accounts for 5.5 percent of the retail price variance, in terms of the demand, the operational costs, and the category manager's strategies. These studies are taking into evidence the retail practice that is similar to the standard business of observed DIY retailer within this analysis. It is necessary to reflect on the creation of the retail price and its variation because the competition is taking a big part in setting the price strategy during the price wars or when the new product in one category enters the market (Kopalle et al., 2009). Kopalle et al. (2009) state that the price dispersion is the result of the model and ordered levels of the competitor's price on the market will randomly fluctuate over time. This is also reflecting on the demand and it is presented that sometimes the demand and competitor's price has the unusual behavior producing the nonlinear relationship between the competitor's promotions and consumer demand. It is quite studied in the literature about the retail price mixed strategies of high and low prices, today mostly present in online retailing affecting the price-promotion intensity. In this analysis, it is evident that the price-promotion intensity of leading products often on promotions directing the demand and producing the high variations during the seasonal sales (Figure 4-2). The optimal retail price depends on the competitor's reaction or aggressiveness of competitive pricing, it can have a significant impact on the optimal pricing and in the markets where competitors are sensitive to the price changes, the optimal modeling should adapt accordingly (Shankar and Bolton, 2004; Natter et al, 2007). Usually, leading products in one category that are highly monitored by the competitors should be excluded from the automated optimal price modeling. The competitive information is not relevant to insert if there is no interest by the market and profit leaders. In the research out of the 275 articles two product categories there are selected 22 products (Table 4-3.) that competitors react imposing the promotions that directly influence the focal demand. In the next example, it is observed the change of the focal promo price with the competitors' promotions. In Figure 4-3. it is possible to see the high variations of demand, focal promotional price, and competitors discounts in the seasonal sales periods (July-August), and the evident effect of the observed retailer to follow the competitor's actions where after imposing the discounts the retailer reduced the price. Monitoring can be vice versa, knowing that in that period the observed retailer will reduce the price due to the seasonal sales, the competitor anticipates the discounts and demand decreases slowly. The effects of the competition's reactions are possible to see in the weeks after where its demand curve has slower reactions to the internal promotions than before. The slowdown effect of the demand can be also related to the unobserved variables as the type of the products, the brand, the availability on the shell, weather, etc. In this research it is concerned that the retailer and its main competitors operate in the “business as usual”, saturated market of “do-it-yourself” retailers, in the product segments where it is high elasticity of the substitution, the low level of the affiliation to the brand, high costumer’s sensitivity to the price.

When competing starts between the in-same-channels retailers, the intensity of reaction can range from accommodating to no reaction or strong retaliation (Steenkamp et al., 2005). No reaction is the less common and accommodating is the least common (Leeflang and Wittink, 1996; Steenkamp et al., 2005; Nijs et al., 2007). Because of the lack of time and limited span of attention, managers take only certain actions as a response to the competition. Usually, price and promotions strategies are the most visible actions and generate stronger competitive retaliation than any other action. Steenkamp et al. (2005) suggest that managers use instant actions as price-promotions to respond in a short time to the competitive attack. The price –promotion strategies are the visible, sensitive and immediate instrument that can affect the profit margins and their effect bring the results quickly. That is why the price-promotions actions have the same effect on changing the demand of the product when the sales are affected by the seasonality. In the literature, distinguished three underlying factors motivate the behavioral drivers to the competition attack. Those drivers are the awareness of the competitive action, motivation to respond and able to react (Chen et al., 1992; Chen, 1996; Steenkamp et al., 2005)

These drivers influencing directly the competitive reaction to the price and promotions and result in certain reaction patterns.

