Aalto University
School of Science
Degree Programme in Mathematics and Operations research

Joonas Vauhkonen

Cross-Price Elasticities in Fashion Retail Pricing

Master's thesis
Espoo, April 16, 2018

Supervisor: Professor Harri Ehtamo
Thesis advisor: M.Sc Sophia Hemstad

Aalto University
School of Science

Abstract of master's thesis

Author	**Joonas Vauhkonen**
Title of thesis	**Cross-Price Elasticities in Fashion Retail Pricing**
Master's programme	**Mathematics and Operations Research**
Thesis supervisor	**Professor, Harri Ehtamo**
Major or Minor/Code	**Systems and Operations research / SCI3055**
Department	**Engineering Physics and Mathematics**
Thesis advisor(s)	**M.Sc Sophia Hemstad**
Date: April 16, 2018	**Number of pages: 5 + 71** **Language: English**

Abstract

This thesis aims to produce insight into the price and demand relationship in a fashion retail setting. The research goals are to create cross-price elasticity values for product groups and to use these values in price optimization. The obtained results will be used as suggestions when making product pricing decisions. Any results obtained are estimates, not exact due to the nature of elasticity estimations.

This research is based on a thesis work by another student, in which it was found that many of the products offered by the Company X have elastic demand. That notion motivated this research project, which extends the scope of the previous work and builds upon it. The data processing and modeling practices were determined in the early stages of the thesis project and stayed somewhat unaltered throughout the work. The extension from single products to product groups required changing from linear to multiple regression which produced the cross-price elasticity coefficients. The price optimization was run on several optimization methods and the results were compared and reflected on reality. The optimization methods are covered briefly in the mathematical section of the thesis. The thesis provides a brief look into the retail business and market theory as well as the mathematical aspects of the project. The most interesting concept for this thesis is price elasticity of demand, which describes how changing products' prices affects their demand. This concept and its interpretations are covered in the theoretical section of the thesis.

Unfortunately, the results obtained from the computations were not of very good quality when statistical significance and the level of explained variation is considered. The values computed for single products were of better quality, which would encourage further research on individual products. However, the results provide the desired insight into pricing. Furthermore, it was noted that more development is needed to reach an even higher level of quality and more usable results especially with product groups.

Keywords: Retail, Pricing, Elasticity of demand, Optimization

Aalto-yliopisto
Perustieteiden
korkeakoulu

Diplomityön tiivistelmä

Tekijä	**Joonas Vauhkonen**
Työn nimi	**Cross-Price Elasticities in Fashion Retail Pricing**
Koulutusohjelma	**Teknillisen fysiikan ja matematiikan koulutusohjelma**
Valvoja	**Professori, Harri Ehtamo**
Pää tai sivuaine/koodi	**Systeemitieteet / SCI3055**
Työn ohjaaja(t)	**M.Sc Sophia Hemstad**
Päivämäärä: 16.4.2018	**Sivumäärä: 5 + 71** **Kieli: Englanti**

Tiivistelmä

Tämän diplomityön tarkoituksena on tuottaa laskennallisia tuloksia, joiden perusteella voidaan saavuttaa syvempää ymmärrystä myyntihinnan ja tuotteiden kysynnän välisestä suhteesta. Tutkimus on tehty yhteistyössä muotialan vähittäiskauppaa harjoittavan Yritys X:n kanssa. Tavoitteena on luoda tuoteryhmäkohtaiset kysynnän ristielastisuusarvot, joita käytettäisiin myös hintaoptimoinnissa. Laskennan tuloksena saatuja arvoja käsitellään suuntaa antavina lukuina hinnoittelupäätöksiä tehtäessä. Kaikki tulokset ovat arvioita, koska elastisuuksia ei voida määrittää riittävällä tarkkuudella.

Tässä työssä tehtävä tutkimus pohjautuu toisen opiskelijan opinnäytetyöhön ja siinä saavutettuihin tuloksiin, joiden mukaan monien Yritys X:n tarjoamien tuotteiden kysynnässä on havaittavissa elastisuutta. Kyseisen tutkimuksen tulokset johtivat tähän projektiin, jossa edellistä tutkimusta laajennetaan yksittäisen tuotteen käsittelemisestä tuoteryhmiin. Laajentaminen tuoteryhmiin vaati siirtymistä lineaarisesta regressiosta usean selittäjän regressioon (multiple regression), jonka tuotteena ristielastisuusarvot saadaan laskettua. Hintaoptimoinnissa hyödynnetään useampaa optimointimenetelmää, joiden tuottamia tuloksia vertaillaan ja peilataan todellisuuteen. Eri menetelmät esitellään työn teoriaosiossa. Datan käsittelyyn ja mallintamiseen liittyvistä menetelmistä sovittiin projektin alussa ja ne pidettiin lähes muuttumattomina loppuun saakka. Työn teoriaosio tarjoaa näkökulmaa muotialan vähittäiskauppaan ja yleisesti markkinateoriaan, sekä työn matemaattisiin käsitteisiin. Tämän diplomityön kannalta tärkein ja mielenkiintoisin käsite on kysynnän elastisuus, joka kuvaa tuotteen kysynnän muutosta myyntihinnan muuttuessa. Kyseinen ilmiö käsitellään tarkasti työn teoriaosiossa.

Työssä saavutetut tulokset jäivät valitettavasti laadultaan heikoiksi etenkin tilastolliselta merkitsevyydeltään. Käytetyt mallintamismenetelmät eivät pystyneet selittämään datassa esiintynyttä variaatiota riittävän hyvin tuoteryhmiä käsiteltäessä. Yksittäisiä tuotteita tarkasteltaessa laatua saatiin parannettua huomattavasti, mistä johtuen tulevaisuudessa voitaisiin keskittyä tuoteryhmien sijaan laajempaan yksittäisen tuotteen tarkasteluun. Tuloksista saadaan toivottua näkemystä hinnoitteluun, vaikkei toivotussa määrin. Lisäksi, tätä tutkimusta tulisi kehittää edelleen, esimerkiksi kohti dynaamista hinnoittelumallia, jotta voitaisiin siirtyä tuoteryhmien käsittelyyn ja saavuttaa laadukkaampia tuloksia niiden hinnoittelua varten.

Avainsanat: Vähittäiskauppa, Hinnoittelu, Kysynnän jousto, Elastisuus, Optimointi

Acknowledgements

This thesis project was done in collaboration with Company X. The project is a continuum for another thesis. I was lucky enough to meet the person who wrote the preceding thesis work as well as receive an interesting and motivating thesis topic from a company in a field that I find very interesting.

First of all, I would like to thank my thesis supervisor Harri Ehtamo for his guidance and encouragement during this thesis work. I also thank my thesis advisor Sophia Hemstad for answering for the project. The circumstances were not optimal with this project, but I hold it in high regard that you took the ownership of this project.

Further thanks go to Lasse, Peppe and Denise for helping me and taking time to listen to the advancement of my work in addition to your own tasks and busy schedules. Your insight and ideas as well as sparring have been valuable throughout this project. Special thanks to the person who took me on for this project, David. Thank you for giving me the opportunity and trusting me to be able to finish this project. Your advice and guidance in the beginning of this thesis work has been extremely valuable. Your enthusiasm for business development was contagious to say the least. Ever since I started this project, the company's local office has welcomed me as a part of the team. Huge thank you to everyone I have had the pleasure of working with in the organization.

Finally, without naming everyone separately, I want to thank everyone who has taught me and guided me during my years in Aalto-university, from teachers and lecturers to the course assistants and faculty staff alike.

Espoo, April 16, 2018

Joonas Vauhkonen

Contents

1 Introduction

This thesis is based on a project, conducted in collaboration with a fashion retail company. The project was initiated by this fashion retail company (from here on referred to as Company X), due to high competition on the field and the desire to set some statistical basis on pricing. We have access to real transactional data, as well as other deliberately selected tables of information that are used to refine the raw data and develop our model further. Most of the data, graphs and exact results have to be left out of this thesis due to confidentiality issues, but we will do our best to describe them in order to provide the reader a good picture of what was accomplished. The goal of this thesis is to compute elasticity of demand values to both individual products and product groups, as well as look for optimal pricing suggestions to presented product groups.

