Cannibalization Effects of Products in Zara's Stores and Demand Forecasting

by

Evelyne L. Kong

B.S., Mechanical Engineering
Ecole Centrale Paris, 2007

Submitted to the MIT Sloan School of Management and the Engineering System Division Department
in Partial Fulfillment of the Requirements for the Degrees of

Master of Business Administration
and
Master of Science in Engineering Systems

In conjunction with the Leaders for Global Operations Program
at the
Massachusetts Institute of Technology

June 2015

Signature redacted

Signature of Author ....
MIT Engineering Systems Division, MIT Sloan School of Management
May 8, 2015

Signature redacted

Certified by ....
Georgia Perakis, Thesis Supervisor
William F. Pounds Professor of Management Science
MIT Sloan School of Management

Signature redacted

Certified by ....
Bruce Cameron, Thesis Supervisor
Director of the System Architecture Lab

Signature redacted cturer in Engineering Systems

Accepted by ....
Maura Herson
Director of MIT Sloan MBA Program
MIT Sloan School of Management

Signature redacted

Accepted by ....
Munther A. Dahleh
William A. Coolidge Professor of Electrical Engineering and Computer Science
Chair, ESD Education Committee

Cannibalization Effects of Products in Zara's Stores and Demand Forecasting

By

Evelyne L. Kong

Submitted to the MIT Sloan School of Management
on May 8, 2015 in Partial Fulfillment of
the Requirements for the Degrees of Master of Business Administration
and Master of Science in Engineering Systems

Abstract

Because of short product life cycles, large product offerings and fickle consumer tastes, product demand in the apparel retail industry is volatile and difficult to predict. It is also thought to be impacted significantly by the presence of other products in the store, which may cannibalize or complement sales. This thesis explores the interactions among products in Zara's stores, and uses this information to improve the existing demand forecast. It focuses on attribute-based and price-based cannibalization effects for the trousers product category.

A necessary preliminary step to study cannibalization is to obtain an accurate classification of all products according to distinct features. We present such a classification for the trousers product category according to three main characteristics—fit, color and color lightness. The text mining and picture color detection techniques used here are of independent interest and can be used for the classification of other product categories.

We then use Zara's RFID system and sales databases to estimate cannibalization effects among products in stores. To this purpose, we introduce similarity factors and define a linear regression model with fixed and time effects. We obtain statistically significant cannibalization effects. However, these effects tend to account for only a small portion of demand variations.

Finally, we propose various improvements to the demand forecast method, incorporating display, availability and cannibalization effects. The resulting demand forecast represents a significant improvement over the base forecasting model. It reduces the forecasting error by over 10% for the sample stores in Madrid, Barcelona and London and by over 7% with only partial display information.

Thus the implementation of the present work will reduce overstock and lost sales. It also constitutes the first step towards product assortment optimization at Zara.

Thesis Supervisor: Georgia Perakis
Title: Williams F. Pounds Professor of Management Science - MIT Sloan School of Management

Thesis Supervisor: Bruce Cameron
Title: Director of the System Architecture Lab, Lecturer in Engineering Systems - MIT Engineering System Division

This page is intentionally left blank

Acknowledgments

During six months in 2014, I had the opportunity and immense pleasure to work in La Coruña at Inditex's headquarters. This was possible thanks to the long-standing collaboration between Zara and the Leaders for Global Operations program at MIT. I feel blessed for having the chance to discover the fashion industry at Zara and for the friends I have made during my internship. The environment at Zara has been very supportive of my work since my very first day in June 2014 and I am very grateful for the exciting project the distribution team has given me.

I would like to thank my supervisors and colleagues Martín Nóvoa Vidal, Ane Insausti Altuna and Iván Escudero Rial, who were always available for any questions on a daily basis. Their insights and experiences were invaluable for my work and I could not have done this thesis without their contribution. I am thankful to Alberte Dapena Mora who did a fantastic job helping me understand the structure of Zara's database. Thank you also to Begoña Valle Vilela, Sergio Marquez, Patricia Arribas Davila, Yago Vera Cuartero and the rest of Zara's distribution team for creating a joyous atmosphere at work and for constantly bringing pastries from Panadería Rozas.

The project would not have been possible without the support of the LGO Program and of my advisors. Georgia Perakis has been instrumental in providing structure to my thesis and I would like thank her for giving a considerable amount of her time on a weekly basis. I am very grateful to Bruce Cameron, for his invaluable advice on how to define the objectives of my internship early on, and how to keep my focus throughout.

Finally, my family and friends have always been present for me during my six months in Spain. I am particularly thankful to my amazing husband, who was there at every step of my experience at Zara and has always encouraged me to pursue my goals in life.

This page is intentionally left blank

Contents

List of Figures

List of Tables

1 Overview

1.1 Thesis Motivation and Problem Statement

Every year, Zara delivers 11,000 distinct products to 2,000 stores across the world. With new products arriving in the stores twice a week and an affordable product offering, Zara has successfully attracted customers to its stores to shop for the latest fashion trends and has become the world largest fashion retailer. To respond quickly to the demand of its customers, the company has developed a very responsive supply chain. Still, even with a design to store delivery lead time of a few weeks, predicting customers' volatile demand is key to Zara's success. Given the short life cycle and fickle customer taste, it is not uncommon to observe forecasting errors of 60% of more at the SKU level in the apparel industry. This leads traditional fashion retailers to sell up to 50% of the merchandise on markdowns ([1]).

Therefore, an accurate forecast of the store-level demand is essential to minimize stock-outs and overstocks. In Zara's current demand forecasting method, each product is considered in isolation. In practice, however, products interact with one another in the store. For example, the presence of a product on the shop floor may reduce or increase the sales of other products. Similarly, there can be cross-price effects, with sales of a product being affected by the prices of other products. Product cannibalization within a store is the subject of considerable research ([2, 3]), especially for the study of product assortment for the supermarket retail industry. However, no significant research has been done on finding evidence of cannibalization among products in the apparel retail industry.

Identifying cannibalization patterns among products requires access to detailed data about product attributes, sales, inventory or display information. For example, a retailer may not perfectly know if a given product is present on the shop floor, as data only show the total inventory in the store, which includes inventory on the shop floor and in the stock room. However, the recent deployment of the RFID technology in Zara's stores has given us access to a wealth of information which was not available before and is crucial to know the product assortment at any given time in a store. The current work presented here is one of the first projects using the new RFID database, which can ultimately help optimize product assortment planning for Zara.

Thus, the present thesis focuses on understanding these cannibalization and complementarity effects among products within Zara's stores and how this information can be used to improve the store demand forecast at the product level.

1.2 Proposed Approach and Contributions

To understand cannibalization and improve Zara's product demand forecast, we adopted the following approach to the problem:

- ***Collect data and build a new product classification***. The project requires using store sales data and information collected by Zara's new RFID system. Because we expect more significant cannibalization or complementarity effects among similar products, we created a new product classification that allows to identify easily similar or dissimilar garments. The product classification was implemented using as much as possible automated methods, such as text mining and color detection techniques.

- ***Develop, test and validate a model of demand at the product level, and of the cannibalization and complementarity effects among articles***. This step consists in developing a model to quantify cannibalization effects on demand. The model focuses on cannibalization patterns based on similarities between products among one product category: trousers. For this step, we defined an attribute-based regression model using similarity factors among products, and fixed and time effects to isolate specific store and seasonal impacts.

- ***Develop, test and validate a demand forecast model at the product level.*** We used the most significant covariates determined in the step above to improve Zara's current forecasting methods. We compared the accuracy of the different models tested against the base model.

- ***Create a forecast model with estimated display information.*** We also improved the forecast using estimated display information, given that the actual display in a given week is not always known. Finally, we defined a simple method to estimate the display based on past information.

1.3 Insights and Summary of the Results

The results of the thesis are summarized below.

We estimated cannibalization and complementarity effects using a demand model with fixed and time effects across stores, for three different cities and for the Fall-Winter 2013 and the Spring-Summer 2014 seasons. Table 1 summarizes the main findings of the model. The sales of products with a similar fit cannibalize each other, whereas the sales of products with similar color or color lightness are complementary. The price of similar products also has an impact on the sales of a given product - an increase in the price of the former tends to increase the sales of the latter. All of these effects are consistent across seasons and stores; however, the measured effects are fairly small. This in turn can be rationalized by attenuation bias from measurement error, but also by other effects, such as product location in the store.

Coefficients	Effect
$\alpha_{subs, fit}$	-
$\alpha_{subs, color\ lightness}$	+
$\alpha_{subs, color}$	+
$\alpha_{subs, price}$	+

Table 1: Summary of the cannibalization or complementarity effects by attribute and price

In a second step, we used these results on cannibalization to improve the demand forecast. A model with display, availability and price effects results in substantial improvement in forecast accuracy, compared to the base model, which is a simplified version of Zara's existing forecast method, as one can see in Table 2. Adding cannibalization effects also improves slightly the accuracy of the forecast. However, it also requires a significant amount of data preparation. Hence, Zara should consider whether to add these effects to its forecasting model in light of this tradeoff. We generated the new forecast using both the actual product display information as well as a model that estimates the display given the information available to Zara's distribution team. The results with the estimated display are less accurate than the ones with the accurate display, but still present a significant improvement compared to the base method.

wMAPE reduction		
City	Actual Display	Estimated Display
Madrid	-18%	-11%
Barcelona	-10%	-7%
London	-43%	-26%

Table 2: Improvement in wMAPE by city with the new demand forecast model relative to the base model

These results can be implemented in Zara's processes, by improving the integration of the various databases used and the current product classification efforts. More precise results can also be obtained with more historical data for various stores.

1.4 Thesis Layout

The thesis is organized with the following structure:

- After the present overview of the thesis, we introduce background information on Zara in Chapter 2. We describe characteristics of fashion products, Zara's supply chain and the recent implementation of the RFID technology in Zara's stores.
- Chapter 3 presents Zara's existing forecasting method used for the short term replenishment and a literature review related to the study of cannibalization.
- Chapter 4, describes the work done to create additional attributes for the products. It also explains why we limit our study to one year of sales for the pants product category.
- The modeling of the product demand with cannibalization effects and the results are presented in Section 5.
- Chapters 6 and 7 focus on the demand forecast, by testing different models and comparing their accuracy with the one of the base model.

- Lastly, Chapter 8 summarizes the results. It also discusses the implementation of this project into Zara's processes and the next steps of the project.

2 Background

Zara has grown in forty years into the largest fashion retailer in the world, by offering a large choice of affordable clothes. Despite the short life cycle and high demand variability of fashion products, the company has been able to respond very quickly to customers' demand by developing a fast supply chain, from the design stage to the delivery in the stores. Moreover, the recent development of RFID technology has provided additional operational benefits, which gives valuable insight into the product interactions within a store.

2.1 Zara in the Fashion Industry

Zara is a world leading fashion brand, with close to 2,000 stores in about 90 different markets. It is the flagship company within the group Inditex, which also includes seven other brands (Massimo Duti, Zara Home, Berksha, Oysho, Stradivarius, Pull & Bear and Uterqüe). Altogether, the group represented more than €16B of sales in 2013 ([4]), leading in revenues ahead of the rest of the fashion retailing companies in the world.

Zara is very well known in the industry for having successfully implemented the fast fashion concept. It offers a large collection to cosmopolitan, fashion-conscious and budget-minded customers, with new products delivered every week in the stores, to appeal to its most fashion-forward audience. One of Zara's competitive strengths is its ability to respond very quickly to customers' demand by offering the latest trends, with a turnaround from design to store delivery of a couple of weeks. Zara is able to offer 18,000 distinct products each year, while traditional competitors propose 2,000 - 4,000 models ([5]). This strategy allows Zara to develop a customer base that returns to the store several times during the year to check out the latest fashion trends.

Inditex has a very strong presence in its home country, Spain, which represents 20% of its yearly sales. The United Kingdom, France and Italy also are large markets for the company in Europe, which totals 46% of Inditex's sales ([4]), while the company is expanding its Asia markets very fast, notably with its online sales.
Thus, Zara's success is based on following quickly the demand for fashion products. The next section presents the main characteristics of this demand, which is notoriously volatile.

2.2 Fashion Product Demand and Product Categories

Fashion garment products generally have a short product life cycle, high variability and very seasonal sales.

Products usually only last at most a season, i.e. five or six months - a trend that the fast fashion movement has accentuated. Zara defines two broad categories of products: the basic clothes, which have a longer life in the store and appeal to a large customer base; and the "fantasía" products. The latter have a trendier customer target, limited stock availability, and (as a result) may last in the store only for a couple of weeks. The two types of products are illustrated in Figure 1.

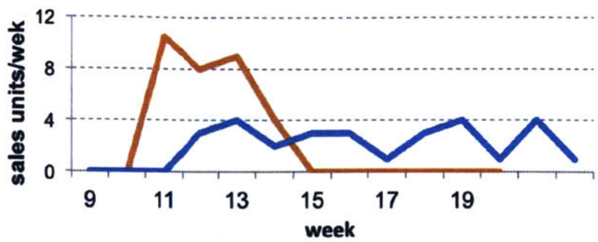

Figure 1: Examples of product life cycles in a store.
The blue line represents the demand for basic products, whereas the orange line represents an example of demand for "fantasía" products.

The weekly sales at the product level in each specific store also present a high variability. A product usually sells at a rate of a few units per week. Popular products sell a dozen units per week in certain stores, while many products are not purchased in a week - those articles tend to be replaced in Zara's stores with new references. Given this low average demand level, the demand sees a high variability, as a one unit difference represents a high percentage of the product weekly sales.