4.4.2 Relation between competitors' promotions and demand

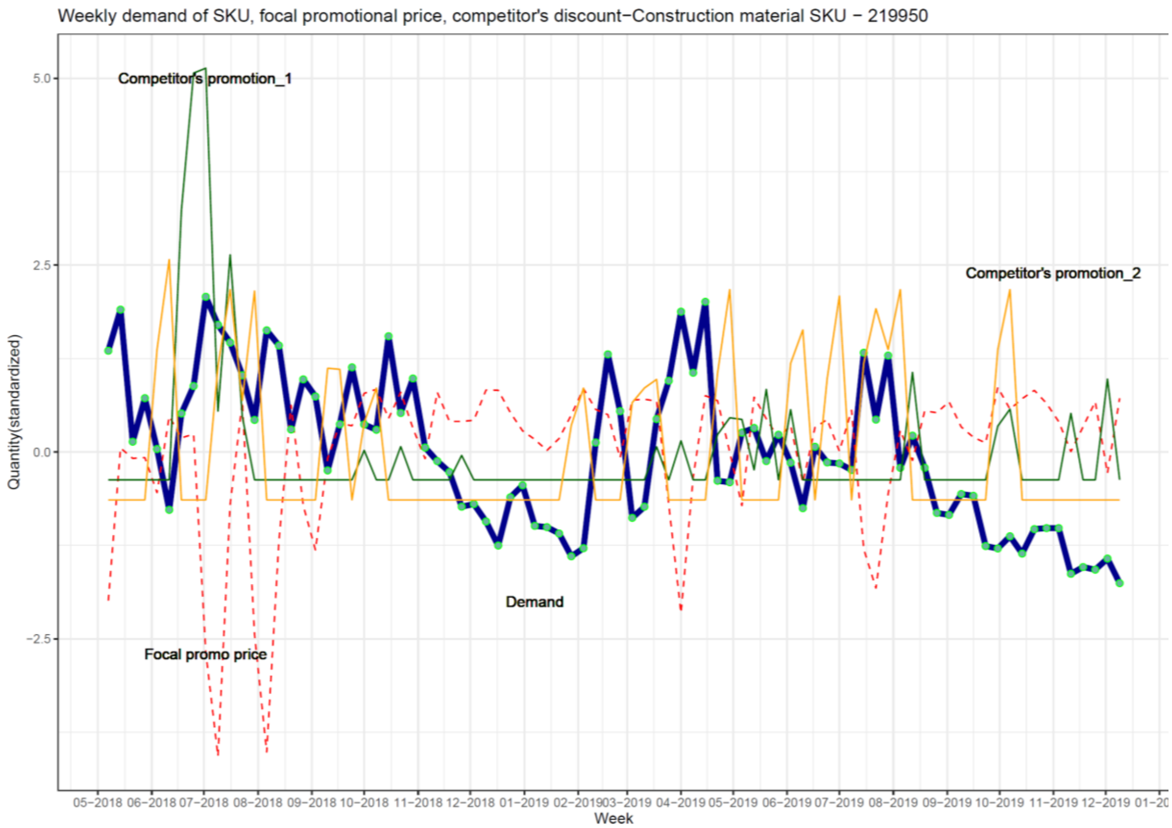
According to the linear model (4.1), the graphical representation (Figure 4-3) is reflecting the functional form created between the external variable of competitors' promotion where the negative coefficients tend to explain that increasing the promotions for one percent the demand tends to decrease to the value of the beta coefficients keeping all other predictors fixed. This is the particular example of the model (4.2) where it appears the significant negative sign of the coefficient related to promotions at the *competitor_2* and positive significant coefficient at the *competitor_1*. This is due to the different promotional strategies at both competitors where the at the case of the *competitor _1* the product is on promotion for two weeks while at *competitor_2* the product is on promotion 25 weeks. The positive significant parameter can be explained by the potential same promotional strategies the retailer and its *competitor_1* applying, matching the period of applying the seasonal discounts. As the products from the Construction materials, tend to be demanded during the seasonal periods, applying for the promotions at the same time may boost the demand of both retailers if the product is not available at one store may be at the competitor's store. The distance from the competitors can also have a significant impact.

$$Demand_{219950(t)} = 77680.5571 + (-48552.6132) * promo\ price + 0.4056 * lag_{t-2} + 0.1250 * lag_{t-3} + (-37.4080) * upward\ trend + (-226.9549) * downward\ trend + (-5585.6310) * dummy\ holiday + (61730.2739) * promo\ discount\ competitor\ 2 + 99465.6356 * promo\ discount\ competitor\ 1$$

(Equation 4.2)

Monitoring the competitors' actions retailers can have the possibility to create different price-demand strategies and apply them according to the necessities. The nature of the action can be mixed, for the certain products, retailers rate can have the aggressive nature and attack the competitor responding fast applying for the special promotions and anticipate the competitors' action, on another way it can be accommodating playing as "cooperative" partner. It has been demonstrated that both actions can boost the demand and add value to the multivariate linear models.

Figure 4-3. Weekly demand in relation to focal promotional price, intensity of two competitors' promotions*



*standardized: x-mean(x)/standard deviation(x)

Source: Author's elaboration

4.4.3 *Predictive performance*

In the first part, the discussion is related to the importance of the competitor's promotions and the possible effect on the demand. In the retail business, it is necessary to talk about planning and forecasting especially automated forecasting tools that can be useful for the thousands of articles in the assortment and planning of the stock. Prediction results are the basis for the retail supply chain and maybe not always precise they give the possibility to organization to look at the pattern of data and plan promptly the inventory stock.

The variables in the multivariate regression model and way of forecasting are previously discussed and aligned with the DIY retailer in the way where it can see if the additional external data benefit the forecasting models or no.