The retail market is a highly competitive field, with numerous companies, retail chains, online retailers and other wholesalers all competing with each other for the same consumers. Each field under the concept of retail has its' own characteristics and unique aspects. One of such interesting fields is fashion retail. Namely, seasonality and ever changing trends have a significant effect on fashion retail among other variables, such as brand and store preferences. The latter ones might be rooted deeper in the individual, whereas the first ones are more universal, quickly changing and affecting all potential consumers. Nowadays, the development of Internet provides new information outlets to consumers and thus increases the consumer awareness and cost consciousness, since there are increasingly large amounts of information available. The fashion retail companies must be able to handle quick seasonal and trend changes, answer the consumers' demand and recognize the consumer's awareness while upholding their own image in order to gain an upper hand over the competition. This thesis addresses one of the issues, that retail companies can do to achieve advantage through more effective pricing.

Often, retail companies follow some "rules" when pricing their products and deciding on markdowns (Levy et al. 2004). Such practice may work since it is based on past experiences of successful decisions. Even though such easy way is lucrative, it might be more efficient to utilize the collected data to more extent. Especially nowadays, there are huge amounts of data available that can be analyzed. The transactional data that retail companies have, can be regarded as one of the most valuable assets in their disposal. It might be rather effortless to lower a price for one product and observe the increase in demand, whereas on the other hand it could be far more difficult to observe the effect that same price adjustment has on other items. The data analysis could provide insight to that dimension as well. Recently, such ideas have been embraced by some companies using customized dynamic pricing models for their own markets, utilizing, among other factors, the historical data available to them (Grewal et al. 2011).

Levy and his colleagues (2004) list some aspects that need to be taken into account when creating a pricing model in retail. Some of these factors are explicitly related to elasticity of demand, which is the main point of interest for this thesis. The elasticities are a product of statistical analysis of the data available to us. They depict the effect a price adjustment has on the demand of the items. This is called the price elasticity of demand, and we will observe both the items' own-price elasticities and cross-price elasticities in a fashion retail setting. Similar research has been conducted for example in the electricity market (Kirschen et al. 2000). More extensive list of other elasticity of demand research is provided by Dawit Mulugeta and his colleagues in their presentation from the SAS Global forum in 2013.

As always, the Company X strives to fulfill the customers' needs and to provide as good a customer experience as possible. Similar motivation is behind this project as well. We want to use the analysis conducted in this thesis to find elasticity of demand values for products and product groups. Further, we venture an optimization problem based on the computed elasticities, with a goal of finding optimal prices for some product groups. Additionally, we are interested in understanding and interpreting the elasticity values, especially in the sense of finding potential substitute and complementary products. The obtained results and possible future research suggestions are discussed in detail in the final sections of this thesis.

The statistical analysis, as well as the later computations are run with a program called Rstudio. The software enables R-programming, which is a well documented open source programming language, developed especially for statistics and data analytics. The R environment is easily extended by installing packages that can be downloaded from specific sites. All code written for this thesis is considered confidential and thus will not be presented.

2 Background and Theory

In a traditional sense, retail is understood as a system where retailers function as intermediaries between the producers and consumers (Niemeier, Zocchi & Catena 2013). They create additional value to both and take a profit off of the transaction for themselves. In essence, retailers help producers find the market for their goods by introducing them to consumers through their stores and online sites. Recently, many retailers have started to pay even more attention to creating additional value to the customer by enhancing the customer experience (Grewal et al. 2009; Verhoef et al. 2009). This is a mindset present in the Company X as well. Thus it might be more appropriate to view current retailers as agents creating and co-creating value to the customers through customer experience in addition to the plain transactions and fulfilling demand (Sorescu et al. 2011).

Some date the evolution of retail back to the historic time when people exchanged goods with each other, while others consider retail to have been born with the traveling salesmen of the Middle Ages (Niemeier et al. 2013). Following the latter interpretation, retail has evolved from a single salesman to multinational, internationally significant retail companies through industrial development and globalization. At present, consumers are provided with countless opportunities to research different products and compare prices from several retailers with little effort. Thus, setting competitive prices is one of the ways to stand out from the competition. However, baring in mind the nature of retailing, retailers must also strive to maintain their profitability in order to survive in the highly competitive market, which affects their pricing decisions greatly.

Since retailers are in essence service providers, bringing supply and demand together, they need to listen to and observe their customers to make good management and market decisions. Following the law of demand, pricing decisions are among the critical ones. Lower prices are expected to increase demand while profitability might decrease. On the other hand high prices would lead to a decrease in demand and might even steer consumers toward competition in hopes of lower prices. An important concept in this thesis, following similar logic, is that of price elasticity of demand, originally developed from marginal utility theory (Marshall 2009, referenced in Owen 2012 & Mulugeta et al. 2013). Price elasticity describes how changing the price of a product affects the demand for the same product, or others. The concepts of own-price and cross-price elasticities respectively, will be explained in more detail in this section of the thesis.

This section introduces a basis for our research. We start with a brief review of literature on retail, supply and demand as well as price elasticities. That is followed up by an introduction of a preceding research project, the results of which motivated our research. Finally, we bring in the mathematical expressions for the relevant concepts and give a brief explanation of the relevant optimization methods used in this project.

2.1 Market and Retail

As stated in the previous section, retail can be viewed as a system, where retailers function as intermediaries between producers and consumers. In a sense, retailers are agents connecting the market by bringing the demand to the suppliers. On the other hand they also facilitate a platform through which supply can reach the customers. Figure 1 illustrates this in a simple case of one retailer, three producers, a store network of five stores and a pool of consumers respective to each store. The alternative view on retailers can be explained with the same graph, with the addition of the services and additional value that retailers create and co-create for the customers in the processes linking different levels of the system.

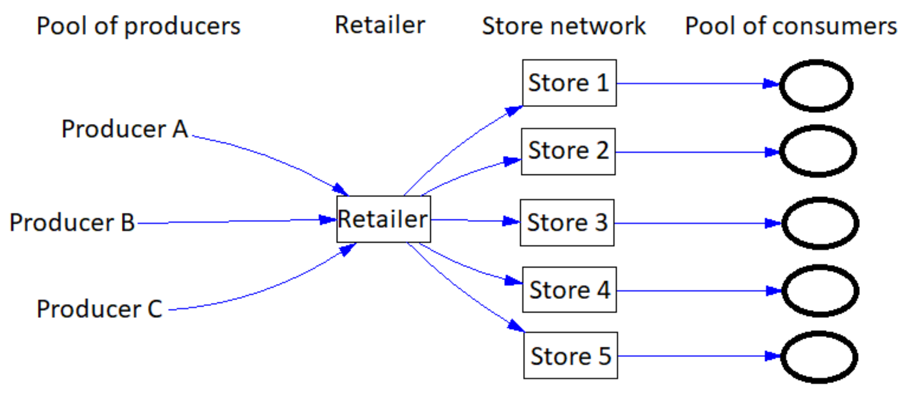

Figure 1: A flow chart depicting how a retailer works between the producer and a pool of consumers.

In a simple case as in figure 1, the retailer buys the products from a manufacturer and distributes the items to the local stores, from where the consumers can purchase a product fitting their needs. The challenge for fashion retailers is to determine a proper distribution of different product lines to stores in differing demographic and geographic areas to maximize the local sales. Furthermore, they need to be able to move the existing inventory to make room for new products in a fast changing fashion retail market. In order to win the consumer over, the retailers have the possibility to offer their customers promotional and other discounts and offers. An increasing amount of retailers have started to provide an online store option in order to reach the customers who prefer to search for information and make their purchases online instead of visiting an actual bricks-and-mortar store. Additionally, the present retail market has welcomed a significant amount of online retailers who are utilizing the power of Internet and computing to take the demand and supply online (Niemeier et al. 2013). With these recently introduced retail business models, the same retailer can provide the customer with several options to utilize for different purposes (Sorescu et al. 2011). We concentrate our effort on the more traditional retailers in this phase of the project and leave the online options to future research topics.