Moreover, the product demand can present a seasonal profile. The year is structured in two main seasons - the Fall-Winter season, which runs from July to December, and the Spring-Summer season, from January to June. The two seasons offering, and therefore the demand for particular products, are adapted to the weather at that time of the year. Special events during the year also impact the demand. For example, the store sales increase during Christmas in the US or Ramadan in the Middle East (a more detailed discussion on the demand seasonality is available in [6]). The end of the fashion season is also marked by significant markdowns in Europe. These periods have their own sales dynamics, and are the subject of a previous MIT thesis [7].

Therefore, the demand at the product level can show a high variability from week to week, while presenting seasonal patterns.

We now introduce Zara's current classification, which will be a useful reference for the rest of this thesis. Zara Women uses the following terminology[1]:

- At the higher level, Zara Women is organized in three *collections* - Woman, Basic and Trafaluc.
- The largest product category within a collection is the *product family*, such as shirts, skirts, pants. Zara defines about thirty different product families.
- The following level in the classification is the *product* level, also called MCC at Zara (in Spanish, Modelo/Calidad/Color).

[1]Because the project was in Zara Woman Distribution team, we looked specifically at the classification in the Women Department of Zara but similar classifications have been pursued in the other departments.

- Finally, the *SKU* is represented by a product in a specific size - this level is also called MCCt in Zara's jargon (t for "talla").

2.3 Zara's Supply Chain

To be able to respond quickly to demand, Zara has a very centralized and rather vertically integrated organization that focuses on regular customer feedback through the supply chain.

At the headquarters in La Coruña, hundred of designers, in close collaboration with the buyers, decide on the collection 6 to 8 months in advance, taking into account the latest fashion trends and the customer feedback received at the stores. The buyers procure the garment through three channels. For basic items, a long circuit sourcing is chosen - products are manufactured in Asia in large quantities for a low cost, with a lead time of about 12 weeks. With proximity sourcing, on the other hand, buyers procure garment manufactured in European and Mediterranean countries, such as Morocco, Turkey, Portugal or Spain. This channel allows for a shorter lead time (4 to 6 weeks) for a moderate manufacturing cost. Finally, in-house manufacturing is used for higher-priced "fantasía" products and allow Zara to react quickly to new trends. In-house manufacturing refers to a network of local suppliers in the region of La Coruña that have built long-term partnerships with Zara. Between these three channels, Zara can adapt its supply chain based on the profile of its product sales.

All the products, whether they are manufactured in Asia or locally, are transported to the distribution centers in Spain for the distribution to the stores. This centralized structure, though costly, also enables Zara to control its stock more closely and optimize the distribution to the different stores depending on the local demand. The garment is then transported, almost always by air, to the 2,000 stores in the world.

Lastly, all the stores in the world receive deliveries of new products or replenishment twice a week. Stores manage the in-store replenishment and product assortment processes. Every evening and every morning after the potential arrival of the trucks, they rearrange the store product assortment. They remove products for which the main sizes (S, M, L) are no longer available, follow the merchandizing guidelines by reproducing windows as shown in pictures of pilot stores in La Coruña, and place the new products on display. During the day, the sales associates regularly replenish the shop floor from the stockroom, as products are bought by customers. More generally, the stores play a crucial role in transmitting the customers' feedback back to the product managers at La Coruña, who communicate with the designers and buyers.

Thus, Zara is structured to be able to collect feedback from its stores and react quickly to the demand. Its organization and operational strategy allow the company to keep lower inventory than its competitors, which results in a lower percentage of its merchandise sold on markdowns (according to [1], 15% versus 50% for traditional fashion retailers). The RFID deployment in Zara's stores has also contributed to improving Zara's operations in the stores and to enhancing the agility of its supply chain.

2.4 Implementation of the RFID Technology

Recently, Zara has been deploying the RFID technology in its stores. The technology offers many benefits to the company's operations for the in-store operations:

- Easy and fast in-store inventory - at the moment of the delivery of new products in the store, the RFID technology allows for an almost instantaneous inventory, as opposed to the long manual process that consists in checking every box number against the listing and counting the number of articles received for each reference.

- Instant knowledge of units displayed on the shop floor for each product - this is particularly useful for the shop floor replenishment during the day. The store can immediately know how many units of each product it needs to bring from the stock room. In the past, the store checked throughout the day which products were sold within the last hour and used this information to replenish the shop floor.

- Instant knowledge of the inventory level - the regularly audits of the inventory in the store is largely facilitated by the RFID chips on the products. In addition, the stores also know the inventory levels of the other stores in the region. They can reorient a customer looking for a particular product to the nearest store where the product is still in stock, therefore improving the customer experience in the store.

- Finally, the RFID technology can in the long run provide invaluable information about the product display in the store. In particular, the information collected from the RFID database are essential for the study of the interactions among products in the stores or can be used to understand the impact of marketing or merchandising strategies within the stores.

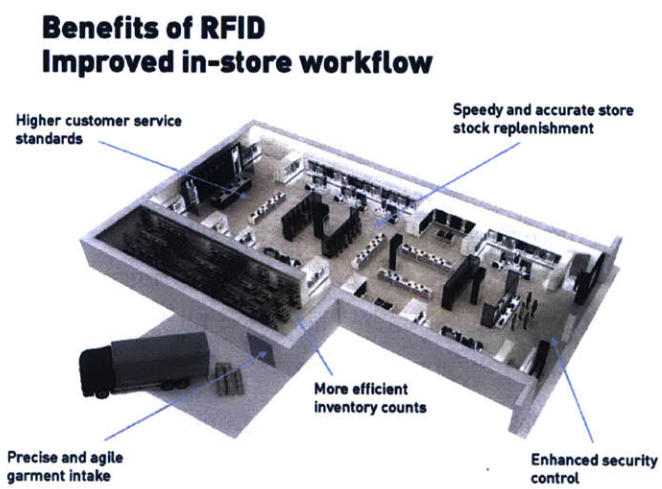

Figure 2: RFID technology benefits
(source [9])

[8] provides additional details on the benefits and business value of RFID deployment for fashion retailers and Figure 2 illustrates the benefits of RFID in the stores.

Almost all the clothes in the stores carry an RFID tag. Currently, Zara has rolled out the RFID technology in more than 700 of its stores and expects to complete the deployment in all its 2,000 stores by 2016.

Zara's strategy relies on a flexible, responsive supply chain to respond rapidly to its customers' needs. The RFID deployment in Zara's stores enhances the customer's value chain, but also provides access to new information that can be used to improve the understanding of the demand in the store. The present project focuses on that objective, by analyzing the cannibalization or complementarity effects among products in a store. We start in the following chapter by reviewing Zara's current forecasting method and the existing literature on cannibalization and describing our approach to the problem.

3 Description of the Approach

Before adopting an approach to study cannibalization and improve the demand forecast, one should understand the current demand forecast at Zara and the existing literature on cannibalization effects.The short-term replenishment from the distribution centers to the stores requires to predict the demand a week ahead. Zara's existing forecast method, detailed below, currently does not take into account any interaction among products displayed together in the stores. We then review in Chapter 3.2 the current literature on cannibalization modeling and finally describe our approach to understand cannibalization effects on the shop floor.

3.1 Existing Replenishment Method and Demand Forecast at Zara

Demand forecast is needed in various processes in Zara's operations: it is needed for financial planning; buyers prepare long term procurement plans using 6 to 8 month ahead demand forecast; for the distribution team, the forecast horizon is one week ahead. The work presented in this thesis focuses on the product demand forecast for the distribution team, and therefore only tackles one week ahead demand forecast.

The decision process for the replenishment from the distribution centers in Spain and for the first delivery of new products is centralized in La Coruña for all 2,000 stores. Zara sends products twice a week to the stores. At the beginning of the replenishment cycle, the demand of each product for the week ahead is estimated. The amount to be delivered to the store is the difference between the demand, adjusted for a given target inventory level, and the current stock at the store, as expressed in Equation 1 for each product k in each store s:

$$Replenishment\,Quantity_{k,s} \approx Daily\,Demand_{k,s} * target\,inventory\,level\,(days) - store\,inventory_s \tag{1}$$

The products are then sent to the stores in about 48 hours, before the beginning of a new replenishment cycle. Figure 3 illustrates the replenishment cycles from the distribution to the stores.

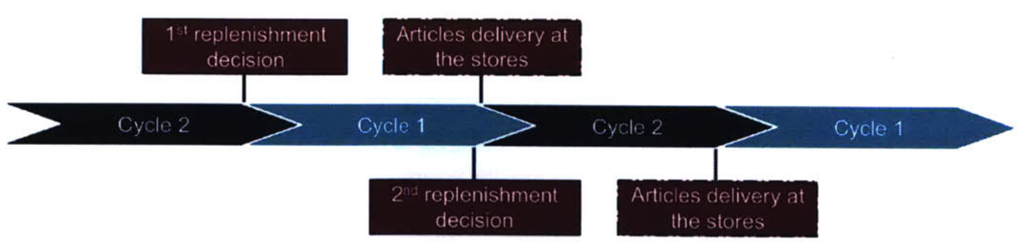

Figure 3: Representation of the two cycles for the decision making process of the distribution at Zara

For products already in the stores, the daily product demand is determined using the last two weeks of sales. For simplification purposes, Zara's existing forecasting method can be approximated as a two week rolling average of the demand. Adjustments can be made to account for stock-outs, the differentiated weights of the weekdays, and, when it comes to the demand at the size level (MCCt), for the profile of size sales at the store. The detailed calculation of the adjusted demand is explained in [10]. For the present thesis, we are interested in the demand at the product level. Therefore, we will consider that the current demand forecast for product k, in week w, in the store s, is expressed as follows:

$$\text{average daily sales}_{k,w,s} = \frac{\text{average daily sales}_{k,w-1,s} + \text{average daily sales}_{k,w-2,s}}{2}, \tag{2}$$

where the average daily sales of the previous weeks are adjusted for stockout. For example, for week $w-1$, the average daily sales are as follows:

$$\text{average daily sales}_{k,w-1,s} = \frac{\sum\limits_{day_{w-1}\, in\, week\, w-1} \text{sales}_{k,day_{w-1},s}}{\text{DPV}_{k,w-1,s}}, \tag{3}$$

where $\text{DPV}_{k,w-1,s}$ represents the sales adjusted by possible days of sales (DPV - días posibles de ventas). The possible days of sales exclude the days of the week when the product was out of stock or when the store was closed.

Finally, a different process is in place for the demand forecast of new products. This was the object of a previous MIT project with Zara ([10]), which the reader can refer to for further details. The forecast of new products relies on the use of products similar to the new product going to the store. The sales of these comparable products serve as the basis of the new product demand forecast for the first weeks.

Thus, the forecast of the product demand one week ahead is crucial to Zara to minimize stock-outs and overstocks. To estimate the product demand, Zara uses a method based on two-week rolling average of the product sales. With this method, it assumes that products are independent of each other in the store. However, this assumption is debatable. Customers have a limited budget going into the stores and have therefore to choose among products to stay within their budget. One can also think that if many similar products are displayed together, it may negatively or positively impact the sales of each of these products individually. Thus, products interact with each other in the stores and this may impact of the sales. The study of the cannibalization and complementarity effects among products is one of the objectives of the present work.

3.2 Literature Review

The study of cannibalization is the object of multiple fields of research in operations research (for product assortment problems) and in economics. We first take a look at how cannibalization is defined, in particular in relation to substitution; we then present various commonly used models to measure cannibalization.

3.2.1 Definition of cannibalization and substitution

Cannibalization is defined in different ways in the literature and is closely related to substitution.

Cannibalization has been defined as "the process by which a new product gains sales by diverting the sales from existing products" ([11, 12]). This definition, fairly consistent across the marketing literature ([13, 14]), is often employed for new product introduction. However, one can also be interested in studying the decrease in sales of existing products due to the mere presence of other products. This is closer to the definition of substitution, term that is more often used in the economics literature. Substitution is more broadly defined as the extent to which a customer chooses a product to replace another product, whether it is because the latter was not available or because the customers had a personal preference. Consequently, the sales of the existing or unavailable product decrease. Hence, the definitions of cannibalization and substitution are closely related and in the rest of this thesis, we will indiscriminately use the two terms.

The literature has defined a number of distinct substitution effects. In the context of assortment planning for retailers, A. Gürhan Kök ([2]) defines three distinct patterns of substitution. The first type of substitution, called *stock-out based substitution*, is based on the assumption that a customer will replace an unavailable good that he/she was planning to buy with a similar product available in the store. Another substitution, based on the assortment choice made by the retailer, sees the customer choose among the products displayed in the store, to replace an article seen elsewhere. These two patterns are often encountered for food products and the second one is of less interest for our study, given the scope of the thesis. Finally, the third type of substitution listed in [2] can be more often encountered for fashion retailers. It is a *substitution based on consumer preferences.* The customer looks at the available products in the store. He/she considers an article based on the utility he/she gains from its purchase and selects the article which maximizes his/her utility. Because the latter is very common in the fashion industry, we will focus on this particular pattern of substitution.

Within consumer preference-based substitution, [15] identifies further types of cannibalization. The *multi-product pack cannibalization* and *combo-product cannibalization* both relate to different products which, because they are packaged together impact negatively the sales of products sold individually. In the context of our project at Zara, we are more interested in the two other types of cannibalization defined in [15]. The first is *inter-product substitution*, which deals with patterns of substitution between articles belonging to different product categories. For example, it would be the purchase of a pair of pants instead of a skirt. The *intra-product cannibalization* is related to competing products within the same product category, for example two different shirts.