To test the predictive performance of the models with and without competitors' promo information the 20% of data or the last 17 weeks are used for evaluating the forecasting models. The hold-out sample for the testing has a mix of the promoted and non-promoted weeks to have the representative sample and more accurate predictions. For the evaluating forecasting models, Mean Absolute Error (MAE) and Mean Absolute Percentage Error (MAPE) are calculated. MAE (4.3) and MAPE (4.4) are computed across the time over each SKU individually.

$$\text{MAE} = \frac{1}{N}\sum_{t=1}^{N}|y - \bar{y}|$$

Equation 4.3

$$\text{MAPE} = \frac{1}{N}\sum_{t=1}^{N}\left|\frac{y - \bar{y}}{y}\right|$$

Equation 4.4

MAPE is frequently used in evaluating the forecasting models in business practice and when the quantity to predict is known to be far away from 0 (Myttenaere et al., 2016). Two models, with and without the external competitors' promotion for the selected 22 articles are evaluated on the hold-out sample of the last 17 weeks. To compare the models with and without the competitor's promotion difference between the MAE and MAPE models is calculated (Table 4-4).

Table 4-4. Evaluating the performance of models on hold-out-sample (17 weeks)

Article (nr)	***MAPEc**	***MAPEnc**	**MAEc**	**MAEnc**	****Difference MAPE**	****Difference MAE**
889020	24.85	25.35	29.34	29.84	-0.5	-0.5
19208	16.38	17.19	9.56	9.87	-0.8	-0.31
609235	21.38	20.47	80.54	78.01	0.91	2.53
389674	20.34	21.28	15.9	16.52	-0.94	-0.61
463141	28.32	29.68	102.69	104.79	-1.36	-2.1
339665	19.38	20.75	792.65	847.14	-1.37	-54.49
339644	37.21	35.43	190.38	210.07	1.79	-19.7
219950	72.52	74.47	8711.26	8979.13	-1.94	-267.87
397138	54.35	39.25	2438.68	1776.2	15.1	662.48
346592	34.27	36.51	109.93	118.23	-2.24	-8.3
258428	28.17	30.93	47.9	50.96	-2.76	-3.06
102476	16.97	19.77	73.6	83	-2.8	-9.39
281656	31.68	34.55	621.37	662.04	-2.87	-40.67
174189	63.25	66.99	25.08	31.88	-3.74	-6.81
339643	88.01	49.99	57.72	37.83	38.02	19.89
134494	34.82	30.44	23.01	20.35	4.38	2.66
300490	51.28	45.9	1368.59	1220.29	5.38	148.3
221073	44.71	39.01	1504.59	1340.52	5.7	164.07
102905	141.64	147.56	26.98	27.7	-5.92	-0.72
493304	70.61	64.3	25.26	22.61	6.32	2.65
132446	Inf	Inf	1247.02	1228.53	Undefined	18.49
132447	Inf	Inf	524.12	506.81	Undefined	17.31

*multiple by 100

(MAPEc-MAPE nc), (MAEc-MAEnc); **c-with competiors' information, **nc**-no competitors' information.

Source: Author's elaboration

Results of the evaluation of models on the test set show that 12 models out of 22 with the external competitors' promotions produce less MAPE and MAE errors than the model without external variables. The linear models that have the major difference in MAPE and MAE are related to SKUs that are more than 10 weeks on promotion at the competitor's store either internal store (*102905, 174189, 281656*). The more SKUs are on promotion, the better evidence of testing the models will be. Even at this small portion of available SKUs, it is possible to see the advantage of assuming the competitors' promotional activities inside the forecasting model.

4.5 Conclusions and future research

In this work, the aim is to identify and quantify the effect of the competitor's promotions on the focal demand. The previous research (Huang et al., 2014; 2019) considers the competition's information in terms of latent variables or indexes where there the forecasting model can suffer from the loss of the information. Results of this research give a clear representation of the relationship between the competitor's promotional actions and demand, the carry-over effect of the internal promotions, and a comparison of the accuracy of the models with and without the competitor's information. The forecasting using the linear method, where SKUs are selected from two product segments, 22 SKUs are a satisfying condition where the model underlines the effect of the competitor's promotion and demand. This approach is based on the simple linear time series methods taking the lag effects of the promotions and evaluating the effect on the demand from the internal and competitor's promotions. In literature, it is possible to find the research based on the sophisticated methods that better capture the effect of promotions (Cooper et al.,1999; Alon and Swadowski, 2001; Aburto and Weber 2007; Silva et al.,2018) used the additive combination of the neural networks and sarima models to capture the promotions for forecasting grocery sales. Besides, these methods are complex for interpretation and often do not take into account the leftover of the promotions and dismiss the competition's information (Huang et al., 2014).