The retailers make their profit by taking a piece of the transaction between the consumer and the producer. In essence they subtract a commission from the ultimate sales price that the consumers pay for any given product. The challenge is to manage the transactions and win the consumers from competition. The more information is available to the consumers, the more they know and learn and thus become more aware of different options and prices available. The consumer behavior is affected by a great deal of other factors as well, but this thesis will not address them directly. We are more interested in the way prices affect the demand for goods. Furthermore, the level of quantity demanded, or ultimately the quantity sold, affects the retailer's revenue which is related to another point of interest in this thesis, finding optimal price suggestions for product groups.

2.1.1 Pricing

The demand created by consumers is the guiding force in the retail market. Customers determine the pool of products that sell. Of course this pool includes similar products, some of which are considered inferior while some are considered luxurious items. The demand for these products depends on several variables, such as the wealth and income level of the consumer demographic. Producers and retailers can try to affect the demand by pricing and moderating the supply. Additionally, since customers are exposed to different fashion trends through media, magazines and various public figures, the producers and retailers can affect such trend setting outlets through collaboration. Although the producer and retailer both aim to fulfill the demand created by the consumers, they also have a goal of selling their respective top line products which usually create more revenue and likely more profit in the sense of higher gross marginal. Consequently, the collaborations with trend setters and other fashion and style media exist. In the end however, the consumer is the force that creates and determines which products are of higher demand and both producers and retailers need to listen to and react to their wishes.

In addition to price, there are other underlying reasons that can cause fluctuations in products' demand, which include for example aforementioned changing fashion trends, brand preferences and seasonality. The first of these is something that the retailers and manufacturers could try to affect as mentioned previously. Opposite can be said of the seasonal demand fluctuations that can be caused by the environmental changes due to the time of year. For instance, the demand for snow boots increases after the first snowfall while the demand for skateboards might drop simultaneously. On the other hand some products, such as T-shirts, are not affected as strongly by such changes. Furthermore, seasonal effects can be seen for instance around Christmas and Valentine's day. The retail industry has many international companies that need to take into account the global differences in addition to the seasonal ones. For instance, the demand for some product groups is probably different in the northern Finland when compared to that in Spain due to very different environmental and seasonal aspects.

There are such factors affecting demand, that are not necessarily dependent on pricing or seasonality. One such factor is the consumer's brand preference. For instance, the demand in a product group of specific type of trousers can be significantly emphasized on one option regardless of the pricing in that product group, due to that specific model or brand being the more preferred one. Additionally, some brands might launch exclusive and highly limited product collections that, for example, have been designed in collaboration with some famous fashion designer or a well known public figure. Such special collections might experience high demand regardless of their price being significantly higher than any similar product from the basic collections. These kinds of preferences are individual to the customer and thus difficult to model and will be left outside the scope in this thesis.

When the demand and supply are in balance, in mathematical sense when their functions intersect, we have reached a market equilibrium. In the equilibrium the supply equals the demand exactly. The market sets some level of demand and the suppliers need to react to that demand by providing products for some price. When the supply and demand meet, we have an equilibrium price. This relation is presented in figure 2. Deviating from this price might give a boost to a retailers sales due to consumers choosing the lower price. This kind of consumer behavior will motivate a quick reaction from the competing retailers and they will more than likely give the same or a substitute product a markdown to neutralize that competitive advantage. Thus the equilibrium will shift accordingly. However, some retailers use permanent markdowns in order to tempt consumers with lowered prices and thus increase the sell-through on respective products. With an unlimited supply, this habit would not be profitable, but retailers rarely have a supplier with endless inventory to offer them more products. Thus the permanent markdowns usually last until the products in question are sold out of stock. These permanent markdowns could be a consequence of a trade deal offered to the retailer by a manufacturer (Hall, Kopalle & Krishna 2010). In such a deal the retailer could have an opportunity to purchase a large quantity of some product for a lower than normal price. Consequently, the sales price could be subject to a permanent markdown.

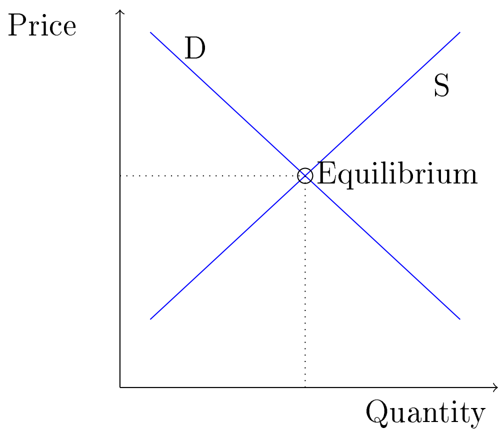

Figure 2: An illustration of the market equilibrium with demand and supply functions.

As figure 2 shows, the market equilibrium and the respective equilibrium price is found when the demand and supply functions intersect. Deviating from this price causes instability on the market since one retailer has shifted their supply function toward right in the graph. The shift is due to the retailer offering the similar quantity of products for a lower price. Said deviation would also shift the intersect to lower equilibrium price, which would yield higher quantity demanded. This could be, for example, due to consumers choosing this retailer instead of the competition or attracting the more cost conscious buyers with the lowered prices. At a price lower than equilibrium, the demand exceeds supply, which leads to products selling out. On the other hand prices higher than equilibrium cause a supply surplus due to low demand.

The market equilibrium can shift for several reasons, one of which was explained above. Any changes in consumer behavior and price sensitivity, as well as supplier or retailer driven changes can cause the equilibrium to shift. The retailers could gain valuable insight to the market from research and utilizing data available to them. Especially in a competitive market, such as fashion retail, it is important for retailers to give thought to pricing. Through pricing it could be possible for retailers to increase their sales as well as gain consumers from the competition. However, pricing too low could eat away the profitability and lead the retailer to problems in the market. On the other hand, pricing lower in order to gain market share and increase future traffic on the expense of current profits could be a good strategy since consumers are becoming ever more sensitive to price (Fox, Postrel & Semple 2009). This approach would require the retailer to be in such an economic state that it can sacrifice short term profits for long term gains. As the market theory and the equilibrium suggest, the market is interesting for new entrepreneurs and companies when prices are forecast to increase, while many are tempted to leave the market when prices are decreasing. Thus driving the prices lower could motivate some of the competition to abandon the market (Levy et al. 2004) and in doing so to increase the retailers market share and future traffic.

Pricing does have an effect on consumer behavior and thus affects the buying decisions. Lower prices are often connected with a perception of low quality, while higher prices usually reflect higher quality (Dodds, Monroe & Grewal 1991; Levy et al. 2004) in the minds of consumers. On the other hand, prices lower than customers' preferred price range would yield an increasing net value for the product while prices higher than the preferred price range would yield a decreasing net value. Therefore, in consumers mind the perceived quality and the gained value form a tradeoff which affects their buying behavior. For example if the price is too high, the customer feels that there is no net value available by purchasing the product, since the sacrifice would be too great, albeit for a quality product (Dodds et al. 1991). Finding the acceptable price range and creating additional value through customer experience have the potential to affect such tradeoff situations to the mutual benefit of the customer and retailer alike.