Substitution can occur at various levels. It happens between competitors - for example, the opening of a new store can cannibalize the sales of competitors in the neighborhood. For a same brand, the substitution can occur between the various stores or even between difference channels of sales. For instance, the fast growth in Zara's internet sales may impact the sales in the stores[2].

[2]The brick and mortar and e-retail sales can present both substitution and complementary effects. Customers may prefer to buy online instead of in the stores, but the integration of the brick and click businesses can also complement

However, in the scope of this work, we will restrict the substitution to patterns occurring within the same store of Zara.

We still need to determine which articles within a store are most probable to be substituted. For the case of Zara, given the large number of articles on the shop floor at the same time (500 to 2,000 articles, depending in a given store), and the very limited time an article is displayed, it is virtually impossible to measure *all* the substitution patterns of a given article with the rest of the articles displayed on the shop floor at the same time. S. R. Srinivasan et al. in [15, 16] discuss that the cannibalization happens most often based on product attributes. These attributes include brand, product family, and individual product attributes (for example color, pattern, fabric or price for garment products). Therefore, as explained in [17], two products from a different product family with different functions, such as a scarf and a skirt, are seen as less likely to substitute each other. [18] is an example of paper that adopts the same attribute-based approach to model its demand.

Cannibalization and substitution are often very similarly defined in the literature, and we will used both terms in the rest of the thesis without any distinction. A large number of substitution patterns exist within a store; however, the present work will focus on patterns based on consumer preferences between products with similar features. Given this focus, we present models of substitution in the next section.

3.2.2 Cannibalization Models

Different approaches can be adopted to model product demand with cannibalization effects. We will detail the most widely used model, the Multinomial Logit Model, then we will give a brief overview of other approaches that can be taken.

Multinomial Logit Model A very popular model of choice is the multinomial logistic (or logit) model, as described in [19, 20]. The multinomial logistic (MNL) model relies on the assumption that customers maximize their utility by choosing among a finite choice set. For a given alternative j, the utility of the customer n can be written linearly as follows:

$$u_{nj} = \beta x_{nj} + k_j + \epsilon_{nj}, \tag{4}$$

where x_{nj} is a vector of variables that describe alternative j for individual n, β are the vector coefficients for these variables, k_j is a constant that is specific to alternative j and ϵ_{nj} is the unobserved portion of the utility. Assuming that each ϵ_{nj} is independently and identically distributed according to an extreme value distribution, the logit choice probability that the individual n chooses alternative j is:

$$P_{nj} = \frac{e^{\beta x_{nj} + k_j}}{\sum\limits_i e^{\beta x_{ni} + k_i}} \tag{5}$$

each other, thanks to better brand awareness and higher level of service for the customers.

As discussed in [20], the MNL model has limitations. Notably, it only allows proportional substitutions across alternatives. The model assumes that the Independence from Irrelevant Alternatives (IIA) property is valid. This entails that when a new alternative appears in the choice set, the rest of alternatives probability decreases by the same percentage. This model is appropriate for many situations, but inadequate for others such as the famous example of the *red bus-blue bus problem.* Let us imagine that an individual has a choice of either taking a car or a blue bus and that a new choice, the red bus, is available to the individual. Under the IIA assumption, the probabilities that the individual chooses the car and the blue bus decrease by the same percentage. However, it seems logical that the probability for the blue bus decreases more, given the similarity of the blue bus and red bus alternatives. In this situation, the MNL overestimates the proportion of individuals that take the bus and underestimate the car option.

A second important restriction of this model is that it provides *shares* of the market for the various products, without indicating if the overall market has increased with the introduction of the new products. In other words, we do not know if the sales of the new product have cannibalized the sales of the existing products or has increased the size of the market. A final limitation of the MNL model is that it is usually used for a fairly small set of choices and therefore it would be computational intensive to use the model for the thousand of products displayed in a Zara's store.

Despite these limitations, the MNL model has been extensively used in the literature to study substitution patterns. To overcome the restrictions on the substitution patterns made by the MNL, other generalized linear models can be used. The nested multinomial model is very common, allowing to provide a hierarchy in the set of choices - for example the red bus and blue bus on the one hand, and the car on the other hand. [20] describes other models for discrete choice models similar to the MNL, and gives many examples.

Therefore, for the present work, MNL models presented some issues. First, with a large number of displayed products (at least 100 of products displayed at the same time), the estimation of the parameters for the MNL models proceeds to be be time-consuming. In addition, MNL model outputs provide percentages of sales shares for the various products, which may not always be appropriate. With only the prediction of percentages, it is difficult to know if a new product has cannibalized the sales of other products or has increased the volume of sales without impacting the sales of the other products.

Combined Clustering and Regression Methods Another approach to model cannibalization is to use data mining methods to group products together, and then study the cannibalization within the product groupings.

Several classifications methods can be used to create clusters of products. Two methods easy to implement are the k-means and k-medoïds clustering. The first step of both methods is to measure the inter-product distances, which is a measure of how similar, in terms of attributes, two products are. Different weights can be attributed to the different attributes. The next step is to try to find, for a given number of clusters k, centers that are representative of the clusters. Products are then associated with the cluster whose center they are the nearest to. Other methods are available, such as hierarchical clustering, decision regression trees or support vector machines. The reader can refer to [21] to get an overview of the various techniques.

Once the products are grouped together, analysis using regression methods can be used to measure cannibalization. Iterations may be necessary to find the best groups of products, depending on how much weights is given to each attribute. For this purpose, [22] adopts a technique that seems particularly appropriate for the current problem, as it enables to cluster and run the regression simultaneously.
However, the method based on clustering presents the drawback of being difficult to implement dynamically, as clusters have to be recalculated as new products are introduced in the stores.

Attribute-based Regression Models The last approach discussed uses the classical linear regression method. [18] uses a log-log model to estimate the product demand and uses similarity factors to measure cannibalization among products. This method has the benefit of using few parameters. In the study of [18], the products considered are detergents sold in supermarkets, which is an easier product to study. Detergents generally present a smoother demand and have a longer shelf life than fashion products. In addition, the number of displayed products (only 31) is limited. However, this method can be adapted to fit the context of this thesis.

Hence, the various methods described above presents various alternatives to model substitution. MNL is a very popular model but may not be adapted to a very large choice set. Methods combining clustering and regression are also used in the literature, but may be difficult to use in the case of short product life cycles. Finally, regression models using similarity factors have also been employed but need to be adapted to fashion products. The latter approach is the basis of our model, which is presented in the following sections.

3.3 Proposed Approach

We describe below the approach taken to study cannibalization and to improve the demand forecast at the product level.

The literature review showed that for all the methods to study the cannibalization among products, it is essential to obtain attributes for each product. This becomes even more a necessity when the number of products studied at any time reaches one hundred or more. We reviewed the current product classification at Zara in detail and we found that additional work was required to be able to distinguish the products based on their attributes. In fact, limited information was available on the specific features for each product. Therefore, the first step consisted in collecting data to create the appropriate product classification. This was done using text mining methods and color detection techniques on the pictures of the products. We limited the study here on the trousers product category, but this approach can be adapted to other product categories. The product classification and results are presented in Chapter 4.

Once an adequate product classification was determined, we focused on the study of cannibalization or complementarity effects. We chose to adopt an approach similar to the attribute-based regression models described in Section 3.2.2. The differences with the model in [18] reside in how the similarity factors was calculated and how we defined product clusters to account for the products that have the exact same set of features. We made the study on various stores in different cities to benefit from a high statistical power coming from a large sample. To identify the cannibalization or complementarity effects and isolate the store-, feature- or time-specific effects, we chose to introduce fixed and time effects. Using this model, we estimated the attribute-based and price-based cannibalization effects. The results are discussed in Chapter 5.

After the study of the cannibalization, we created a demand forecast model (see Chapter 6). We had to adapt the model to study the cannibalization effects, as time and fixed effects were inadequate for forecasting. Therefore, to integrate information about the products, we included sales lags of order 1 and 2. We specified and estimated models with incremental improvements. We then compared and discussed their forecasting accuracy.

Another step necessary to implement the findings of this thesis into Zara's processes was to work with imperfect display information for the products. The reasons to have only partial information on the display from the perspective of Zara's distribution department are explained in Section 7.1. We first estimated whether a product is displayed in the week w using the display information of the week $w-1$ and the information about the planned delivery of new products for the week w. This approach proved to be sufficient to obtain an estimate of the product display one week ahead. We defined again forecast models using the estimated display information. We also compared their forecast accuracy against the existing forecast and the forecasts using the actual product display.

Finally, we summarized the results in the last chapter and discussed potential improvements or directions to improve the accuracy of the demand forecast. We also identified the steps necessary to implement the work on this thesis into Zara's tools.

4 Data Collection and Product Classification

Before being able to study cannibalization effects, it is necessary to create a classification of the products, based on their attributes. The following chapter discusses the current classification and the work done to obtain additional information on the product attributes.

4.1 Data Description and Selection

Part of the initial work was to select the appropriate and available dataset for the study of cannibalization.

Inditex has an extensive data collection system which provides information on each article at the SKU level (each reference, at the size level) in every store. This database provides daily sales, product family information, and an estimate of the available stock, for each store and each product broken down by size. The size of the daily product-level sales data for one store for a period of one year can amount to 3 million rows. If we were to work with the sales data of all Zara's stores at the same granularity, the data size would exceed 2 billion rows.

The second source of information available was the one collected from the RFID system. This technology is in the process of being deployed worldwide in all Inditex's stores - at the time of the project, only data for Spain, Portugal and the United Kingdom were available for a period of a year or more. As a result, our analysis was restricted to these three countries. Still, the recorded RFID information results in database of the same size as the sales database described above. The RFID database provides exact information on whether an article is on the shop floor or in the stockroom for each day of sales, the desired inventory position on the shop floor and the stock room inventory. Given the complexity of comparing multiple stores, we also selected stores that did not undergo a major renovation in the past two years, which restricted the eligible stores to a few hundred.

Finally, as we studied cannibalization effects among products having similar features, it was necessary to obtain information related to the product attributes, such as color, fit, existing patterns, neck, fabric, etc. Some of these attributes, such as the fabric, were readily available as a field in the database. However, features like fit or existing patterns on the garment were information that had to be extracted from free description fields of Zara's database. The current classification of the colors of the products were particularly complex and had to be reprocessed to be consistently used in our analysis. The collection of relevant attributes for the articles was therefore a time-consuming task. Given the time constraint of the internship, we limited our study to one product family—we decided to focus on the trousers product category. This product family still provides a wide variety of choices to the customers, with more than 800 distinct references offered throughout one year. Depending on the number of products a store receives, the size of the final dataset for the trousers product category is 20,000 to 35,000 rows per store.

The following section describes the methods used to extract the main attributes from the database.

4.2 Classification Methods Used

4.2.1 Text Mining

We employed text mining methods to obtain attributes that are not yet collected systematically into Zara's database. We used the following methodology:

- Identify fields in the database containing valuable information for the classification.
- Based on the recurrences of the words, identify attributes used by the buyers to describe a product. For the pants product category, the fit, the existence of patterns, the color and the color lightness were attributes very commonly described in the analyzed fields.
- Create lexicons for each attribute. Words found in the various fields were in different languages (mostly Spanish and English) or had typos. Algorithms, available in packages in R, can recognize typos and can be used to process the data faster.
- We added manually missing information for some of the products to complete the classification of the collection for the season Fall-Winter 2013 and Spring Summer 2014.

The benefits of such an approach is that if the description provided by the buyers is accurate enough and defined with a keyword approach, it will be easier to detect trends among the collections or patterns of sales. This method was used to determine the fits and patterns for each product in the trousers product family.

4.2.2 Color Detection Algorithm

Zara already had an existing classification for the colors but this classification was mostly qualitative and difficult to simplify.

To have a more systematic approach, we created a code to analyze the pictures of the products and detect the color of the garment. Two methods were used. The first attempted to detect the garment in the picture and then gave the average color of the area where the garment was detected. The second method used the fact that all pictures had a similar layout and that we could find a fixed area for all the pictures where the garment was positioned. We then calculates the average color of that area. The results of both methods are shown in Figure 4. The second method gave better results given the generic layout of the product pictures. The first method had a disadvantage when it came to detect the contours of light-colored products.

Once the new system of color was obtained, we classified it in three different main colors - grey (which groups black, grey and white), blue, and colored (all other colors). To additionally simplify the classification, we added patterned products as "colors". An additional feature classification was also created to represent the lightness of the color (dark, medium and light).

Figure 4: Illustration of the Color Detection Methods
For each image, the boxes at the bottom-right corner represent the color detected by the algorithm. The bottom box is the color obtained when trying to detect the location of the product in the image. The detection of the product contours in the image is not optimal, especially for light-colored pants. The top box shows the color obtained when assuming that the product can always be found in the same region of the image. Given the generic layout of the pictures, this method produced better results.

4.3 Results of the Classification

Figure 5 below shows the distribution of classified products across the above-defined categories.

For the colors, the more represented groups are blue and grey, whereas there is a smaller number of patterned products and of products of colors other than grey and blue. For the fit, the collection includes a very large number of slim models, which is consistent with the fashion trend of the last few years. Finally, pants in the collection tend to be medium or dark-colored.

The sales of each trend can also be observed in Figure 6 for some stores in Madrid, from September 2013 to May 2014.