This research is starting with evaluating and selecting the potential SKUs that influence of competitor's promotion will affect the dynamics of demand. In the rigorous selection of 22

SKUs based on the intensive sold quantity and applied promotions, it has been discussed the relationship between the competitors'promotions and the demand then the comparison of the predictive accuracy of models with and without competitors' information. It has been shown the trend: while the SKU is in promotion more weeks, the competitors' promotional effect will be capture better and improve the accuracy of the forecasting model. The limitation of this research is data; particularly low demand for the chosen product segments in the given period so the fewer possibilities to forecast for more SKUs. Although, the conclusion that can be outlined that the leading SKUs from the one product segment should be left out from the aggregated forecast for the whole product segment and allowed to evaluate the effect of the competitor's activities on the given portfolio of SKUs. That will give the manager opportunity to better validate the demand in the two weeks and focus the attention on managing the portfolio of SKUs that can be in a competitive advantage. Nevertheless, the future work will be the focus on testing the model on the different product segment especially SKUs with high demand; including the more promotional variables that are the result of the company's politics and employing, comparing with the models that relax the linear assumption; at final, testing models on same SKUs in different settings as different country and competitors.

4.6 References

1. Aburto, L., and Weber, R. (2007). Improved supply chain management based on hybrid demand forecasts. *Applied Soft Computing*, *7*(1), 136-144. https://doi.org/10.1016/j.asoc.2005.06.001
2. Alon, I., Qi, M., and Sadowski, R. J. (2001). Forecasting aggregate retail sales:: a comparison of artificial neural networks and traditional methods. *Journal of retailing and consumer services*, *8*(3), 147-156. https://doi.org/10.1016/S0969-6989(00)00011-4
3. Arunraj, N. S. and Ahrens, D. (2016). Estimation of non-catastrophic weather impacts for retail industry. *International Journal of Retail and Distribution Management*, *44*(7), 731-753. https://doi.org/10.1108/IJRDM-07-2015-0101
4. Baecke, P., De Baets, S., and Vanderheyden, K. (2017). Investigating the added value of integrating human judgement into statistical demand forecasting systems. *International Journal of Production Economics*, *191*, 85-96. https://doi.org/10.1016/j.ijpe.2017.05.016
5. Bahng, Y. and Kincade, D.H. (2012). The relationship between temperature and sales: sales data analysis of a retailer of branded women's business wear. *International Journal of Retail and Distribution Management*, Vol. 40 No. 6, pp. 410-426. https://doi.org/10.1108/09590551211230232
6. Bertrand, J. L., and Parnaudeau, M. (2017). No more blaming the weather: a retailer's approach to measuring and managing weather variability. *International Journal of Retail and Distribution Management*. https://doi.org/10.1108/IJRDM-09-2016-0152
7. Chase Jr, C. W. (2013), *Demand-driven forecasting: a structured approach to forecasting*. John Wiley and Sons.
8. Chase, C. W. (2016). *Next generation demand management: People, process, analytics, and technology*. John Wiley and Sons.
9. Chen, M. J., Smith, K. G., and Grimm, C. M. (1992). Action characteristics as predictors of competitive responses. *Management science*, *38*(3), 439-455. https://doi.org/10.1287/mnsc.38.3.439
10. Chen, M. J. (1996). Competitor analysis and interfirm rivalry: Toward a theoretical integration. *Academy of management review*, *21*(1), 100-134. https://doi.org/10.5465/amr.1996.9602161567
11. Christen, M. Gupta, S. Porter, J.C., Staelin, R. and Wittink, D.R. (1997). Using market-level data to understand promotion effects in a nonlinear model. *Journal of Marketing Research*, pp.322-334. https://doi.org/10.1177%2F002224379703400302
12. Chu, F. L. (2011). A piecewise linear approach to modeling and forecasting demand for Macau tourism. *Tourism Management*, *32*(6), 1414-1420. https:/doi.org/10.1016/j.tourman.2011.01.018
13. Cooper, L. G., Baron, P., Levy, W., Swisher, M., and Gogos, P. (1999). PromoCast™: A new forecasting method for promotion planning. *Marketing Science*, *18*(3), 301-316. https://doi.org/10.1287/mksc.18.3.301
14. Dell, M., Jones, B. F., and Olken, B. A. (2014). What do we learn from the weather? The new climate-economy literature. *Journal of Economic Literature*, *52*(3), 740-98. https://doi.org/10.1257/jel.52.3.740
15. Curry, D. J., Divakar, S., Mathur, S. K., and Whiteman, C. H. (1995). BVAR as a category management tool: An illustration and comparison with alternative techniques. *Journal of Forecasting*, *14*(3), 181-199. https://doi.org/10.1002/for.3980140304
16. Divakar, S., Ratchford, B. T., and Shankar, V. (2005). Practice prize article—chan4cast: A multichannel, multiregion sales forecasting model and decision support system for consumer packaged goods. *Marketing Science*, *24*(3), 334-350. https://doi.org/10.1287/mksc.1050.0135
17. Fildes, R., Goodwin, P., Lawrence, M., and Nikolopoulos, K. (2009). Effective forecasting and judgmental adjustments: an empirical evaluation and strategies for improvement in supply-chain planning. *International journal of forecasting*, *25*(1), 3-23. https://doi.org/10.1016/j.ijforecast.2008.11.010
18. Fildes, R., Ma, S., and Kolassa, S. (2019). Retail forecasting: Research and practice. *International Journal of Forecasting*. https://doi.org/10.1016/j.ijforecast.2019.06.004
19. Fisher, M., and Raman, A. (2018). Using data and big data in retailing. *Production and Operations Management*, *27*(9), 1665-1669. https://doi.org/10.1111/poms.12846
20. Ali, Ö. G., Sayın, S., Van Woensel, T., and Fransoo, J. (2009). SKU demand forecasting in the presence of promotions. *Expert Systems with Applications*, *36*(10), 12340-12348. https://doi.org/10.1016/j.eswa.2009.04.052
21. Harvey, N. (2007). Use of heuristics: Insights from forecasting research. *Thinking and reasoning*, *13*(1), 5-24. https://doi.org/10.1080/13546780600872502
22. Haupt, H., Kagerer, K., and Steiner, W. J. (2014). Smooth Quantile-Based Modeling Of Brand Sales, Price And Promotional Effects From Retail Scanner Panels. *Journal of Applied Econometrics*, *29*(6), 1007-1028.