The market in retail is a fast-changing competitive environment and retailers need to be able to react quickly to changes in the market and competition. Some companies have come up with specialized dynamic pricing models, that utilize market data and their own transactional data to set optimal prices more frequently according to the present or forecast demand (Grewal et al. 2011; Hall et al. 2010). These pricing models have become more frequently used since the amount of data available has grown significantly and the sense of urgency in the market has increased. Additionally, dynamic pricing models enable small scale price discrimination, which could even mean price personalization to individual customers (Grewal et al. 2011). Such opportunities are interesting to retailers since finding the optimal price in the market could maximize their revenue. Which in turn could facilitate larger profits. Cross, Higbie and Cross (2011) wrote that revenue maximization, rather than cost minimization has lead to promising results on various fields, such as aviation, travel, delivery and retail. They talk about revenue management and price optimization using demand forecasts, price elasticities and competitive rates. Their examples of historically significant companies surviving severe difficulties in large part due to revenue management, is further motivation to the importance of analytical pricing.

Now that we have accomplished an idea of the importance and effects of pricing, we will introduce the concepts of price elasticity of demand in more detail. These concepts carry a significant meaning for our research, the main point of which is to find own-price and cross-price elasticity values and interpret the relationships between products within a specific product group. The following section will include some interpretation models of these elasticities as well.

2.1.2 Elasticities of demand

The concept of elasticity has been mentioned on several occasions in this thesis. This section gives a rather economist style explanation to the two types of elasticities this thesis addresses: own-price elasticity of demand and cross-price elasticity of demand. We begin with own-price elasticity and extend that into cross-price elasticity. In this section, we give the basic explanation to the concepts with easy to read equations, the better mathematical expressions of which will be introduced later on in the following sections.

Basically, these elasticity values measure the extent to which the demand for a product is affected by the change in it's price. The concept was originally derived from Marshall's (2009) work, the first edition of which was published in 1890. The applications have since been developed further and for instance, many dynamic pricing models (Grewal et al. 2011) and market analysts use (Cross et al. 2011) and should use (Levy et al. 2004) these concepts when determining proper pricing for their products. The elasticities have been applied in various fields of study, including groceries (Genchev & Yankova 2010; Andreyeva, Long & Brownell 2010), water (Espey, Espey & Shaw 1997; Schoengold, Sunding & Moreno 2006), oil (Cooper 2003), gasoline (Hughes, Knittel & Sperling 2006), electricity markets (Kirschen et al. 2000; Thimmapuram & Kim 2013), housing (Hanushek & Quigley 1980; Ermisch, Findlay & Gibb 1996), and aviation and travel (Cross et al. 2011).

The own-price elasticity value is calculated by comparing the percentage change in the product's demand with the percentage change in the product's price. The relation is presented in the following equation (1). The sign of the elasticity value indicates the direction in which the changes occur. In case we want to only observe the extent of elasticity, it is enough to look at the absolute value of elasticity. An absolute elasticity value below 1 indicates that the product's demand is inelastic for that price change. For example, having an elasticity value of 0.2 would indicate that for a 10% change in price, the response in quantity demanded would be a change of 2%. Elasticity values above 1 indicate elastic demand, meaning that the change in price causes a larger change in quantity demanded. For example an elasticity value of 1.5 would lead to 15% change in quantity demanded when the price is changed by 10%. Elasticity value of exactly 1 is called unitary elasticity, which means that the percentage changes in price lead to equal percentage change in demanded quantity.

$$Own - price\ elasticity \ = \frac{\%\ Change\ in\ quantity\ demanded}{\%\ Change\ in\ price} \tag{1}$$

An extension to the concept presented in the equation (1) is that of cross-price elasticity of demand. This elasticity value measures the extent to which the demand for product A is affected by the change in price of product B. The magnitude of the cross-price elasticity value has a similar interpretation to that of own-price elasticity. The relation that yields the cross-price elasticity values is presented in equation (2) below.

$$Cross\text{-}price\, elasticity\ (B_A) = \frac{\%\ Change\ in\ quantity\ demanded\ (A)}{\%\ Change\ in\ price\ (B)} \quad (2)$$

The above relation is best understood through examples. Let's assume we have two relatively similar pairs of jeans in price, looks and material. The choice between the two is assumed to not be affected by brand or other preference. Now, if we were to lower the price on one pair of jeans (B), the rational consumer would choose to purchase that pair of jeans (B) due to lower price. This could possibly mean that the increased demand for the discounted jeans (B) leads to a decrease in demand for the other pair of jeans (A). In essence, lowering the price affected the preference of the consumer. Such a dynamic could be described with a positive cross-price elasticity, due to the changes in price for jeans (B) and demand for jeans (A) being both decreasing. A positive cross-price elasticity would also imply that the two products are substitutes, whereas a negative cross-price elasticity would suggest the products to be complements. For example, lowering the price of running shoes could lead to an increase of demand for running socks in addition to the shoes themselves. Finding such relationships is important to retailers, since that could provide more insight into consumer preferences, facilitate more accurate forecasting of demand in the presence of discounts and provide help with product bundling. Additionally, uncovering some unexpected complements could assist retailers in upselling, which means recommending additional products to those a customer has decided to purchase and thus increasing the level of customer experience.

Pricing in general is important for retailers as mentioned earlier. Understanding the effect prices and price adjustments have on demand is equally important in order to gain maximum revenue in a highly competitive market. The advantages range from optimal pricing strategies to discounting, product bundling and marketing (Mulugeta et al. 2013). Levy and his colleagues (2004) noted that different price optimization systems have a common aspect of analyzing price elasticities while being otherwise very different. Due to having access to large amounts of data and highly sophisticated software tools, it is important for retailers to take advantage of these opportunities and to enhance their functions and thus profitability. This project is one of the first steps for Company X to take advantage of the transactional data through price analytics. Company X is hoping to gain additional insight into pricing and the relationships between products across various product lines and groups.

It needs to be noted that the elasticity value depends on the price change and is not a universal value for the respective product across all prices. Furthermore, the elasticity values vary over time, since there are other factors affecting demand in addition to price. Thus changing the observation period might alter the resulting elasticity values significantly. This notion is taken into account in the following analysis by determining the interesting time periods for each product analyzed. Due to price and time sensitivity, the elasticity values are only regarded as estimates. In our case, we will be using weighted means of products' prices to calculate the elasticity estimates. We will also compute the values using different levels of aggregation in the data. The results are described and discussed in the later sections.

2.2 Basis

This thesis is based on an existing research conducted by another Master's student in collaboration with the Company X. That research was a case study that includes a group of stores and a large selection of test products. The aim was to investigate if the demand for the selected products was affected by changes in their respective prices. During the test period, the participating stores were divided into groups that made price adjustments according to the directions provided by the project team. The transactional data was then collected and refined in order to analyze it further. The analysis suggested that most of the test products, regardless of the product group, demonstrated elastic demand behavior. Our project is an extension to the previous research and aims to investigate similar relationships and interactions within product groups instead of individual items.

The previous research is based on the assumption that the price - demand relationship can be modeled with a linear regression model. This assumption has some theoretical basis, since linear regression gives a good approximation on a small price range and it is easier to extend to a longer time as well. This approach was used for example by Owen (2012), Hall, Kopalle and Krishna (2010) and Mulugeta and his colleagues (2013). Furthermore, it is quite easy to add variables to a linear regression model. In this sense linear regression is more efficient than for example mid-point formula when estimating elasticities, as pointed out by for example Dawit Mulugeta and his colleagues in their presentation in SAS Global Forum (2013). However, we realize that the price - demand model is rarely linear, which means the results are suggestive at best. Furthermore, elasticity values obtained from the analysis are approximations, not exact values. Consequently any results and pricing solutions yielded by our analysis are not absolute truths but only give suggestions on possibly better pricing decisions.

Using linear regression to model the alterations in products demand when facing changes in price, yielded results where most test products showed a downward slope, which would suggest the products to have negative own-price elasticity. An example of such graphs is presented in the figure 3. A negative slope is resulted in by increasing demand for a respective product, when the sales price is lowered. The changes are of opposite direction, which causes the slope to be negative and follows the theoretical model of the law of demand. In the figure 3, the vertical axis represents average daily demanded quantities while the horizontal axis has the different sales prices.