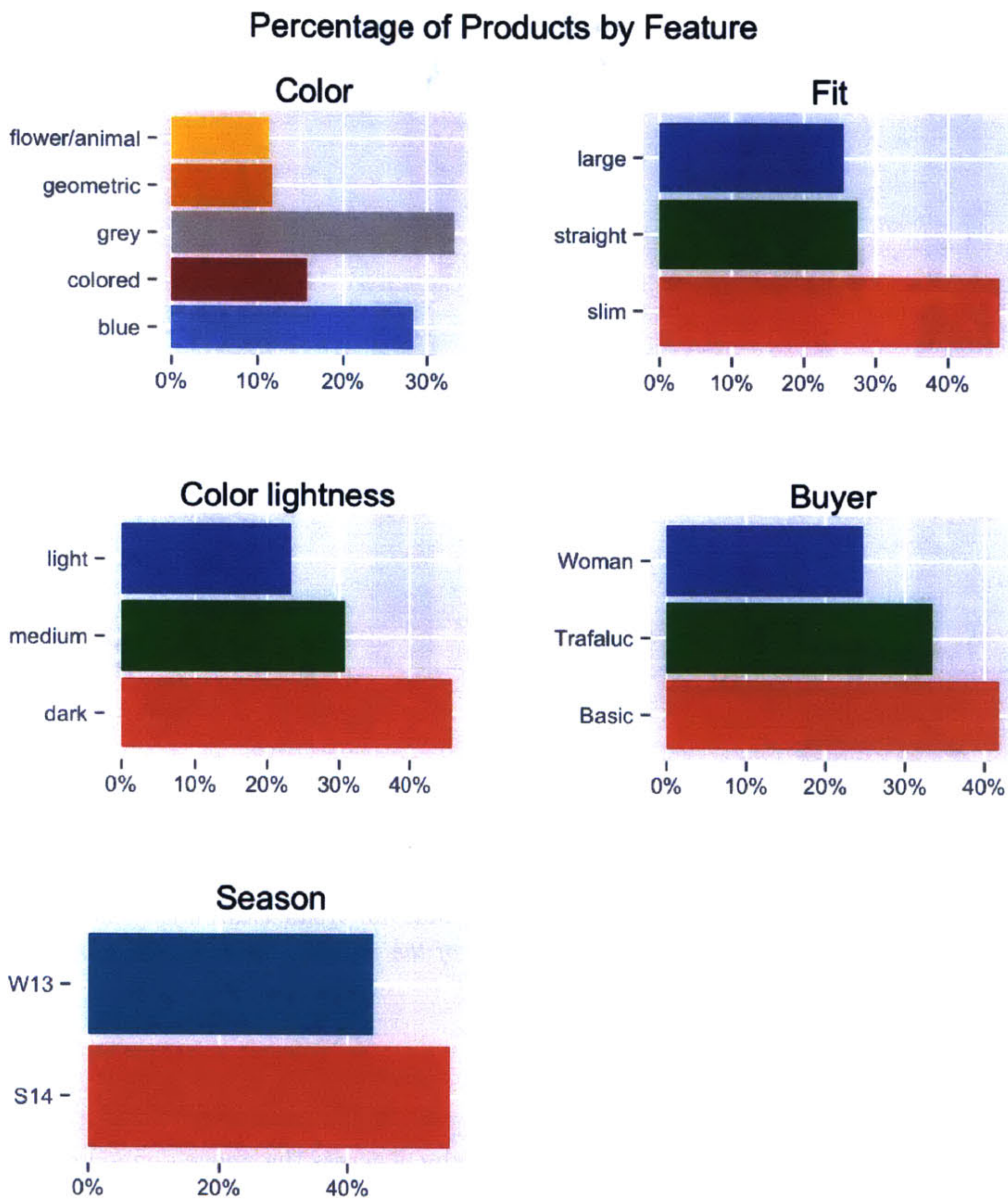

Figure 5: Classification of the products by characteristics (color, fit, color lightness, buyer and season) The buyer classification represents Zara's three product lines.

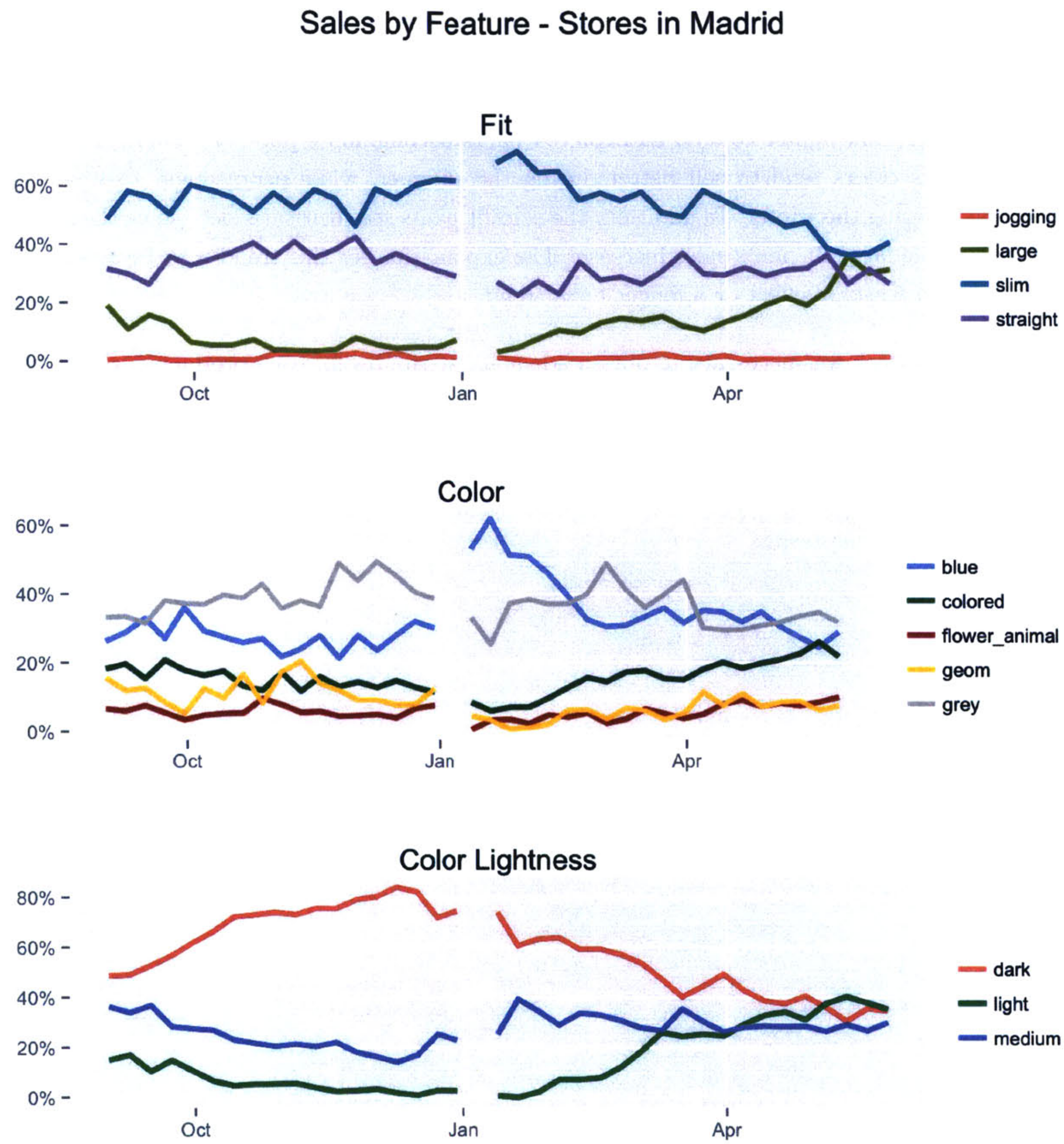

Figure 6: Trends over the year of the products presenting various features for some stores in Madrid (fit, color, color lightness).
The lines represent the percentage of trousers by feature over time.

We present below the findings for the stores in Madrid based on the observations of Figures 5 and 6. However, the same trends were observed for the other stores in the data samples in London and Barcelona.

As expected, dark-colored, blue, grey and slim pants represent the majority of the sales. However, two trends are interesting to note. The sales of the dark-colored, grey and blue pants decrease closer to summer, and lighter-colored and colored pants become more popular. This phenomenon is well-known as colors tend to sell better during the summer, while people tend to wear more neutral colors during the winter. In addition, the slim fit pants see their sales decline as the summer approaches, while large fit pants sales increase. The explanation for this trend may be less obvious - this could be a seasonal effect or a longer term trend.

Thus, additional work was necessary to obtain adequate attributes for each product in the trousers product category. The various methods used in this chapter can be adapted to the other product families. With this classification, we can therefore study the cannibalization or complementarity effects among products in a store.

5 Modeling of the Cannibalization and Complementarity of Products within a Store

To estimate the cannibalization and complementarity effects in a store, we attempted various approaches, described in Section 5.1. We eventually found a simple approach that could be easily adapted for the all product categories, and for the needs of the distribution department at Zara. The Sections 5.2 to 5.4 present the final model adopted, its results and a discussion about its limitations.

5.1 Discussion on the Model Selection

We discuss below the various models that were considered for the study of cannibalization.

Initially, considering the large number of articles present at the same time on the shop floor, we decided to cluster products together to reduce the number of interactions under consideration. As explained in Section 3.2.2, we calculated an inter-product distance matrix, indicative of the degree of similarity between two products. Using the k-medoïd method, we created k clusters. Each cluster was represented by its medoïd and the distances between clusters were given by the distances between the medoïds of the clusters. For a given cluster, we ranked the other clusters from the "closest" cluster to the "furthest" one. We then could run a linear regression measuring the effects of the closest and furthest clusters on the sales of a given cluster. The initial choice of the adequate weights for each feature of the classification and the number of clusters can turn into an optimization algorithm when considering that we can choose the weights to obtain the most significant cannibalization coefficients in the linear regression. However, we tried different optimization methods for this problem but results were not consistent across different samples and therefore we did not pursue this approach further.

We also considered a multinomial logistic (MNL) model to analyze the cannibalization effects. As described in Section 3.2.2, generalized linear models are often used to model consumers discrete choice. It was also necessary to cluster products together, given that the computational time to estimate MNL models increases linearly with the number of clusters. We followed the same clustering approach as the one described in the previous paragraph. The biggest drawback of this method was still the computational challenge of treating a large dataset with a wide choice selection. In addition, we encountered the same difficulty to find the appropriate weights for the calculation of the distance between products as in the previous paragraph. For these reasons, to which we can add the limited patterns of substitution that can be represented with MNL models, we looked into other directions to study cannibalization.

Finally, another option considered was the use of dynamic linear models. This approach, and the related Kalman filter methodologies, is described in details in [23] and has already been adopted in demand forecasting, as shown in [24]. However, the complexity to compute parameters with a large datasets was again a limit to the use of such models.

Consequently, we attempted to use various models to estimate cannibalization effects in Zara's stores. However, the choice of clusters based on distances calculated with arbitrary weights for each attribute, and the computational complexity of the models have been important drawbacks in the approaches described above. The final model adopted, described in the next sections, has the benefit of simplicity and provided consistent results across the various sample datasets.

5.2 Specification of the Product Demand Model with Cannibalization and Complementarity effects

The model adopted for this study introduces similarity factors among products in a store, similarly to [16] and [18]. These terms are the attribute-based cannibalization effects that we aim at estimating. Price-based cannibalization is also of interest in our model. We also define fixed and time effects in our model to neutralize the heterogeneity of the demand across the stores and due to product attributes. We first introduce the notation used through the chapter in 5.2.1 then present the model in 5.2.2.

5.2.1 Description of the Notation

We provide below details on the notation used in the subsequent sections:

- k, k', k'' are indices for products in the store
- c is the index used for clusters of product
- w is the week studied
- s is the store studied
- f is the feature for which we are trying to detect cannibalization (color, color lightness and fit)
- t is the size of the product
- $A_{f,k'=k} = 1$ when the products k and k' have the same feature f and $A_{f,k'=k} = 0$ otherwise
- $N_{w,s}$ is the number of products of the category studied (here, trousers category) displayed during week w in store s
- $display_{k,w,s}$ is the weekly average shop floor display factor of the product k in week w in store s. For example, if the product was displayed on the shop floor all days of the week, $display_{k,w,s} = 1$; if it was not displayed only 3 days out of the 7 days of the week, $display_{k,w,s} = \frac{3}{7}$
- $stock_{k,t,w,s}$ is the stock (in number of units) of product k in size t in week w, and in store s
- $sales_{ac_{k,t,w,s}}$ is the accumulated sales since the first day in the store of product k in size t at the time of week w in store s

- $sales_{ac_{k,t,w,s}}$ is the accumulated sales since the first day in the store of product k for all sizes at the time of week w in store s

- $days\ opened_{w,s}$ is the number of days the stores are opened in the week. Most stores are open 6 days a week, whereas in some touristic places, some stores are open even on Sundays.

We also created additional parameters, to measure similarities among products in the store and product availability.

First, we defined clusters, with the index c, made of products that have the same set of features. For example, a cluster includes only products with a "straight" fit, of color "blue" and of "dark" luminosity. Any light-colored product will be in a different cluster. Appendix 8.3 shows the number of products in each cluster for the seasons Fall-Winter 2013 and Spring-Summer 2014 while Figure 7 illustrates the various levels of product groupings used in this thesis.