https://doi.org/10.1002/jae.2347

23. Hoch, S. J., Kim, B. D., Montgomery, A. L., and Rossi, P. E. (1995). Determinants of store-level price elasticity. *Journal of marketing Research*, *32*(1), 17-29. https://doi.org/10.1177%2F002224379503200104
24. Huang, T., Fildes, R., and Soopramanien, D. (2014). The value of competitive information in forecasting FMCG retail product sales and the variable selection problem. *European Journal of Operational Research*, *237*(2), 738-748. https://doi.org/10.1016/j.ejor.2014.02.022
25. Huang, T., Fildes, R., and Soopramanien, D. (2019). Forecasting retailer product sales in the presence of structural change. *European Journal of Operational Research*, *279*(2), 459-470. https://doi.org/10.1016/j.ejor.2019.06.011
26. Hyndman, R. J., Koehler, A. B., Snyder, R. D., and Grose, S. (2002). A state space framework for automatic forecasting using exponential smoothing methods. *International Journal of forecasting*, *18*(3), 439-454. https://doi.org/10.1016/S0169-2070(01)00110-8
27. Hyndman, R., Koehler, A. B., Ord, J. K., and Snyder, R. D. (2008). *Forecasting with exponential smoothing: the state space approach*. Springer Science and Business Media.
28. Hyndman, R. J., and Athanasopoulos, G. (2015). 8.9 Seasonal ARIMA models. *Forecasting: principles and practice. oTexts. Retrieved*, *19*.
29. Keele L, Stevenson RT, Elwert F (2019). The causal interpretation of estimated associations in regression models. Political Science Research and Methods 1–13. https://doi.org/10.1017/psrm.2019.31
30. Kopalle, P., Biswas, D., Chintagunta, P. K., Fan, J., Pauwels, K., Ratchford, B. T., and Sills, J. A. (2009). Retailer pricing and competitive effects. *Journal of Retailing*, *85*(1), 56-70. https://doi.org/10.1016/j.jretai.2008.11.005
31. Kumar, V., and Leone, R. P. (1988). Measuring the effect of retail store promotions on brand and store substitution. *Journal of marketing research*, *25*(2), 178-185. https://doi.org/10.1177%2F002224378802500206
32. De Livera, A. M. (2010). Automatic forecasting with a modified exponential smoothing state space framework. *Monash Econometrics and Business Statistics Working Papers*, *10*(10), 6.
33. Ma, S. H., Fildes, R., and Huang, T. (2016). Demand forecasting with high dimensional data: The case of SKU retail sales forecasting with intra- and inter-category promotional information. *European Journal of Operational Research, 249*(1), 245-257. https://doi.org/10.1016/j.ejor.2015.08.029
34. Moretta Tartaglione, A., Bruni, R., Bozic, M., and Cavacece, Y. (2018). How the environmental complexity affects forecasting in retail? Actual research trends. *Journal of Marketing Trends (1961-7798)*, *5*(2).
35. Moretta Tartaglione, A., Bruni, R., and Bozic, M. (2019). Exploring the retail industry environment using nonlinear analysis. *International Journal of Retail and Distribution Management*, Vol. 47 No. 4, pp. 453-47. https://doi.org/10.1108/IJRDM-09-2017-0224
36. Murray, K.B., Di Muro, F., Finn, A. and Leszczyc, P.P. (2010), "*The effect of weather on consumer spending*", Journal of Retailing and Consumer Services, Vol. 17 No. 6, pp. 512-520. https://doi.org/10.1016/j.jretconser.2010.08.006
37. De Myttenaere, A., Golden, B., Le Grand, B., and Rossi, F. (2016). Mean absolute percentage error for regression models. *Neurocomputing*, *192*, 38-48. https://doi.org/10.1016/j.neucom.2015.12.114
38. Natter, M., Reutterer, T., Mild, A., and Taudes, A. (2007). Practice prize report—An assortmentwide decision-support system for dynamic pricing and promotion planning in DIY retailing. *Marketing Science*, *26*(4), 576-583. https://doi.org/10.1287/mksc.1060.0217
39. Nijs, V. R., Srinivasan, S., and Pauwels, K. (2007). Retail-price drivers and retailer profits. *Marketing Science*, *26*(4), 473-487. https://doi.org/10.1287/mksc.1060.0205
40. Ord, J. K., and Fildes, R. (2013). Principles of Business Forecasting (International Edition). *South-Western College Publishing, Mason, Ohio*. https://doi.org/10.1016/0377-2217(94)00275-H
41. Parsons, A. G. (2001). The association between daily weather and daily shopping patterns. *Australasian Marketing Journal (AMJ)*, *9*(2), 78-84. https://doi.org/10.1016/S1441-3582(01)70177-2
42. Preston, J., and Mercer, A. (1990). The evaluation and analysis of retail sales promotions. *European journal of operational research*, *47*(3), 330-338. https://doi.org/10.1016/0377-2217(90)90219-2
43. Raju, J. S. (1995). Theoretical models of sales promotions: Contributions, limitations, and a future research agenda. *European Journal of Operational Research*, *85*(1), 1-17.
44. Ramanathan, U., and Muyldermans, L. (2010). Identifying demand factors for promotional planning and forecasting: A case of a soft drink company in the UK. *International journal of production economics*, *128*(2), 538-545. https://doi.org/10.1016/j.ijpe.2010.07.007
45. Ramos, P., Santos, N., and Rebelo, R. (2015). Performance of state space and ARIMA models for consumer retail sales forecasting. *Robotics and computer-integrated manufacturing*, *34*, 151-163. https://doi.org/10.1016/j.rcim.2014.12.015
46. Richards, T. J., and Hamilton, S. F. (2006). Rivalry in price and variety among supermarket retailers. *American*