Figure 3 presents a product with negative own-price elasticity. Many ordinary products have a negative elasticity value, indicating that demand for such product decreases when the price increases. If the elasticity value is strongly negative (< -1) the product is likely to have substitute products (Genchev & Yarkova, 2010). Although, the substitute products are better revealed through cross-elasticity research. Products that have an inelastic demand would have a horizontal line instead of a slope. That would indicate that the product is sold with similar quantities regardless of the price. Such products can be considered absolute necessity products that do not have significant substitutes. Additionally, the slope can also be positive, which would suggest that the respective product experiences higher demand when the price is increased. There are several reasons why and examples of such goods that do not follow the law of demand, such as Giffen (Nachbar, 1998) goods.

The previously obtained results hold a strong suggestion about the own-price elasticity of the observed products. The results have motivated an extension from individual products to entire product groups. It was chosen that we would form product groups instead of for example observing different brands separately, due to intuition and proof of better efficiency of such practice (Hall et al. 2010) as well as better indication of the relationship between competing products (Levy et al. 2004). Within this thesis, we are interested

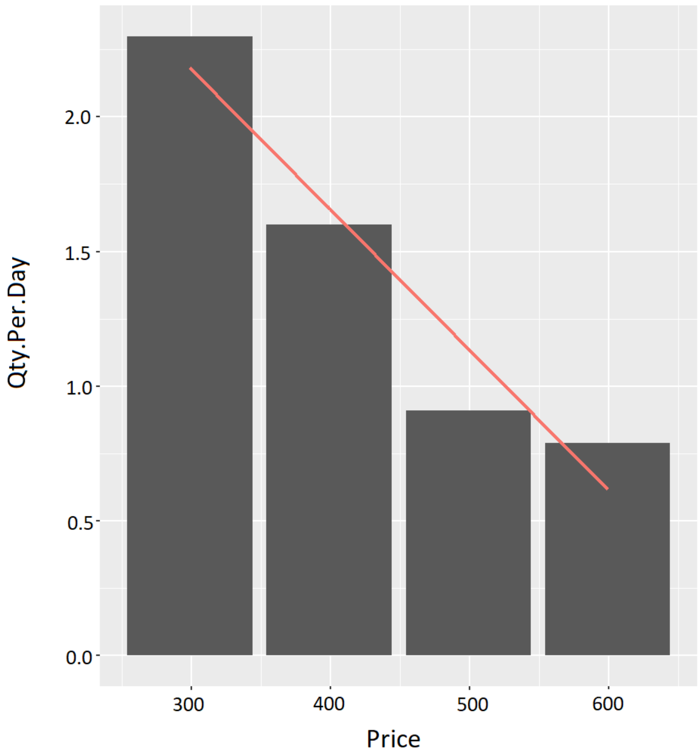

Figure 3: An example of the original graphs suggesting own-price elasticities.

in enhancing the previous computations to fix some issues with the data analytics, as well as extending the elasticity research to product groups. Additionally, we use the cross-price elasticity values in price optimization. This final phase is interesting, since any advantage or extra insight is welcome in the highly competitive retail business.

2.3 Mathematics

In order to understand the process of finding answers to our research questions, the following chapter introduces the relevant mathematical expressions used in this research. We assume the reader to be familiar with the mathematical concept of linear regression and the basic applications thereof. The general presentations of simple and multiple regression models are presented in the appendices in case revision is needed. We will have a brief look at correlation analysis and implications thereof, followed by the mathematical expression of the concept of elasticity, both own-price elasticity and cross-price elasticity. Finally, we will present the relevant optimization methods utilized in this project.

Since the reader has been assumed to be familiar with the concept of linear regression, we will start with correlation analysis, which is related to linear regression. By computing a sample correlation coefficient for two variables, we can find out the level of linear association between the two. The correlation coefficient r gets values $-1 \leq r \leq 1$. The absolute values of r approaching 1 indicate stronger association, whereas those closer to 0 suggest low levels of association. In essence, regression analysis is similar in the sense of observing variables association by fitting a model into a sample of observations. The correlation analysis is not the main focus in this thesis, but it will be used when choosing the proper product groups and specifically when excluding some items from a product group due to low correlation between it's price and the demand of others. The calculation formula for correlation between two variables, dubbed "x" and "y", is

$$r = \frac{Cov(x,y)}{\sqrt{s_x^2 * s_y^2}}, \tag{3}$$

where s_x^2 and s_y^2 are the respective sample variances for x and y. The sample covariance is

$$Cov(x,y) = \frac{\Sigma(x - \bar{x})(y - \bar{y})}{n - 1}, \tag{4}$$

where $\bar{x}$ and $\bar{y}$ indicate sample means, and n stands for sample size. The sum in the numerator includes a product of the differences between each pair of observations and the respective sample means. Applying the equations (3) and (4) in the later phases of this project allowed us to choose the relevant products into the same product group for computations.

Own-price elasticity of demand is similar to correlation between the products price and it's demand. The stronger resemblance and our chosen path to compute the own-price elasticity values is through regression analysis. We fitted a linear regression model to our sample and determined the slope of the obtained model. The slope in turn is needed to determine the own-price elasticity value for the product in question. The following equation presents the formula for own-price elasticity,

$$\varepsilon_A = \frac{\%\Delta Q_A}{\%\Delta P_A} = \frac{\Delta Q_A}{\Delta P_A} * \frac{P_A}{Q_A}. \tag{5}$$

The latter form of the above equation (5) includes a familiar notation from regression analysis. The term $\Delta Q_A/\Delta P_A$, is the formula to calculate a slope of a linear regression model. Since price P_A and quantity Q_A are always positive values in our case, the slope determines whether the elasticity value is positive or negative. By using the slope and choosing proper prices and demand levels for the computation, we are able to obtain the own-price elasticity values respective to those prices and quantities. As noted earlier, the elasticity values are price and time sensitive and thus choosing the price and quantity levels carefully is important and the elasticity values will change with the price and quantity levels chosen to the calculation. The elasticity values describe how the demand is affected by price changes on that exact price point.

Extending the own-price elasticity to span an entire product group yields the concept of cross-price elasticity. Now we observe other products' prices' effect on the demand of another. We need to use multiple regression instead of simple linear regression. Multiple regression model is explained briefly in the appendices. Now the slope is replaced by the variable specific coefficients from the multiple regression model (denoted by b_i in the appendix) when determining the cross-price elasticity values. The calculation formula for the cross-price elasticity values is,

$$\varepsilon_{B_x} = \frac{\%\Delta Q_A}{\%\Delta P_B} = \frac{\Delta Q_A}{\Delta P_B} * \frac{P_B}{Q_A}. \tag{6}$$

The main difference between the computation of own-price and cross-price elasticities lies in the inclusion of other products in the regression model for cross-price elasticity. The computation becomes more complex with more products and the regression model is more difficult to fit on the data. The difference can be seen from the calculation formulas as well. The formula for cross-price elasticity in equation (6) has price change $\%\Delta P_B$ instead of

$\%\Delta P_A$, which indicates that the price change affecting the value of cross-price elasticity of product A is of the price of product B. The notation on the equation (6) is by design ε_{B_x}, which means that the value is the cross-price effect of product B on the demand of product A.

We are interested in the cross-price elasticity values of entire product groups and want to have these values available for interpretation in a form that is easy to read. Consequently, we store the values in a matrix, each row of which includes the cross-price elasticity values for the respective item pairs. The diagonal of the matrix represents the own-price elasticity values yielded by the multiple regression model. For the optimization phase this matrix is filled with the multiple regression coefficients instead of cross-price elasticity values and needs to be transposed. The matrix was deemed easier to read and understand when the elasticity values were stored this way.

The matrix that is filled with the multiple regression coefficients is carried on to the optimization phase. This part of the computations is explained in more detail in the following section.