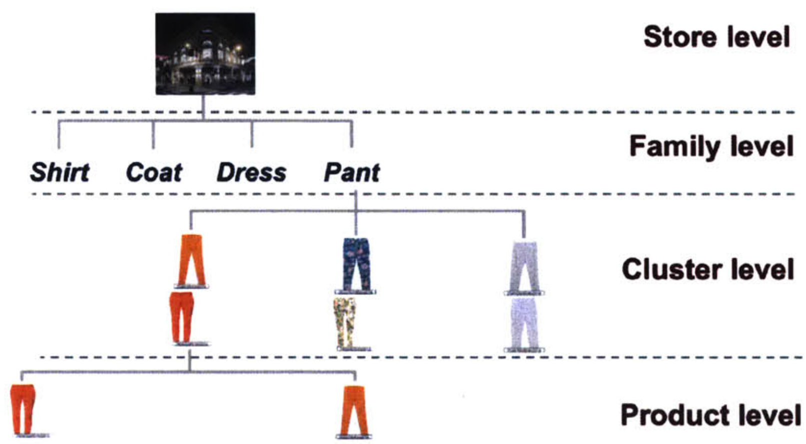

Figure 7: Hierarchy of the various product groupings in a store for the study of the product demand

As discussed in the literature review section, we expect that the cannibalization effects are more important among similar products than among products that present a lot of differences. Therefore, it was necessary for the study to define a similarity factor that translates how similar one product is with the rest of the store assortment. We define a similarity factor $SIM_{f,c,w,s}$ for cluster c in the week w in the store s for the attribute f as follows:

$$SIM_{f,c,w,s} = \left(\sum_{c' \neq c} display_{c',w,s} A_{f,c'=c} \right) * \left(1 - \frac{1}{N_{w,s}} \sum_{c''} display_{c'',w,s} A_{f,c''=c} \right) \tag{6}$$

The similarity factor defined in (6) follows the approach of [16] and [18]. It differs from [18] by the introduction of the display factor to reflect that the display is now averaged over the week and by the fact that the similarity factor is here defined for a cluster of products. The first term of the similarity factor reflects the fact that a cluster will have a high similarity factor if many products in other clusters displayed in the store at the same time have the same attribute f as a given cluster c. The second term of the right-hand side of the equation takes into account that a rare attribute should have a higher weight than a very common attribute among articles.

The availability factor of the product $availability_{k,w+1,s}$ is defined in the equation below as a global indicator of the stock per size of the product k at the time of week $w+1$ in the store s:

$$availability_{k,w+1,s} = \sum_t \frac{sales_{ac_{k,t,w,s}}}{sales_{ac_{k,w,s}}} * \min\left(1, \frac{stock_{k,t,w,s}}{sales_{day_{k,t,w,s}} * days\ opened_{w+1,s}}\right) \tag{7}$$

Each size is given a weight equal to the percentage of sales of that specific size in the total sales of the product k. If the stock of the size t is sufficient for the following week, the availability of that size is equal to 1. The overall availability of the product k at the time of week $w+1$ in the store s is the weighted average of the availability of the product in each size.

5.2.2 Demand Model

To study the cannibalization, we defined a linear fixed-effect and time-effect model. We grouped together into clusters the products with the same set of attributes (fit, color, luminosity). The idea of the model is that the cannibalization occurring between clusters is due to the similarity of attributes between these clusters, whereas the cannibalization within a cluster, where products have the same set attributes, is due to the difference in price. We modeled the sales in week w and in store s for product k as:

$$\begin{aligned}
\log(\text{sales}_{k,w,s}+1) = & \underbrace{\sum_{f,w} \lambda_{f,w}\delta_{f,w} + \sum_{s,w} \lambda_{s,w}\delta_{s,w}}_{\text{time effects}} + \underbrace{\sum_{f,s} \lambda_{f,s}\delta_{f,s}}_{\text{fixed effects}} \\
& + \underbrace{\sum_{\text{f}} \alpha_{subs,f} \text{SIM}_{f,k,w,s}}_{\text{similarity-based cannibalization}} + \underbrace{\alpha_{subs,price} \log(\text{Price}_{other\ products\ in\ cluster}+1)}_{\text{price-based cannibalization}} \\
& + \underbrace{\alpha_{display\ product} \log(\text{display}_{k,w,s}+1)}_{\text{display effect}} + \underbrace{\alpha_{availability\ product} \log(\text{availability}_{k,w,s}+1)}_{\text{availability effect}} \\
& + \underbrace{\alpha_{self-price\ product} \log(\text{price}_k+1)}_{\text{self-price effect}} + \epsilon_{k,w,s}
\end{aligned} \tag{8}$$

The right hand side of the equation includes the following terms:

- time effects and fixed effects: these terms were introduced to control for changes over time of the popularity of some features uniformly across all stores, difference in traffic in the stores for each week and finally to reflect different tastes for the different features.
- similarity-based cannibalization: these terms, one for each attribute, are the cannibalization effects due to the presence of products with the same attributes.
- price-based cannibalization: here, we looked at the effect of the average price of the other displayed products with the same attributes as product k.
- display effect: we controlled for the time the article is displayed in the store.
- availability effect: the impact of how available the product is in the store was also measured. The availability is measured with a number from 0 to 1 based on the stock available in the different sizes for a given product.
- self-price effect: this term measured how the price of a product impacts its sales - it is the price elasticity of the product.

5.3 Results

The model defined in the previous section was tested on several selected data samples.

First, we expected the cannibalization effects to be rather small given the significant noise in the sales data. To find some significant effects, we needed to have a large data sample to be able to have some statistical power for our model. Therefore, we chose a sample that regrouped several stores. This explains the indexing of the various parameters with the store index s in the previous section. We also chose stores that are located in the same city. This was intended to neutralize the weather effects. Finally, after discussion with the team, we chose stores that are located in a commercial center. The street stores and the stores located in a commercial center have different patterns of traffic, which should be accounted for. In addition, the layout across commercial center stores is more uniform than the layouts for street stores. To attempt to control for the difference in store layout, we selected only commercial center Zara stores. Therefore, for our model, we selected three groups of stores, located respectively in Madrid, Barcelona and London. Moreover, the model was run separately for the Fall-Winter 2013 season and the Spring-Summer 2014 season, which resulted in six distinct data samples.

Table 3 shows the parameters estimation of our model on the various samples. The model was estimated by ordinary least squares.

We present below the main results of the parameters estimation:

Coefficients	Value (Significance)					
	Madrid		Barcelona		London	
	Winter	Summer	Winter	Summer	Winter	Summer
$\alpha_{subs,\ fit}$	-0.011 (***)	()	-0.016 (***)	-0.008 (***)	-0.006 (***)	0.005 (***)
$\alpha_{subs,\ luminosity}$	0.010 (***)	0.008 (***)	0.012 (***)	0.012 (***)	0.008 (***)	0.014 (***)
$\alpha_{subs,\ color}$	0.001 (*)	-0.001 (**)	0.003 (***)	0.002 (***)	0.002 (***)	()
$\alpha_{subs,\ price}$	0.118 (***)	0.060 (***)	0.070 (***)	0.066 (***)	0.055 (***)	0.019 (**)
$\alpha_{display\ product}$	1.686 (***)	1.809 (***)	1.680 (***)	1.768 (***)	1.561 (***)	2.038 (***)
$\alpha_{availability\ product}$	0.345 (***)	0.359 (***)	0.310 (***)	0.355 (***)	0.317 (***)	0.192 (***)
$\alpha_{self\text{-}price\ product}$	-0.146 (***)	-0.231 (***)	-0.101 (***)	-0.255 (***)	-0.086 (***)	-0.293 (***)

Table 3: Value of the Coefficients obtained for the Cannibalization Model
*Signif. codes: 0 ≤ '***' < 0.001 ≤ '**' < 0.01 ≤ '*' < 0.05 ≤ '.' < 0.1 ≤ ' ' < 1*

- The substitution effect due to the similarity in the fit is generally negative, except for one sample data; it is therefore a cannibalization effect. The more products with the same fits are available on the shop floor, the lower the sales of a given product are.

- In contrast, the substitution effects due to the similarity in color luminosity and color tones are generally positive. Therefore, those effects are complementary and the display of more products with the same color or luminosity tends to increase the sales of the product considered.

- The coefficient related to the price-based substitution is positive, which seems logical. If the average price of the other products in a given cluster increases, the sales of the product considered tend to increase.

- The display and availability effects are positive and very significant for all the data sample.

- Finally, the self-price elasticity is negative, which is again consistent with our expectations and with most economic models. In addition, the absolute value of the elasticity is lower than 1, which suggests that an *increase* in the price of the product would generate more revenues despite the slight decrease in sales. Note, however, that this estimate reflects the short-term elasticity only. We may expect larger values of this elasticity in the long run. Therefore, an increase in price may increase revenue in the short-run, but might have longer-term negative consequences and negatively impact Zara's revenues.

Thus, our model of cannibalization generally provided statistically significant cannibalization or complementarity coefficients for our different samples. We also looked at the results of the Anova (type II) test to analyze the contribution of the various parameters to the variance. The results obtained for the Sum of Squares is given in Table 4.

Anova Test (Type II) - Sum of Squares (Significance)						
Coefficients	Madrid		Barcelona		London	
	Winter	Summer	Winter	Summer	Winter	Summer
$\alpha_{subs, fit}$	19.6 (***)	0.0 ( )	17.8 (***)	2.8 (***)	3.0 (***)	0.9 (***)
$\alpha_{subs, luminosity}$	19.4 (***)	3.8 (***)	10.4 (***)	6.4 (***)	10.2 (***)	6.9 (***)
$\alpha_{subs, color}$	0.3 (*)	0.4 (**)	1.7 (***)	1.2 (***)	0.6 (***)	0.2 ( )
$\alpha_{subs, price}$	13.3 (***)	5.7 (***)	2.5 (***)	3.3 (***)	2.2 (***)	0.4 (**)
$\alpha_{display\ product}$	110.2 (***)	132.8 (***)	41.6 (***)	60.8 (***)	69.9 (***)	154.0 (***)
$\alpha_{availability\ product}$	23.8 (***)	20.6 (***)	9.8 (***)	9.4 (***)	13.2 (***)	3.5 (***)
$\alpha_{self\text{-}price\ product}$	7.2 (***)	11.9 (***)	1.7 (***)	8.8 (***)	2.3 (***)	17.7 (***)
$\lambda_{s,w}$	4.1 ( )	7.2 (***)	2.2 ( )	3.6 (**)	6.0 (***)	8.9 (***)
$\lambda_{w, fit\ feature}$	4.7 (***)	5.7 (***)	3.4 (**)	4.7 (***)	4.0 (***)	5.1 (***)
$\lambda_{w, luminosity\ feature}$	7.6 (***)	4.9 (***)	10.3 (***)	1.2 ( )	6.0 (***)	2.3 (*)
$\lambda_{w, color\ feature}$	6.7 (***)	6.4 (***)	2.4 ( )	3.4 ( )	2.8 ( )	5.3 (***)
$\lambda_{s, fit\ feature}$	3.6 (***)	3.8 (***)	3.9 (***)	3.5 (***)	9.0 (***)	2.5 (***)
$\lambda_{s, luminosity\ feature}$	2.9 (***)	1.3 (**)	2.9 (***)	2.0 (***)	8.8 (***)	5.1 (***)
$\lambda_{s, color\ feature}$	2.6 (***)	2.5 (***)	0.4 ( )	1.0 (*)	3.6 (***)	2.5 (***)

Table 4: Results of the type II Anova test - Sum of Squares for all covariants
*Signif. codes: $0 \leq$ '***' $< 0.001 \leq$ '**' $< 0.01 \leq$ '*' $< 0.05 \leq$ '.' $< 0.1 \leq$ ' ' < 1*

As shown in Table 4, the parameter for the display explains a great part of the variance in the sales of a given product. The product availability also provides a significant contribution to the variance. The self-price, price, fit and luminosity-based cannibalization terms also seem to have some explanatory power, whereas the contribution of the cannibalization term associated with the color attribute is more modest.

5.4 Limitations of the Model

The proposed model in this chapter provided a fairly simple method to measure cannibalization. The cannibalization appear to be statistically significant, but its economic magnitude is generally small. This may come from the very nature of the sales data. The sales at the product level, even averaged weekly, for each store, display a large variability of the product demand. Other factors may also contribute to the limitations of the model:

- The exact location of the products in the shop floor is not known. Still, it is believed that this location may greatly affect the sales of a given product. An article displayed at the entrance of the store usually generates more sales than a product located in the back of the store. The knowledge of "cold" and "hot spots" in the stores as well as the exact location of the products in the store could help understand some of the variation in the demand. However, due to the lack of information on product location, we considered that in the current model, the position of the products generated a noise that was identical for all the products.

- Another source of information that has not been used in our model are the "themes" defined by the designers for the season. The store managers arrange some products according to themes defined by the designers and the Merchandising team in La Coruña. These themes group several products of similar or complementary colors, patterns, as shown in Figure 8. These groupings are expected to have a correlated impact on the sales of the products belonging to a same theme, but this effect was ignored in our model, as information on the themes were limited at the time of the project.

- Another factor that limited the estimation of the cannibalization effect was the limited amount of data available. At the moment of the project, only limited RFID data was stored in Zara's database. In addition, we did not have enough time to reproduce the work involved to classify the products of additional seasons (as it was done in Chapter 4). The inclusion of additional data would help detecting seasonal trends. This can be done in a similar way to the work in the MIT9 project [6].