Journal of Agricultural Economics, *88*(3), 710-726.
47. Shankar, V., and Bolton, R. N. (2004). An empirical analysis of determinants of retailer pricing strategy. *Marketing Science*, *23*(1), 28-49. https://doi.org/10.1287/mksc.1030.0034
48. Silva, A. T., Moro, S., Rita, P., and Cortez, P. (2018). Unveiling the features of successful eBay smartphone sellers. *Journal of Retailing and Consumer Services*, *43*, 311-324. https://doi.org/10.1016/j.jretconser.2018.05.001
49. Smith, A. N., Rossi, P. E., and Allenby, G. M. (2019). Inference for product competition and separable demand. *Marketing Science*, *38*(4), 690-710. https:/doi.org/10.1287/mksc.2019.1159
50. Steenkamp, J. B. E., Nijs, V. R., Hanssens, D. M., and Dekimpe, M. G. (2005). Competitive reactions to advertising and promotion attacks. *Marketing science*, *24*(1), 35-54. https://doi.org/10.1287/mksc.1040.0069
51. Taylor, J. W. (2003). Short-term electricity demand forecasting using double seasonal exponential smoothing. *Journal of the Operational Research Society*, *54*(8), 799-805. https://doi.org/10.1057/palgrave.jors.2601589
52. Taylor, T. A., and Xiao, W. (2009). Incentives for retailer forecasting: Rebates vs. returns. *Management Science*, *55*(10), 1654-1669. https://doi.org/10.1287/mnsc.1090.1045
53. Thompson, B. (1992). Interpreting Regression Results: beta Weights and Structure Coefficients are Both Important.
54. Trapero, J. R., Pedregal, D. J., Fildes, R., and Kourentzes, N. (2013). Analysis of judgmental adjustments in the presence of promotions. *International Journal of Forecasting*, *29*(2), 234-243.
55. Van den Broeke, M., De Baets, S., Vereecke, A., Baecke, P., and Vanderheyden, K. (2019). Judgmental forecast adjustments over different time horizons. *Omega*, *87*, 34-45. https://doi.org/10.1016/j.omega.2018.09.008
56. Van Heerde, H. J., Leeflang, P. S., and Wittink, D. R. (2004). Decomposing the sales promotion bump with store data. *Marketing Science*, *23*(3), 317-334. https://doi.org/10.1287/mksc.1040.0061
57. Wang, W., Xu, Q., Gao, C., and Liu, X. (2015). Supply Chain Collaborative Forecasting Modeling. In *LISS 2014* (pp. 317-322). Springer, Berlin, Heidelberg. https://doi.org/10.1007/978-3-662-43871-8_47
58. Ziglari, L. (2017). Interpreting multiple regression results: β weights and structure coefficients. *General Linear Model Journal*, *43*(2), 13-22.

Chapter V

Conclusions, limitations and future research

The last section of the dissertation offers a synthesis of the main conclusions obtained from the research. At first, the main concepts that contribute to the academic research, at second the practical implications of the research and, at final, the limitation, the future ideas for the development of research, and general considerations.