2.3.1 Optimization

The cross-price elasticities, and more precisely the multiple regression models are utilized in the optimization phase of the project. While the elasticity values are extracted by utilizing the coefficients in the regression model, the coefficients themselves are used in the objective functions in the optimization. This section presents the optimization methods used in this project. We used several different methods in the optimization phase to test multiple approaches and algorithms in order to compare the results. The optimization methods were then applied in Rstudio environment and proper options were selected for the in-built optimization tools to make the algorithms work as wanted. The following will introduce the methods one by one.

Nonlinear conjugate gradient method

The conjugate gradient method is usually used when solving large problems that would otherwise require large amounts of storage memory. Such problems could prove problematic for example when implementing quasi-Newton methods (Kelley 1999). The conjugate gradient methods have, however, demonstrated some reliability issues when the initial conditions and the initial guess are not carefully specified.

For a quadratic optimization problem, the conjugate gradient method utilizes a linear combination of the residual r, which is not to be confused with the correlation coefficient from the previous section, and the search direction d in the iteration update. Since the search direction is itself a linear combination of the previous residuals, the only storage needed is for the residuals and search directions. The nonlinear method is an extension to the linear conjugate gradient method. The linear method minimizes

$$f(x) = \frac{x^T H x}{2} - x^T b, \tag{7}$$

with the residual being simply the negative of the gradient:

$$r = b - Hx = -\nabla f(x). \tag{8}$$

The extension to nonlinear problems becomes such that the initial state is $r_0 = d_0 = \nabla f(x_0)$, and the iteration steps

$$r_k = \nabla f(x_k), \tag{9}$$
$$d_k = r_k + \beta_k d_{k-1}. \tag{10}$$

The update of the decision variable x can be done by an analytic minimization in a quadratic problem, while a nonlinear problem requires a line search. The update for the decision variable x becomes

$$x_{k+1} = x_k + \alpha_k d_k. \tag{11}$$

In equation (11), the α_k denotes the step size in the iteration, which will terminate when the change in the function value decreases below a set termination parameter t. The final part of the initial specifications is to choose a proper value for the constant β_k. There are a few common choices which will not be discussed here. For example, Kelley (1999) presents two of those common options.

We will implement the conjugate gradient method on our quadratic problem and compare the results with other methods that utilize a different approach on finding an optimum.

Nelder-Mead method

The following method of finding an optimum is called the Nelder-Mead method. This is usually a reliable option, although it might be slower than other, more developed methods. Nelder-Mead simplex algorithm does not need any gradients, but evaluates the function itself. In order to find a minimum for a function of n variables, we need to construct a simplex of $n+1$ vertices. The function is evaluated in each vertex and the function values are sorted accordingly,

$$f(x_1) \ \leq \ f(x_2) \ \leq \ \cdots \ \leq \ f(x_{N+1}). \tag{12}$$

The vertex x_1 gives the "best" objective function value and x_{N+1} gives the "worst" one. The algorithm replaces the "worst" vertex with a new one, which is of the form:

$$x(\mu) \ = \ (1 \ \mu)\overline{x} \ - \ \mu x_{N+1}, \tag{13}$$

where $\overline{x}$ is the centroid of a convex hull of $\{x_i\}_{i=1}^{N}$,

$$\overline{x} \ = \ \frac{1}{N}\sum_{i=1}^{N} x_i. \tag{14}$$

Thus, the simplex is updated by replacing the vertex that yields the "worst" function value with a new one. The new vertex should yield a function value at least better than the previously removed vertex. For simplicity, we give an example for a case with two variables, which would have a simplex of three vertices. The first step of such simple case is presented in figure 4 below. Further and more accurate information on the Nelder-Mead algorithm can be found from Kelley's (1999) work on iterative optimization methods.

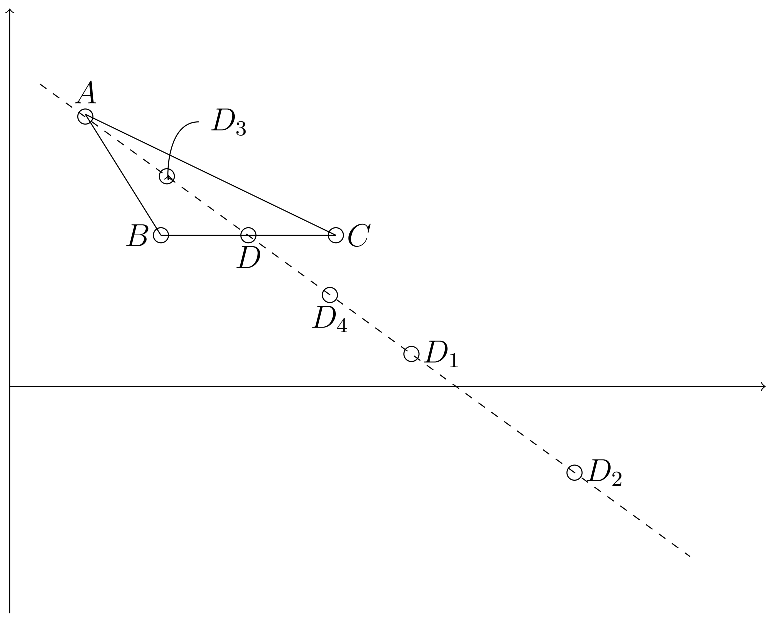

Figure 4: An illustration on the first step of a Nelder-Mead simplex algorithm for two variables.

Figure 4 presents a simplex with three vertices. The objective function would thus be of two variables. The algorithm evaluates the objective function in each of the original vertices A, B and C. The vertix that yields the "worst" value needs to be removed and replaced. Lets assume vertix A yields the worst value. The candidates for a new vertix lie on the dashed line drawn from A through the centroid of B and C, which is denoted by D. Now the function value is evaluated in vertix D_1. If the value here is better than the best value from A, B and C, then this line is deemed a good direction of descent. Next the function value is estimated in D_2. If this vertix provides an even better value than D_1, then A is replaced with D_2, otherwise with D_1. On the other hand, if D_1 gives a better function value than A, but not the best, then A should be replaced by D_4. In case D_1 is worse than A, we choose D_3 to replace A. We end up with a new simplex, which will be subject to a similar procedure. This iteration is continued until we have found a good enough approximation of an optimum. This method was chosen to be used as a reference for it's reliability and simplicity, even though it does not necessarily converge in every case.

Quasi-Newton methods

The more important group of optimization methods for our purposes is that of quasi-Newton optimization methods. These methods use the second derivatives of the objective function, namely $\nabla^2 f(x^*)$, by updating it's approximation during the iteration. Generally, the algorithms determine the updated approximations by first finding a feasible direction for a line search procedure to find an updated decision variable. This is then used to update the Hessian matrix. The general update iteration for the approximates is presented below. Different quasi-Newton methods are set apart by the manner in which the Hessian is updated (Step 3 in equation group (15)):

$$\begin{aligned}
&1: \text{ Determine feasible direction } d = -H_-^{-1}\,\nabla^2 f(x_-);\\
&2: \text{ Update the decision variable } x_+ = x_- + \lambda d \text{ by line search;}\\
&3: \text{ Update the Hessian to } H_+ \text{ by using } x_-,\ x_+, \text{and } H_-.
\end{aligned} \tag{15}$$

The method we introduce is called the BFGS method. The abbreviation stands for Broyden, Fletcher, Goldfarb and Shanno method (Kelley 1999). This method, among others grouped in quasi-Newton methods, satisfies the secant equation and can thus be characterized as a secant method. The BFGS method requires a more complex Hessian update, than that of a standard nonlinear equation which only requires rank-one update. The BFGS method needs a rank-two update, which is an extended version of the standard Broyden's method. This update method is,

$$H_+ = H_- + \frac{yy^T}{y^T s} - \frac{(H_- s)(H_- s)^T}{s^T H_- s}, \tag{16}$$

where $s = x_+ - x_-$, and $y = \nabla f(x_+) - \nabla f(x_-)$. In order to reach the best performance from this algorithm, some strong assumptions need to be made of the objective function and the initial iterate value. Given that the set $D = \{x| \; f(x) \leq f(x_0)\}$ is convex, the objective function f is Lipschitz twice continuously differentiable in D and that there are $\lambda_+ \geq \lambda_- > 0$, that satisfy $\sigma(\nabla^2 f(x)) \subset [\lambda_-, \lambda_+]$, $\forall x \in D$, we can reach superlinear convergence to x^*. These assumptions would also imply x^* to be a unique minimizer to f. (Kelley 1999)