Figure 8: Examples of themes in Zara's stores for the Spring-Summer 2014 collection

Thus, as Zara continues to collect better information made available by the RFID deployment in its stores, a more detailed study can be achieved to complement the current cannibalization study. Still, the existing information and our product demand model were useful to improve Zara's demand forecast.

6 Forecast of the Product Demand

Using the analysis done for the cannibalization study, we improved the demand forecast for the products in Zara's stores. We defined the various forecast models in the first section, analyze the coefficients obtained in the second section and finally compare their accuracy in the third section.

6.1 Specification of the Forecasting Models

The model used to study the cannibalization in Chapter 5 introduced time effects and fixed effects to control for changes over time of the traffic in the store, in the popularity of the various attributes and for the variation of taste from one store to the next for the different attributes, fixed over time. These terms could not be used easily for forecasting. To take into account trends in the articles demand, we replaced in our specification the fixed and time effects by the lag sales of order 1 and of order 2. Throughout this chapter, we used the notation defined in Section 5.2.1.

6.1.1 M0 - Base Forecast Model

We defined as the base model M0 a simplified method of Zara's current forecasting model[3]. In the model M0, the sales of the product k at the time of week w in the store s was estimated as the average of the sales of the two previous weeks:

$$\text{sales}_{k,w,s} = \frac{\text{sales}_{k,w-1,s} + \text{sales}_{k,w-2,s}}{2} \tag{9}$$

Here, the sales terms represent the sales adjusted by possible days of sales (DPV - días posibles de ventas), similarly to Equation 3. An additional adjustment was necessary to account for a possible difference in the number of days the store is open in the two weeks considered. Thus, the sales of the week $w-1$ were defined as:

$$\text{sales}_{k,w-1,s} = \frac{\sum\limits_{day_{w-1} in\, week\, w-1} \text{sales}_{k,day_{w-1},s}}{\text{DPV}_{k,w-1,s}} * \text{DPV}_{k,w,s} \tag{10}$$

Other changes in the sales demand are discussed in [10]. We used the lag sales terms as defined in Equation 10 in the whole chapter but the parameters can be recalibrated if a different definition of the sales demand is chosen. The model M0 was used to compare the improvement of accuracy of the various forecasting models.

6.1.2 M1 - Pure Auto-regression Model

The second model, M1, uses a log-log regression model and lets the regression model calibrate the coefficients for the lag sales of week $w-1$and $w-2$.

[3]Zara's forecasting method also includes other adjustments not detailed in this thesis.

$$\log(\text{sales}_{k,w,s}+1) = \beta_s + \underbrace{\alpha_{sales,w-1,s}\log(\text{sales}_{k,w-1,s}+1)}_{\text{sales lag (week w-1)}} \underbrace{+\alpha_{sales,w-2,s}\log(\text{sales}_{k,w-2,s}+1)}_{\text{sales lag (week w-2)}} \quad (11)$$

where β_s is an intercept specific to each store.

6.1.3 M2 - Model with Lag Sales, Display, Availability and Price Parameters

For the following model, we introduced the display, availability and self-price effects as defined in Chapter 5:

$$\log(\text{sales}_{k,w,s}+1) = \beta_s + \underbrace{\alpha_{sales,w-1,s}\log(\text{sales}_{k,w-1,s}+1)}_{\text{sales lag (week w-1)}} \underbrace{+\alpha_{sales,w-2,s}\log(\text{sales}_{k,w-2,s}+1)}_{\text{sales lag (week w-2)}}$$

$$+ \underbrace{\alpha_{display\,product,s}\log\left(\text{display}_{k,w,s}+1\right)}_{\text{display effect}} + \underbrace{\alpha_{availability\,product,s}\log\left(\text{availability}_{k,w,s}+1\right)}_{\text{availability effect}}$$

$$+ \underbrace{\alpha_{self-price\,product,s}\log(\text{price}_k+1)}_{\text{self-price effect}} \quad (12)$$

This assume that we have knowledge of what products were displayed in the stores in the week w. Chapter 7 discusses this assumption but for now, we consider that the product assortment in the store is a decision that is perfectly known a week ahead. The availability factor was estimated based on the previous week stocks and estimated demand, as shown in Section 5.2.1.

6.1.4 M3 - Model with Cannibalization and Complementarity Effects

Finally, we defined a last model that accounts for the cannibalization and complementarity effects among products in the store:

$$\log(\text{sales}_{k,w,s}+1) = \beta_s + \underbrace{\alpha_{sales,w-1,s}\log(\text{sales}_{k,w-1,s}+1)}_{\text{sales lag (week w-1)}} \underbrace{+\alpha_{sales,w-2,s}\log(\text{sales}_{k,w-2,s}+1)}_{\text{sales lag (week w-2)}}$$

$$+ \underbrace{\alpha_{display\ product,s} \log\left(\text{display}_{k,w,s}+1\right)}_{\text{display effect}} + \underbrace{\alpha_{availability\ product,s} \log\left(\text{availability}_{k,w,s}+1\right)}_{\text{availability effect}}$$

$$+ \underbrace{\alpha_{self-price\ product,s} \log\left(\text{price}_k + 1\right)}_{\text{self-price effect}}$$

$$+ \underbrace{\sum_f \alpha_{subs,f} \text{SIM}_{f,k,w,s}}_{\text{similarity-based cannibalization}} + \underbrace{\alpha_{subs,price} \log(\text{Price}_{other\ products\ in\ cluster} + 1)}_{\text{price-based cannibalization}} \quad (13)$$

This equation is identical to Equation 8 in Section 5.2, except that the fixed and time effects are replaced by the two sales lag terms of the weeks $w-1$ and $w-2$.

6.2 Results - Parameter Estimation

The models defined in the previous section define coefficients for each store, but one can easily use the same model for a group of store. For the reader's ease of understanding, we grouped the stores by city.

Table 5 shows the coefficients obtained by running the regressions for the models M1, M2 and M3 for three cities (Madrid, Barcelona and London), both for the Fall-Winter 2014 and Spring-Summer 2015 seasons. From these results, we can make the following general comments:

- As expected, the sales lag term of order 1, $\log(\text{sales}_{k,w-1,s}+1)$, is the regressor that has the most explanatory power in our regression. For the model M1, the parameter estimation gives a weight even more important than for the model M0 (0.5). However, its coefficient declines as we add additional covariates in our models M2 and M3.

- The sales lag term for week $w-2$, $\log(\text{sales}_{k,w-2,s}+1)$, has a lesser weight than $\log(\text{sales}_{k,w-1,s}+1)$ and its coefficient also decreases as we add regressors. The coefficient is in the range 0.05-0.2, which is significant less than the one specified in the base model M0.

- The display effect (coefficient $\alpha_{display\ product}$) is of the same order as the one measured in our cannibalization study (see Table 3 in Section 5.3), though the coefficients found for the summer are lower than in Table 3. The display variable here also explains a significant amount of the variance in the sales. Note that although the coefficients is significant, the contribution to the variance of sales lag term of order 1 is higher than the contribution of the display effect. This is explained by the fact the covariates in our regression models are not standardized.

Coefficients	Value of the coefficients (Significance)					
	Winter			Summer		
	M1	M2	M3	M1	M2	M3
Madrid						
β	0.023 (***)	-0.044 (***)	-0.042 (***)	0.029 (***)	-0.060 (***)	-0.059 (***)
$\alpha_{sales, w-1}$	0.576 (***)	0.399 (***)	0.376 (***)	0.547 (***)	0.333 (***)	0.330 (***)
$\alpha_{sales, w-2}$	0.223 (***)	0.144 (***)	0.137 (***)	0.246 (***)	0.143 (***)	0.142 (***)
$\alpha_{display\ product}$		1.485 (***)	1.421 (***)		1.555 (***)	1.533 (***)
$\alpha_{availability\ product}$		0.216 (***)	0.206 (***)		0.242 (***)	0.241 (***)
$\alpha_{self\text{-}price\ product}$		-0.151 (***)	-0.166 (***)		-0.129 (***)	-0.178 (***)
$\alpha_{subs, fit}$			-0.008 (***)			-0.002 (***)
$\alpha_{subs, luminosity}$			0.005 (***)			0.003 (***)
$\alpha_{subs, color}$			0.001 (***)			-0.001 (*)
$\alpha_{subs, price}$			0.055 (***)			0.027 (***)
Barcelona						
β	0.022 (***)	-0.048 (***)	-0.047 (***)	0.037 (***)	-0.073 (***)	-0.070 (***)
$\alpha_{sales, w-1}$	0.567 (***)	0.420 (***)	0.402 (***)	0.560 (***)	0.357 (***)	0.345 (***)
$\alpha_{sales, w-2}$	0.258 (***)	0.189 (***)	0.181 (***)	0.237 (***)	0.149 (***)	0.143 (***)
$\alpha_{display\ product}$		1.033 (***)	0.999 (***)		1.494 (***)	1.454 (***)
$\alpha_{availability\ product}$		0.193 (***)	0.189 (***)		0.252 (***)	0.240 (***)
$\alpha_{self\text{-}price\ product}$		-0.079 (***)	-0.079 (***)		-0.121 (***)	-0.188 (***)
$\alpha_{subs, fit}$			-0.007 (***)			-0.005 (***)
$\alpha_{subs, luminosity}$			0.005 (***)			0.005 (***)
$\alpha_{subs, color}$			0.002 (***)			0.001 (***)
$\alpha_{subs, price}$			0.027 (**)			0.034 (***)
London						
β	0.020 (***)	-0.044 (***)	-0.043 (***)	0.037 (***)	-0.060 (***)	-0.053 (***)
$\alpha_{sales, w-1}$	0.549 (***)	0.372 (***)	0.366 (***)	0.582 (***)	0.268 (***)	0.239 (***)
$\alpha_{sales, w-2}$	0.165 (***)	0.122 (***)	0.121 (***)	0.116 (***)	0.059 (***)	0.051 (***)
$\alpha_{display\ product}$		1.312 (***)	1.209 (***)		1.574 (***)	1.588 (***)
$\alpha_{availability\ product}$		0.118 (***)	0.114 (***)		0.145 (***)	0.123 (***)
$\alpha_{self\text{-}price\ product}$		-0.084 (***)	-0.154 (***)		0.018 (*)	-0.285 (***)
$\alpha_{subs, fit}$			-0.005 (***)			0.002 (***)
$\alpha_{subs, luminosity}$			0.006 (***)			0.010 (***)
$\alpha_{subs, color}$			0.003 (***)			0.003 (***)
$\alpha_{subs, price}$			0.025 (***)			0.015 (.)

*Table 5: Coefficients and their significance for the Models M1, M2 and M3. Negative coefficients are displayed in red and positive coefficients are shown in blue. Signif. codes: $0 \leq$ '***' $< 0.001 \leq$ '**' $< 0.01 \leq$ '*' $< 0.05 \leq$ '.' $< 0.1 \leq$ ' ' < 1*

- The product availability and self-price effects also retain some significance for our forecast models and have, in general, a consistent sign - positive for the availability factor, and negative for the price elasticity. Only for one data sample (London store for the Fall-Winter 2013 season) does the price elasticity appear positive, though not very statistically significant.

- Finally, despite a few exceptions, the coefficients for the cannibalization/complementarity effects based on attributes and price of other products also tend to show the same signs as the ones found in Table 3 in Section 5.3. The sales of the products with the same fit tend to cannibalize each other, while the sales of products with the same color or color luminosity tend to complement each other. Finally, an increase in the price of the other products with the same set of attributes as a given product tends to increase the sale of that product.

The addition of new covariates improved the fit of the our models, as shown in Table 6. The adjusted R^2increased significantly between the models M1 and M2 but barely between models M2 and M3. The same conclusion can be drawn while observing Table 7.

Adjusted R^2						
Model	Madrid		Barcelona		London	
	Winter	Summer	Winter	Summer	Winter	Summer
M1	0.62	0.58	0.66	0.58	0.65	0.58
M2	0.70	0.68	0.70	0.67	0.71	0.72
M3	0.71	0.68	0.71	0.67	0.72	0.74

Table 6: Adjusted R^2 for the different models

AIC						
Model	Madrid		Barcelona		London	
	Winter	Summer	Winter	Summer	Winter	Summer
M1	-5166	-1923	-2947	-682	-1528	4686
M2	-9813	-6974	-4531	-3157	-4365	-562
M3	-10359	-7017	-4720	-3280	-4712	-1525

Table 7: Akaike's Information Criteria for the different models

To be able to measure an improvement in the forecast, it is necessary to measure the accuracy of the various methods and compare them together.

6.3 Results - Comparison of the Accuracy of the Forecasts

To measure the accuracy of the forecast, we used the wMAPE, as was done in previous projects [10, 6]. The wMAPE can be expressed as follows:

$$\text{wMAPE} = \frac{\sum \mid \hat{y} - y \mid}{\sum y} \tag{14}$$

where $\hat{y}$ is the predicted value of the dependent variable, y is the actual value of the dependent variable. In the following sections, to calculate the wMAPE and predict the demand, we used the store-specific coefficients.

6.3.1 In-sample wMAPE

First, we measured the in-sample wMAPE for the six data samples. The wMAPE was calculated at the product level for the group of stores in each city. Table 8 compares the wMAPE of the various models and the wMAPE change compared to the base model, M0.