5.1 Academic contributions

The concept of analyzing the environment for retail forecasting is emerging in recent academic research, and the trend of conceptualizing the environment, the element that can define the business dynamics, is increasing. In the first part of the research (Chapter I) where the literature has been analyzed in a traditional way combining with network analysis of the relevant papers about the topic, the trend is passing from the past research which generalizes the concept of the environment and defining its main characteristics: complexity, uncertainty, and dynamism to recent where research observe market as heterogeneous where the level of the competition intensity depends on the competition hostility with the necessity for the continuous scanning of the environment. The contribution has outlined the concept of the environment in retailing where the decision-maker has to scan and filter the right patterns then observe and quantify the relevant relationships between environmental factors and the demand in the forecasting model. The concept of the retail environment outlines the trend in research to 1) improve the collaboration and information sharing inside the supply chain 2) centralize the forecasting decisions on the retailer’s side, 3) manage better the flux of data 4) to detect the pattern of the volatile demand 5) include the competition’s price and promotions 6) monitor the impact on weather 6) develop the flexible forecasting algorithm 7) consider the pattern from the online channel and internet sources 8) detecting sharply the consumer profile and anticipate shopping behavior. The research contributes to the change from the recent perspective in the literature that has seen the complex model less useful and efficient to the future perspective in which the innovative tools and digital transformation indicating that the complex forecasting model will be able to perform under thousand of variables and millions of observations smooth and efficient.

The second contribution (Chapter III) is in exploring the causality between external variables and sales using the flexible models that permit the nonlinear function (additive models). The patterns in sales have to be detected, especially when sales of large retail chains are affecting the environment. This analysis, using the various modeling approaches it is isolated and measured the causal effect of the macroeconomic variables (CPI, Fuel Price, Unemployment) and the weather conditions (Temperature). While these macroeconomic indicators can be correlated and interdependent in the real scenario, the general additive model isolates independently the relationships between external drivers and sales.

The research contributes to the statistical approach of resolving the collinearity of the entangled macroeconomic indicators permitting the latent variables (principal components) in the model instead of the variables themselves. In benefits of additive models is that nonlinearity, non-normal errors, and heteroscedasticity are automatically resolved by nonlinear techniques. (Soquero-Ruiz et al., 2014).

The third contribution to academic research is the evaluation of the competitions' promotions inside of the weekly demand forecasting. Previous research (Huang et al., 2014; 2019; Ma et al., 2016) studied the effect of competition's information blended in the latent variables where it is possible to lose the information about the effect of the real variable, this research (Chapter IV) reveals the direct effect of the competitor's promotion on the weekly demand on the focal SKU. The contribution relies on the necessity of the retail environment to better incorporate external factors in the decision-making process.

Developing the linear model, based on the internal reduced price, lagged demand effect, and main's competitor's promotions the results show the tendency: when the SKUs are more on promotions internally and at competitor's stores, the model with included competitor's promotions shows better predictive accuracy. The inclusion of the competition's promotion can add value to the weekly demand-forecasting model at the SKU level.

5.2 Implications for the retail management

The dissertation itself can be the guiding compass for the retail managers, related to the research design, the statistical approaches, and data used in the analysis. Precisely the retail manager can find the utility of the results of the research on the corporate and operational planning level.

On the corporate level, the retail manager has always to have the broader picture of positioning the retail chain in one market. The market in which retailers operate is the direct environment where the retail manager receives information about the actions of the competitors, customers, the related institutions, and the regulations of the distribution channels. The decision-maker (retailer) may shift the modeling strategy directing the business dynamics according to the self-purpose aims. Today, when the retail expanded business multinational where many decision-makers have possibilities to strategize the retail company the goal is slightly different: to balance optimally the market share, to keep the position, and to maximize retail chain's profit, regardless to the store and its performance. The concept of the retail environment (Chapter I) can help the retail managers to make their decisional schemes and focus on the right path of filtering the important information from the environment, whether it will be the collaboration and centralizing the forecasting decisions, following the competitor's actions or the detecting the consumer's shopping behavior and better manage the flux of data.

Retailers are focusing on delivering the product or service to diverse multichannel forms. The second implication (Chapter II) for retail management is understanding behind the customers' visits to different store type and their choice of the assortment, switching mood in the shopping behaviors as the critical issue for retail managers to create flexible strategies at the operational level. The results of the research (Chapter II) implicate that using the clustering of existing data in the retail company can deliver significant conclusions as that consumers prefer rather purchase in neighborhood markets that offer a limited assortment. The convenience stores show the highest level of sales profitability during promotions and school holidays periods on other-sided shopping centers, which are those that consumers prefer to, visit. The results of regression analysis on store clusters show how the size of the assortment can affect the sales per square meter: basic assortments stimulate sales per square meter and making convenient for the smaller "neighbor's" store to increase the profitability of the sales. The extended levels of assortment have negative effects in any type of store. Regardless of the type of store, retailer's main key tactics are promotions for boosting and predicting sales with

taking into account the geolocation of the store and its position to the competitors, the distance appears to have a significant effect on the sales in any type of store.