Our research requires constraints on the decision variables x_i. For this reason we need to extend the standard BFGS method to allow bound constraints. The resulting method, with an addition of limited memory usage, becomes the L-BFGS-B method, which stands for limited memory BFGS method with box (bound) constraints. The constraints can be of form $l_i \leq x_i \leq u_i$. This algorithm utilizes a simple gradient method to determine free variables on each iteration and then performs the limited memory BFGS method on those variables to increase the accuracy of the result. Limited memory BFGS differs from the standard BFGS in that it does not store all iteration steps, but "forgets" the oldest step on each iteration. That does not necessarily lead to better efficiency however.

3 Research

This section introduces the project, which has lead to this thesis. We start by explaining in detail the data available and where it is from. Then we introduce the model we have created and go through it step by step to give the reader a good understanding on each part of the model.

The project was started together with the Company X's business development department. It is an extension to a previous research that yielded some promising results. This project however, was originally aimed at creating large tables filled with cross-price elasticity coefficients. The plan was to then use these coefficients to calculate optimal prices for the items through maximization of, for example, revenue or profit. During the project, it was noticed that such large tables were not a proper way to store the elasticity coefficients. Rather, the coefficients are easier and, more importantly, more accurately stored in smaller tables. These tables would include only one product group respectively.

The model we created is a simple multivariate model that is dependent on items' prices. The multivariate elasticity factor is built in the model as a matrix coefficient. The other factors are product specific coefficients derived from the raw data extracted for the purposes of this research. The following subsections will explain these constant factors in more detail.

In order to conduct realistic optimizations, we need to form constraints to our model, that is in fact the objective function we want to optimize. This section is concluded with the explanation of the different constraints and why they are needed to obtain realistic results. The constraints were built in the model as the testing progressed, since changing the data input unveiled new issues rather frequently. Many of these issues required further analysis of the results in order to figure out how to develop our scripts.

3.1 Data

We started out this project with similar data to that of the earlier research. The transactional data was gathered with similar variables. Instead of creating an identical test scenario, we mined the transactional data from a span of several years, including dozens of store locations and thousands of products. Additionally, we extracted several tables which were helpful in refining the transactional data further and when creating the model variables and constraints. The final addition to our data came from the cross-price elasticity computations, since the cross-price coefficients were used as a factor in the objective function in the optimization phase. These cross-price coefficients will be covered in more detail in the later sections.

3.1.1 Transactional data

The most important data for this research is the transactional data collected from the stores. This data includes every sold item and service from the chosen time span. In essence, this data shows us each sales transaction and includes information about different prices, promotions and product categories. The transactional data is in its' raw form, meaning that it is completely unrefined and unusable as it is.

The transactional data is stored in large tables with which it is possible to select the included variables. Some of these tables were extracted by targeted SQL-queries from the database and saved as tab delimited text files, while others were extracted as power pivot tables. These tables include many variables that are not used in this phase of the project, but carry an interest for the future analyses. The largest of the transactional data tables includes more than five million rows of observations.

The most important variables for this project are different sales prices, dates, quantities and product IDs. We need the product IDs (including article numbers and names) to single out different products for analysis. Further, the product IDs are used on the visual presentation of the results, not however in this thesis due to confidentiality issues. The dates determine time periods, during which the products have been sold and with them it is possible to calculate how frequently different products have been sold. Finally, the sales prices and quantities are needed for our analysis of own- and cross-price elasticities. They are essential in the regression model that was created to compute the elasticity coefficients.

The transactional data includes other variables as well. These variables include for example product category, store numbers, receipt ID and promotion category. These variables can be used in more detailed analyses, such as product category or store specific elasticities and frequent product combinations to reveal complementary products. Additionally, buy-in prices were included for optimization purposes. These variables are needed to form some of the constraints and product specific factors for our model and optimization problem. The meaning of some of these variables is explained in more detail in the following section.

For the purposes of this project, it was deemed necessary to refine the transactional data further to ease the computations. This of course causes the results to differ from reality to some extent. The more accurate modeling was left for the next phase of this project, which would include the addition of new tools and software, and most importantly, more computational power. Thus we would be better equipped to handle the unrefined transactional data. The refinement was deemed necessary, since some items were found to have been sold for a large amount of different prices during a short time span. This is due to overlapping promotions and the manner in which the prices are stored in the transactional data tables.

To conclude, the transactional data is the most important and the most leveraged set of input for our analysis. The data cannot be used as it is though. Instead it was necessary to refine it further. For that purpose, several other tables were extracted, the function of which is explained in the following chapters.

3.1.2 Additional data dimensions

This section introduces the remaining data tables used in our analyses and data refinement. These tables include different product and organization specific dimensional data and data regarding prices and promotions. Some of the data is left to the following phases in the project and thus not regarded in the following calculations.

Firstly, the price dimension table includes definitions for different price codes. These price codes can be specific to, for example, a single organization, promotion or loyalty program. This table was referenced when refining the transactional data. This dimension was used to account for some of the odd prices. It provides us the valid price categories which can be cross referenced with other dimensions to obtain a table of valid prices with the respective dates. This table is then used when refining the transactional data.

The additional dimensions include organization, period and product dimensions as well. The first of which lists different organizational departments and store numbers among other information. This table is not relevant in this research but could be interesting for comparison between different units within the Company X, in the fashion of the research conducted by Stephen J. Hoch and his colleagues (1995). Similarly, the periodic data includes information about different sales periods, quarters and their actual date intervals. These tables carry an interest for seasonal analyses, but they are not our main interest in this research. These tables are referenced to determine some date intervals and for instance, the termination of sales for some products due to seasonality.

The largest dimensional table is of the product dimension. This table includes product specific information, such as different prices, item groups, size and weight. Some of these attributes are irrelevant for our research. However, this table is referenced to determine interesting product groups and to extract additional information on products' pricing. Although, most of the needed price data is extracted from price adjustment tables. These tables show the product specific prices, validity dates and price categories. Each row represents an individual price adjustment for one product. This data was utilized together with the price dimension table when refining the transactional data and to recognize some of the so called odd prices that do not belong to the valid price categories.

Finally, a table with information of the product stock levels was extracted. This table was referenced to determine the date when a product ran out of stock. We use this information to cut the observation dates when the item is no longer available. Thus we are able to avoid some distortion in the computations, best visible in the own-price elasticity graphs. In addition, the effects of this refinement will carry on to the later cross-price elasticity computations and thus the optimization as well.

3.2 Model

The collected data is utilized to create a model of quantity demanded. In order to keep this phase of the project simple and easy to build upon, a linear regression model based on the transactional data was chosen as the basis for our model. Further motivation on this choice is given in the preceding sections. For the optimization phase, we build on the multiple regression model in order to get the desired objective function. These objective functions and their constraints are presented later in this section. We will cover the most common constraints that apply to most product groups.

Before we can start creating the regression model though, we need to refine the data as mentioned. First of all, it was decided to exclude all items with less than 20 units of sales during the time span of observation. This decision was made to avoid having too few data points for our regression to work properly. Similar decision was made, when excluding products that had only been sold for one price. Having only one price would mean that the product is irrelevant to other products demand in the cross-price elasticity computations, since it has no price adjustments. Additionally, it was noticed that some products and services have a default price of 0. These items and services were excluded as well, since their pricing follows different methods than the regular products.