Model	WINTER		SUMMER	
	wMAPE	%Δ	wMAPE	%Δ
Madrid				
M0	70%		70%	
M1	65%	-8%	64%	-9%
M2	59%	-16%	58%	-18%
M3	59%	-17%	57%	-18%
Barcelona				
M0	63%		66%	
M1	61%	-3%	63%	-5%
M2	58%	-8%	57%	-14%
M3	58%	-8%	57%	-15%
London				
M0	87%		84%	
M1	63%	-27%	66%	-22%
M2	58%	-34%	55%	-35%
M3	56%	-35%	54%	-36%

Table 8: In-Sample wMAPE for the six data samples
The wMAPE columns represent the measure of accuracy, while the %Δ columns represent the reduction in wMAPE compared to the model M0.

As shown in Table 8, the models M1 and M2 present significant gains in accuracy compared to the base model M0. However, the model M3 barely improved the wMAPE compared to the model M2. Therefore, the inclusion of the cannibalization terms in the model did not seem to improve the accuracy of the forecast significantly.

6.3.2 Out-of-sample wMAPE

This time, we calculated the out-of-sample wMAPE on the six data samples. To that purpose, we used the parameters estimated on one season to predict the demand of the other season. For example, the coefficients calibrated on the Fall-Winter 2013 season for the stores in Madrid were used to predict the Spring-Summer 2014 season for the same stores. Therefore, the six data samples are associated with six out-of-sample datasets.

The results are shown in Table 9.

Model	WINTER		SUMMER	
	wMAPE	%Δ	wMAPE	%Δ
Madrid				
M0	70%		70%	
M1	66%	-7%	64%	-9%
M2	60%	-14%	58%	-18%
M3	60%	-15%	58%	-17%
Barcelona				
M0	63%		66%	
M1	63%	-1%	62%	-7%
M2	60%	-5%	57%	-14%
M3	59%	-6%	57%	-15%
London				
M0	87%		84%	
M1	64%	-26%	65%	-23%
M2	64%	-27%	59%	-30%
M3	61%	-29%	58%	-31%

Table 9: Out-of-Sample wMAPE for the six data samples
The wMAPE columns represent the measure of accuracy, while the %Δ columns represent the reduction in wMAPE compared to the model M0.

The out-of-sample wMAPE table was similar to the in-sample wMAPE table (Table 8). The models M1 and M2 provide an improvement in the accuracy of the forecast, whereas the model M3 again does not deliver significant reduction in wMAPE compared to the model M2.

In conclusion, the existing forecast method used by Zara can be improved with the display, availability and product price information. Despite the slight increase in accuracy, the model with attribute-based cannibalization and complementarity effects requires much more information than the model with only the display, availability and product price information. Therefore, it may be reasonable to only use the latter in the short-term to improve the demand forecast. All of the work in this chapter assumed a perfect information of the display for the week ahead, which is not as obvious as what it may appear. The following chapter discusses this issue.

7 Demand Forecast with the Estimated Display

7.1 An Uncertain Display

So far, we have assumed that the information on which product is displayed on the shop floor is perfectly known a week ahead. However, the reality is not that simple and the product assortment is not always deterministic for different reasons.

First, the store managers retain a very large power of decisions over what is displayed on the shop floor. Historically, the store managers had a lot of independence to decide which product they wanted to receive from the distribution centers or they could leave in the stock room. Zara's strategy still relies on rapid feedback from the customers across the supply chain and store managers are the first contact with the customers, and so the first to be able to react to the demand. Hence, despite the general operation guidelines provided by the Merchandising and Store Operations Team in La Coruña, the final decision on product assortment belongs to the stores.

Moreover, even for the store managers, predicting exactly one week ahead the product assortment at any time of the coming week is not possible. For example, because of the highly fluctuating demand, it is difficult to predict when a product is likely to stock out during the week and therefore to be taken out of the shop floor.

Last, the information on the product assortment collected from the database is a snapshot of the state of the stores at one moment of the day. However, even within a day, the store managers display new products or remove the ones for which only a few sizes are available. Therefore, the display information currently used in the project is only an approximation. The team carefully chose the time at which the product assortment information is retrieved from the database so that the data are representative of the display during the whole day. It was timed to be after the delivery of new articles to the store, after the daily product changeover on the shop floor, but early during the day to capture how the customers sales react to the current product assortment.

Given these reasons, the state of display of a product is not perfectly known and need to be estimated a week ahead[4]. The following subsections describe the approach taken to estimate this covariate and the results of the demand forecast using the estimated display instead of the actual one. Note that, as in the previous sections, we focus on the pants product category in Zara's stores in Spain and the United Kingdom.

7.2 Estimation of the Display a Week Ahead

To evaluate the display of week w at the time of the week $w-1$, we considered a couple of approaches.

A possible forecasting model is logistic regression. The decision to display a product on a particular day can be seen as a binary variable which is made considering factors such as past sales,

[4] This can appear counterintuitive when considering that in the product assortment literature, the display decision is seen as an input to optimize the operating margin. However, the need to estimate the display can be interpreted as an additional uncertainty factor on how the system operates in reaction to the decision of displaying an article in the store.

past display state or the product stock. Substitution terms can be added in these models, as we think of the store managers facing a limited shop floor capacity and more products than they can display in the store to the customers. The weekly average display variable is then the probability that a given product is displayed in any day in the week w, given the information available the previous week.

However, the simplest method showed very adequate results in this case. We first started by assuming that the average display of a product in the week w was identical to the average display of the previous week, or to the last display information known for the product in the week w. The measured accuracy, using the wMAPE for comparison between the methods, were in the range of 17-23% for the first method (method a) in Table 10) and of 12-17% for the second one (method b) in Table 10). However, a closer look at the accuracy for different quantiles of weekly product sales showed that the wMAPE for the products not displayed the week $w-1$ had a wMAPE of 100%. The explanation for this value comes from the fact that by taking the display of the previous week, we do not account for the introduction of new products in the stores, which is a process done twice a week in all Zara's stores. The information of delivery of new products at the store is available the previous week based on decisions made by the distribution team, the products managers and buyers at Zara's office in La Coruña. In addition, the store managers have the instructions of displaying the new products in priority for at least a week when receiving them. Consequently, we can correct the display of the week by $w-1$ with this new piece of information, as shown in methods c) and d) in Table 10. The results of the various display estimation methods are shown in Table 10.

	wMAPE			
City	a) display$_{w-1}$	b) display$_{w-1}$ (correction for new product intro.)	c) display of the last day in week w-1	d) display of the last day in week w-1 (correction for new product intro.)
Madrid	18%	13%	12%	9%
Barcelona	17%	11%	12%	8%
London	23%	15%	17%	12%

Table 10: wMAPE for the two methods to estimate the display of the week $w+1$

As shown in Table 10, the methods with the corrected display taking into account the information for new product introduction give a good forecast of the future display[5]. The method using the last known display information of a product logically provides the best accuracy (method d) in Table 10 and in Figure 9). This confirms that the stores mainly make marginal changes to the product assortment in the store, to make room for new products. The accuracy of the forecast also changes significantly depending on the sales of the product (see Figure 9 below).

The forecast is more accurate for the best selling articles. This is logical, since the store managers

[5]Forecasts made with the logistic regression displayed a lower accuracy. This may be due to the large amount of products displayed for the entire week or not displayed at all. In this case, most of the values of the display variable take a value of 0 or 1. With the logistic regression, the probability can never reaches 0 or 1 and therefore the wMAPE appears to be higher.

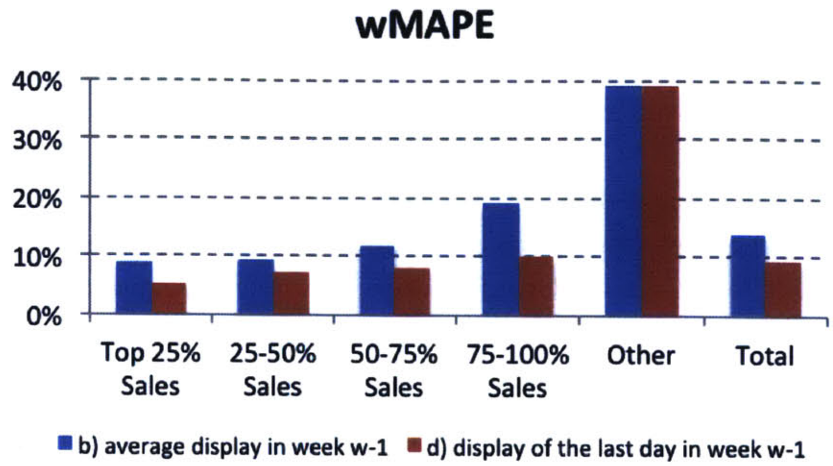

Figure 9: wMAPE of the display estimation, by quartile of sales. The "other" category groups the products that were not displayed the previous week. The two methods correct the forecast of the display for new product introduction.

actively manages its product assortment to make sure that the top 50% best selling articles are displayed in their stores. On the other hand, the display products belonging to the fourth quartile in terms of sales ranking is more difficult to forecast, as those articles have a higher probability of going back to the stock room to leave room for the new products coming in.

Thus, the information of the display of the week $w-1$ combined with the information of the new product introduction in week w can be used to estimate the display of week w with a satisfactory accuracy. The next step is to use the estimated display to forecast the sales.

7.3 Forecasting Models with the Estimated Display

We used the same models described in Section 6.1 with the estimated display. We focus here on models M2 and M3 (see Equations 12 and 13).

The models M2p and M3p use the same equation as, respectively, the models M2 and M3 except that the actual display is replaced by the estimated display discussed in the previous section. The similarity factors and price-based cannibalization terms were also recalculated with the estimated display. The value of the coefficients obtained with the regressions for the model are displayed in Table 11.

The signs of the coefficients are consistent with the ones observed in Table 5 and the same cannibalization, complementarity effects can be observed with these new results. The coefficients for the sales lag of order 1 and 2 are slightly higher than in Table 5 but still comparable.

Regarding the forecast accuracy, we expected that the wMAPE of the models M2p and M3p were higher than the ones of models M2 and M3, to reflect the loss of information on the product

Coefficients	Value of the coefficients (Significance)			
	Winter		Summer	
	M2p	M3p	M2p	M3p
Madrid				
β	-0.043 (***)	-0.042 (***)	-0.067 (***)	-0.067 (***)
$\alpha_{sales, w-1}$	0.426 (***)	0.406 (***)	0.372 (***)	0.371 (***)
$\alpha_{sales, w-2}$	0.160 (***)	0.154 (***)	0.154 (***)	0.154 (***)
$\alpha_{display\ product}$	1.693 (***)	1.611 (***)	2.024 (***)	1.841 (***)
$\alpha_{availability\ product}$	0.290 (***)	0.283 (***)	0.346 (***)	0.342 (***)
$\alpha_{self\text{-}price\ product}$	-0.249 (***)	-0.225 (***)	-0.282 (***)	-0.282 (***)
$\alpha_{subs,\ fit}$		-0.007 (***)		-0.002 (***)
$\alpha_{subs,\ luminosity}$		0.004 (***)		0.002 (***)
$\alpha_{subs,\ color}$		0.001 (***)		0.000 (*)
$\alpha_{subs,\ price}$		0.040 (***)		0.023 (***)
Barcelona				
β	-0.045 (***)	-0.043 (***)	-0.080 (***)	-0.077 (***)
$\alpha_{sales, w-1}$	0.455 (***)	0.438 (***)	0.408 (***)	0.398 (***)
$\alpha_{sales, w-2}$	0.202 (***)	0.194 (***)	0.172 (***)	0.167 (***)
$\alpha_{display\ product}$	1.254 (***)	1.122 (***)	1.653 (***)	1.451 (***)
$\alpha_{availability\ product}$	0.256 (***)	0.252 (***)	0.381 (***)	0.372 (***)
$\alpha_{self\text{-}price\ product}$	-0.173 (***)	-0.157 (***)	-0.223 (***)	-0.230 (***)
$\alpha_{subs,\ fit}$		-0.007 (***)		-0.004 (***)
$\alpha_{subs,\ luminosity}$		0.004 (***)		0.004 (***)
$\alpha_{subs,\ color}$		0.002 (***)		0.002 (***)
$\alpha_{subs,\ price}$		0.029 (**)		0.026 (**)
London				
β	-0.040 (***)	-0.039 (***)	-0.086 (***)	-0.081 (***)
$\alpha_{sales, w-1}$	0.426 (***)	0.421 (***)	0.376 (***)	0.352 (***)
$\alpha_{sales, w-2}$	0.129 (***)	0.128 (***)	0.083 (***)	0.076 (***)
$\alpha_{display\ product}$	0.996 (***)	0.344 (*)	1.892 (***)	0.681 (**)
$\alpha_{availability\ product}$	0.201 (***)	0.196 (***)	0.424 (***)	0.413 (***)
$\alpha_{self\text{-}price\ product}$	-0.096 (***)	-0.058 (*)	-0.208 (*)	-0.181 (***)
$\alpha_{subs,\ fit}$		-0.006 (***)		-0.003 (**)
$\alpha_{subs,\ luminosity}$		0.006 (***)		0.009 (***)
$\alpha_{subs,\ color}$		0.003 (***)		0.004 (***)
$\alpha_{subs,\ price}$		0.038 (***)		0.022 (*)

Table 11: Coefficients and their significance for the Models M2 and M3 with an estimated display covariates.

Negative coefficients are displayed in red and positive coefficients are shown in blue.
*Signif. codes: $0 \leq$ '***' $< 0.001 \leq$ '**' $< 0.01 \leq$ '*' $< 0.05 \leq$ '.' $< 0.1 \leq$ ' ' < 1*

display. The table showing the In-Sample wMAPE and Out-of-Sample wMAPE for the two new models confirm our expectation. Still, M2p and M3p offer a consistent improvement in wMAPE compared to the base model M0.