In contribution to the retail forecasting, the results (Chapter III) show that the nonlinear methods can reflect the real retail dynamics and that retailers cannot base decisions only on the internal indicators. The result of this research is to outline the necessity for the large retail chains to consider the changes in the macroeconomic environment. Looking at decision-making at the corporate level, the decision-maker has to evaluate the long-term positioning of the retailer on the market. Changing the perspective from linear to nonlinear, the decisional framework on how to evaluate the environmental effect can benefit the efficiency of the retail strategies. A pattern or relationship exists that can be identified in the environment and modeled with the tendency to continue. Each pattern or relationship must be identified and quantified before it is possible to create a forecast (Chase, 2016; p. 153). Results of the general additive models in which it is isolated and measured the causal effect of the macroeconomic variables (CPI, Fuel Price, Unemployment) and the weather conditions (Temperature) benefit better predictive accuracy. While these macroeconomic indicators can be correlated and interdependent in the real scenario, applying the general additive model, the function between the between external drivers and sales can be measured independently, that can give the real interpretation to the retailer how the sales are affected due to the changes in the macroeconomic indicators.

The last, the implications to the operational forecasting especially to the weekly demand forecasting (Chapter IV) on the SKU level (Stock Unit Level) show that by including the competitor's promotion in the demand forecasting the predictive accuracy improves on the test set. The results shows also that the dynamics of the SKUs that roll-out faster and have high demand implicates the effect of the competitor's promotions to be stronger and affect the focal demand. This approach can help the forecaster when planning, while the weekly forecasting should be different for the full product segment the leading products should represent the portfolio of articles that benefit the larger sales to the category management and on which promotional actions the competitors are sensitive and ready to react and affect the focal demand.

Nevertheless, the data wrangling and modeling with meta language *R* are taking the great part of the analysis in the dissertation especially the time, dedicated to developing efficient codes to manipulate the millions of observations, has become the precious added value to complete work.

5.3 Limitations and future aggregations

The general limitation to research in the dissertation is the lack of data to prove the additional testing of the hypothesis.

The main limitation of the research (Chapter II) in which it is investigated the relationships between the customer visits and retail format, the size of the assortment is the lack of information about the consumer's profile, age, gender, history of the purchase behaviors due to, the research cannot describe the reason behind the consumer's behavior. The validity of the results cannot, therefore, be generalized. Future research can be aimed at a combination of field studies together with post-sales analysis. This analysis, with the combination of the classification and regression tools it is possible to manage a large sample and deliver a considerable conclusion about dynamics in the retail context.

Regarding research in Chapter III, including the effect of the environmental drivers often means that the availability and collection of external data can be a challenge; the testing of causality is usually limited on what is available now and what has been collected and linked to the particular sales. The lack of adequate information on the link of the promotional discounts to sales is the burden of the research as well (Chapter III). Future research on the application of additive methods should consider the effects at the operational level (Store/SKU) isolating the external effect on the product segment or individual article. Analyzing the short-term effects weekly or daily and particularly in analyzing the drivers that managers should be an effective strategy for the particular store or the group of stores.

The limitation of this research Chapter IV is a particularly low demand for the chosen product segments in that implies the fewer possibilities to forecast for the more SKUs, when the model shows better performance on more SKUs, then the conclusions will become stronger. For the future work, the focus is on testing model throughout diverse product segments especially SKUs with high demand under intense promotional dynamics; considering a larger number of internal and external promotional variables, comparing to the tools that relax the linear assumption; at final, testing models on same SKUs in different settings as a different country and competitors' environment.

5.4 General considerations

Reflecting on the beginning of the dissertation and main research question (*MRQ*; p. 9) what is and how can be measured the retail environment effect on the sales and demand forecasting, it is possible to outline the final answers and considerations:

i. To deal with uncertainty and complex information, forecasting planning needs the scanned and filtered the drivers from the retail environment, focused on the conceptualized, collaborative and centralized decision-making, managing and analyzing data flux efficiently;
ii. Boosting the sales productivity and customer's visits may be reached managing the portfolio of retail formats and the levels of assortment, the combination of the clustering and regression tools can benefit the data insights and deliver meaningful results where smaller stores, optimal assortment, and competition distance have the significant effect;
iii. Inserting the external data in sales forecasting may be adding value to the retail manager at a corporate level, however, it is crucial to inspect the causality and detect the right external pattern using the right tools, in this case, the additive models that can isolate the nonlinear relationships;
iv. Following the competitor's actions especially in terms of price and promotional strategies can add value to weekly demand forecasting for in-channel competitors that based the profit on promotional discounts of fast-rolling SKUs.