In-Sample wMAPE				
Model	WINTER		SUMMER	
	wMAPE	%Δ	wMAPE	%Δ
Madrid				
M2	59%	-16%	58%	-18%
M3	59%	-17%	57%	-18%
M2p	63%	-11%	61%	-13%
M3p	62%	-12%	61%	-13%
Barcelona				
M2	58%	-8%	57%	-14%
M3	58%	-8%	57%	-15%
M2p	60%	-5%	60%	-10%
M3p	59%	-6%	60%	-10%
London				
M2	58%	-34%	55%	-35%
M3	56%	-35%	54%	-36%
M2p	61%	-29%	63%	-26%
M3p	60%	-31%	62%	-27%

Out-of-Sample wMAPE				
Model	WINTER		SUMMER	
	wMAPE	%Δ	wMAPE	%Δ
Madrid				
M2	60%	-14%	58%	-18%
M3	60%	-15%	58%	-17%
M2p	63%	-10%	61%	-13%
M3p	62%	-11%	62%	-12%
Barcelona				
M2	60%	-5%	57%	-14%
M3	59%	-6%	57%	-15%
M2p	61%	-3%	60%	-10%
M3p	60%	-5%	60%	-10%
London				
M2	64%	-27%	59%	-30%
M3	61%	-29%	58%	-31%
M2p	63%	-28%	64%	-24%
M3p	62%	-29%	64%	-23%

Table 12: In-Sample and Out-of-Sample wMAPE for the six data samples. The variation (%Δ) are calculated compared to the model M0.

In conclusion, the display of a product can be estimated one week ahead just by using the last information on the display and information on new product arrivals. The newly estimated display was used in the demand forecast. This resulted in significant improvement compared to the currently used forecast model, despite the loss of information regarding the display.

8 Conclusions

8.1 Contributions

Our work on sales cannibalization and demand forecasting has made the following contributions.

First, we created a log-log model of sales with cannibalization effects. It focused on cannibalization due to similarities among products displayed in the store and on price-based substitution. Similar models can be found in the literature on supermarket product assortment. However, in this thesis, we applied these models to hundreds of products that have highly variable demand and a short life cycle. We calculated a similarity factor between products, taking into account the average display effect in the stores. We conducted an estimation of the parameters in a data sample of 13 stores, in three different cities. The results we obtained were consistent across the various cities, for both seasons studied in the sample.

Coefficients	Effect
$\alpha_{subs, fit}$	-
$\alpha_{subs, color\ lightness}$	+
$\alpha_{subs, color}$	+
$\alpha_{subs, price}$	+

Table 13: Summary of the cannibalization or complementarity effects by attribute and price

As shown in Table 13, the effect among products with the same fit are negative, which translates into a cannibalization effect, whereas the effects among products with the same color or the same color luminosity are positive—and therefore complementary. In addition, an increase in price of other similar products with the same set of attributes (fit, color, color lightness) as those of a given product has a positive impact on the sales of that product, as one might expect from economic models of multi-product demand. Yet, these cannibalization and complementarity effects appear rather small compared to other explanatory variables, such as the product display or availability. This is for two reasons. First, a lot of noise is present in the data. Second, other effects—such as the exact location of the products in the store—are not measured despite having a significant impact on sales.

The second contribution of this work was on the accuracy of the demand forecast. We tested multiple models, with or without cannibalization terms. The predictive power of the model with cannibalization effects was slightly better than the others but requires significant data preparation. A model with only the display, availability and price effects still achieves substantial improvements in accuracy, and presents the advantage of simplicity. The results of the latter for out-of-sample data are shown in Table 14.

wMAPE reduction		
City	Actual Display	Estimated Display
Madrid	-18%	-11%
Barcelona	-10%	-7%
London	-43%	-26%

Table 14: Improvement in wMAPE by city with the new demand forecast model

The forecast accuracy, measured with the wMAPE, increases by at least 10% when the product display is perfectly known. However, because the display is not always known by Zara's distribution team a week ahead, we calculated an estimated display, which had an accuracy of about 10%. With the estimated display, the demand forecast accuracy was slightly reduced, but nevertheless resulted in significant improvements. Another way to quantify the impact of the new forecast is to estimate the reduction in lost sales. Calculating lost sales can be difficult, since it requires modeling how the new forecast, relative to the existing one, would have differentially impacted replenishment decisions and in turn inventory levels. However, we can still estimate lost sales in the cases where we observe an actual out-of-stock event, using the following formula:

$$\text{avoidable lost sales} = \min\{0, \min\{\text{hypothetical demand,forecast}\} - \text{sales}\} \tag{15}$$

where 'hypothetical demand' is the demand in the absence of stock-out, calculated by assuming that sales on stock-out days would have been equal to average sales on days without stock-outs. Calculations for the trousers product category in the sample stores in Madrid show that an upper bound estimation of the avoidable lost sales amount to 1.0% of the total product category sales.

Hence, the project showed the value of the information provided by the RFID technology to improve the understanding of the interaction of products in a store and the demand forecast.

8.2 Implementation

The results uncovered in this thesis can be very useful for Zara to improve its demand forecast accuracy. However, some additional work is necessary to be able to implement the work into Zara's tools.

First, the focus of this thesis was on the trousers product category. Similar work can be done on all other other products category—a study of the cannibalization and complementarity effects can be undertaken on the various product families, based on the characteristics of the product. However, this will require to improve the product classification. The classification process can be reproduced based on the methodology described in Chapter 4. Another solution would be to develop the classification at the time of the creation of the new products—in other words, at the design phase. This would involve modifying the work process of the buyers when entering the new products of a collection into Zara's database system. However, the accuracy improvement between the models without and with the cannibalization effects (models M2, M2p and M3, M3p in Sections 6 and 7) is small. Therefore, to gain most of the accuracy improvement in the forecast, the classification may not be necessary and the models with the display, availability and price effects (M2 and M2p) could be sufficient for Zara in the short term.

Additionally, a larger test may be needed before the demand forecast improvement is implemented into Zara's replenishment decision process. The project was done with limited historical data for selected stores, due the recent and still undergoing RFID deployment. A larger data sample would allow to better measure the cannibalization effects in the stores, especially if we are able to capture seasonal trends. Currently, with only one year of data, the seasonal effects were included in the fixed and time effects of our models in Chapter 5. In addition, the improvement in demand forecast accuracy was very significant and should be tested in additional stores, in different regions and countries.

An additional step towards implementation is to integrate the current RFID and sales database. A significant amount of time in the project was dedicated to assemble and clean the merged database. This process can be simplified and the information can be collected more efficiently at the database level. This work will be necessary before implementing the improved demand forecast model into Zara's replenishment process. The current code of the demand forecast will also be dramatically simplified as the result of the integration of the databases.

8.3 Future Improvements

The work in this thesis introduced key new estimates of the cannibalization effects among products in Zara's stores, which can help inform display decisions for the store managers. It also presented demand forecast models that can improve significantly the accuracy of Zara's existing demand forecast. Still, a better understanding of the cannibalization and further improvements of the forecast can be achieved. We discuss below a few suggestions of directions for future research.

- *Locational information.* At the current time, the exact location of the articles in the stores is not known. Yet, this piece of information is known to impact greatly the sales of a product. In fact, store managers know very well the existence of "hot spots", where articles are more likely to sell well—and their inverse, "cold spots"—within the stores. Zara's team is currently in the process of adding locational information, which therefore could be used to improve the accuracy of the forecast model. The information should be fairly easy to integrate into the models presented in this thesis.

- *Controlled experiment to study cannibalization.* The difficulty in the current study was to capture cannibalization effects, given the heterogeneity of the stores and customer profiles of each store. Dozens of new products were also introduced during the same week in each location, throughout the whole year, which made it delicate to measure substitution with so many new articles arriving at the same time. The best way to measure accurately effects such as cannibalization and price elasticity would be to run a controlled experiment, where the introduction of a new product would be randomized across stores and its effect on the sales of other items precisely tracked to determine the causal effect of this introduction.

- *Product assortment planning.* The demand forecast is often seen as the first step in optimizing the product assortment in the stores. Currently, the designers and buyers, in close collaboration with the product managers, decide on the composition of a new collection. Understanding better how products interact together, and optimizing the product assortment in the stores to increase sales could be useful for Zara as it prepares its new collections for the future.

Appendix A - Clusters

Cluster id	color	color lightness	fit
1	blue	medium	slim
2	blue	dark	slim
3	grey	medium	large
4	grey	light	large
5	blue	medium	straight
6	colored	light	large
7	blue	medium	large
8	blue	light	slim
9	grey	dark	slim
10	grey	light	slim
11	colored	medium	slim
12	colored	dark	slim
13	colored	light	slim
14	geometric	dark	slim
15	flower/animal	light	slim
16	blue	dark	straight
17	flower/animal	medium	slim
18	flower/animal	dark	slim
19	grey	dark	straight
20	geometric	medium	straight
21	geometric	dark	straight
22	grey	dark	large
23	colored	dark	large
24	geometric	dark	large
25	colored	medium	large
26	geometric	medium	large
27	blue	light	straight
28	blue	light	large
29	colored	light	straight
30	grey	light	straight
31	geometric	light	slim
32	colored	dark	straight
33	colored	medium	straight
34	blue	dark	large
35	geometric	light	straight
36	flower/animal	medium	straight
37	flower/animal	dark	straight
38	flower/animal	medium	large
39	flower/animal	light	large
40	geometric	medium	slim
41	grey	medium	slim
42	flower/animal	dark	large
43	flower/animal	light	straight
44	geometric	light	large

Table 15: list of the product clusters

Figure 10: Examples of products in the clusters (1/2)

Figure 11: Examples of products in the clusters (2/2)

References

[1] Sull, D. & Turconi, S. (2008). Fast fashion lessons. Business Strategy Review, 29(2), 4–11.

[2] A. Gürhan Kök. Assortment Planning: Review of Literature and Industry Practice. Retail Supply Chain Management. Eds. N. Agrawal and S. A. Smith, Kluwer Publishers. April 2006, revised November 2006

[3] M. K. Mantrala, M. Levy et al., Why is Assortment Planning so Difficult for Retailers? A Framework and Research Agenda. Journal of Retailing 85, 1, 2009, 71–83

[4] Inditex Group Annual Report 2013

[5] Tokatli, N. (2008). Global sourcing: Insights from the global clothing industry – the case of Zara, a fast fashion retailer. Journal of Economic Geography, 8, 21–38.

[6] J. M. Garcia. Improving The Forecast Model Of Fast Fashion Items From Central Warehouses To Stores. MIT LGO thesis, June 2014.

[7] R. C Borrase. Optimization of Zara's Clearance Sale Pricing. MIT LGO thesis, June 2009.

[8] Pankaj M. Madhani. Fast fashion retailing: RFID deployment for enhancing customer value proposition. Fast fashion systems: theories and applications. editor, Tsan-Ming Choi. CRC Press/Balkema. 2014.

[9] http://www.inditex.com/media/news_article?articleId=150174 - accessed on February 2nd, 2015.

[10] A. Garro. New Product Demand Forecasting and Distribution Optimization: A Case Study at Zara. MIT LGO thesis, June 2011.

[11] M.G. Harvey and R.A. Kerin, Diagnosis and management of the product cannibalism syndrome; University of Michigan Business Review; 1979; Vol 31 No.6, pp. 18-29

[12] J. Heskett. Marketing. Macmillan Publications, 1976. New York, NY.

[13] C.H. Mason and G.R. Milne. An approach for identifying cannibalization within product line extensions and multi-brand strategies. Journal of Business Research, 1994, Vol. 31, pp. 267-281

[14] W. Copulsky. Cannibalism in the marketplace. Journal of Marketing, 1976. Vol. 2, pp. 103-105

[15] S. Raghavan Srinivasan, S. Ramakrishnan and S. E. Grasman. Identifying the effects of cannibalization on the product portfolio. Marketing Intelligence & Planning; 2005; 23, 4/5; ABI/Inform Complete pg. 359

[16] S. Raghavan Srinivasan, S. Ramakrishnan and S. E. Grasman. Incorporating cannibalization models into demand forecasting. Marketing Intelligence & Planning; 2005; 23, 4/5; ABI/Inform Complete pg. 470

[17] K. Rajaram. Assortment planning in fashion retailing: methodology, application and analysis. European Journal of Operational Research 129 (2001), pp. 186-208

[18] Robert P. Rooderkerk, Harald J. van Heerde, Tammo H. A. Bijmolt (2013) Optimizing Retail Assortments. Marketing Science 32(5):699-715

[19] D. McFadden. Conditional logit analysis of qualitative choice behavior; P. Zarembka (ed.); Frontiers in econometrics, 105-142, Academic Press: New York, 1973.

[20] K. E. Train. Discrete Choice Methods with Simulation; 2nd ed., Cambridge University Press, 2009.

[21] Data Mining and Knowledge Discovery Handbook. Edited by O. Maimon and L. Rokach. Springer, 2005.

[22] D. Bertsimas and R. Shioda. Classification and Regression via Integer Optimization. Operations Research; Vol. 55, No. 2, March–April 2007, pp. 252–271

[23] M. West, J. Harrison. Bayesian Forecasting and Dynamic Models. 2nd ed, Springer, 1997

[24] Y. Aviv, A Time-Series Framework for Supply-Chain Inventory Management. Operations Research; Vol. 51, No. 2, March–April 2003, pp. 210–227