Having refined the data, we can start computing the model for quantity demanded, which is presented below. The quantity demanded of a single item is:

$$q_i(p_i) = p_i \ * \ w_i \ + \ c_i, \tag{17}$$

where the demanded quantity q_i, is a function of the price p_i. The slope of the linear regression model is denoted with w_i, which can be used further to compute the own-price elasticity value for the respective item. The equation (5) shows how we calculate these own-price elasticity values. The final value c_i denotes a constant that represents an intersection with the y-axis. In essence, it tells us how high the demand would theoretically be if the items were given out for free, which of course does not happen.

The own-price elasticity values are calculated by using the slope w_i to replace the first fraction term on the final form of equation (5). The remaining terms are the means of price and quantity sold per day, calculated for the chosen time period. The issue with these values is, that they are only computed for a single product and the model does not include prices of other products or their demand. Including these factors in the model requires a more complex model, in which we include the prices of other products as well. This gives us a new model for the demand, dependent on the prices of several products. The new model looks similar to that in equation (17), but all the factors are vectors instead of scalars. The presented model is given by:

$$q_i(P) \; = W \; * \; P \; + \; c_i. \tag{18}$$

The multivariate model can be used to compute cross-price elasticity values in a similar fashion to the previous example of own-price elasticity. Now we extract the weight vector W of dimension $[1 \times n]$ of model coefficients and enter it to the equation (6) to replace the first term in the final form of the equation. The corresponding factor from W needs to be selected to go into the equation (6) in order to get the desired cross-price elasticity value. For our purposes, it was more suitable to form a vector valued function, which includes a matrix coefficient. This extension means that for equation (18), q_i becomes a vector of demanded quantities Q for all interesting products, P is still the price vector, c_i becomes an intersection vector c that includes all respective intersects from individual regression models, and W becomes an $[n \times n]$ matrix coefficient. Now the W-matrix includes the weights that describe how the demand for one product is affected by the prices of others. The equation (19) below takes into account the dimensional rules of matrix calculus and returns a $[n \times 1]$ vector of demanded quantities:

$$Q(P) \; = W \; * \; P \; + \; c. \tag{19}$$

The created multiple regression model is then used in the objective functions for our optimization problem. Furthermore, we can compute the desired cross-price elasticity values from the created regression models. The values are stored in matrices, where a cell $[i, j]$ would represent the relation between the demanded quantity for product i and the price of product j. The obtained results will be addressed in the following section.

We can choose between several different objective functions, depending on the value we want to optimize. The two most interesting values are profit and revenue. The objective functions become very simple when using the previously computed demand model. The equation for revenue is:

$$R \; = \; P^T \; * \; Q(P), \tag{20}$$

where $Q(P)$ is the demand model from equation (19). The profit function needs an additional constant factor, that includes the buy-in prices for the products in question. Including the buy-in prices requires us to take into account the taxation of sold products. The tax is added into the profit function as a product specific constant variable. Some products and services have a different tax included in their price, thus we need to form a variable that gets a different value depending on the product and product group in question. The tax coefficient is formed as a constant $0 \leq t_i \leq 1$, where $t_i = 1 - \frac{tax\%}{100\%}$.

The objective function, in the simplest case, is that of equation (20). We want to maximize the revenue, which means maximizing the flow of money into the stores. In essence, we want to find prices, that maximize the quantity demanded, but still satisfy the constraints defined for the objective function. Unconstrained optimization is deemed unacceptable in our setting, since the optimum could converge on prices that are not profitable nor realistic.

Some constraining factors were already covered in the "Data" section. Products with few transactions during the chosen time period were filtered out. Additionally, the products with only one active price during that span were excluded from the analysis. Still, the optimization phase requires further constraints to return feasible prices as the optimum.

The constraints that apply to most product groups regardless of the product type or price are the buy-in price and a price that is significantly higher than the base price with an assumption of the demand converging toward zero when approaching this price. The latter one could be substituted to that of the price level offered by the competition, but that is left out of the scope of this research. It was originally planned that the initial linear models for individual products would produce a maximum price level from the $Price$-axis intersection, but that was later abandoned. Instead, an arbitrary maximum price level was chosen so that this level was significantly higher than the original base price for each product. Thus reaching this maximum price would give a clear suggestion on the pricing of the respective product while avoiding any inaccuracies that may have resulted from the own-price elasticity computations. Probably the more intuitive constraint is the buy-in price, which means the price that the Company X has to pay to acquire the products from respective manufacturers, producers and wholesalers. Said prices can be extracted from the raw data in the product dimension. These buy-in prices represent the lowest possible post-tax price in the optimization model, whereas the maximum price is that where the demand equals zero.

In addition to these constraints, we include some constraints that force the values of different variables to either real numbers, integers or binary values. We need to constrain the number of iterations as well. The script includes a constraint that breaks the optimization after a certain number of iterations have been performed or the change between consecutive iteration steps is small enough to be insignificant. In the latter scenario, we interpret that an optimum is found.

The revenue optimization problem can be written as:

$$\text{Objective function} : \max R = P^T \ * \ Q(P) \tag{21}$$

$$\begin{aligned}
\text{Subject to} : P \ &= \ [p_1, p_2, ..., p_n]; \\
&p_i^{buy-in} > 0; \\
&p_i^{max} > 0; \\
&p_i \leq p_i^{max}; \\
&p_i \ * \ t_i \geq p_i^{buy-in}; \\
&p_i^{max} > p_i^{buy-in}, \forall i; \\
&t_i \in [0, 1]; \\
&n \geq 2; \\
&P = [n \times 1]; \\
&W = [n \times n]; \\
&c = [n \times 1]; \\
&Q = [n \times 1].
\end{aligned} \tag{22}$$

The equation (21) shows the short version of the objective function and defines price P as a vector of the individual prices in the product group. The group of equations (22) shows the most common constraints that apply to our objective function. In order of presentation, the constraining factors of a buy-in price p_i^{buy-in} and the maximum allowed price p_i^{max}, are defined as positive prices. The relation of the price p_i is then defined in relation to the former price levels. The included constant t_i stands for the value added tax, which is constrained between $[0, 1]$. Finally, in order to be able to utilize any cross-price elasticities, we need at least two products included in the problem and thus $n \geq 2$. The remaining constraints define the required matrix dimensions in the problem.

The objective function becomes more complex when optimizing profit. Namely because we need to take into account the products buy-in prices and value added tax. Additionally, these factors require further constraints to be added to the optimization problem. For profit optimization, it was decided to add an optional constraint depicting the level of gross margin that the Company X was hoping to maintain. For confidentiality reasons any explicit levels are not disclosed. The optimization problem for profit maximization is presented below:

$$\text{Objective function} : \max Pr = [(P * T) - P^{buy-in}]^T \; * \; Q(P) \tag{23}$$

$$\begin{aligned} \text{Subject to} : P \; &= \; [p_1, p_2, ..., p_n]; \\ T \; &= \; [t_1, t_2, ..., t_n]; \\ P^{buy-in} \; &= \; [p_1^{buy-in}, p_2^{buy-in}, ..., p_n^{buy-in}]; \\ & t_i \in [0, 1]; \\ & p_i^{buy-in} > 0; \\ & p_i^{max} > 0; \\ & p_i \leq p_i^{max}; \\ & p_i \; * \; t_i \geq p_i^{buy-in}; \\ & p_i^{max} > p_i^{buy-in}, \forall i; \\ & t_i \in [0, 1]; \\ & n \geq 2; \\ & P = [n \times 1]; \\ & T = [n \times 1]; \\ & W = [n \times n]; \\ & c = [n \times 1]; \\ & Q = [n \times 1]; \\ & GM\% \geq X\%. \end{aligned} \tag{24}$$

The difference in the objective functions for revenue and profit can be seen in the first term of the function. In equation (23), the first term calculates the post-tax prices for each product and subtracts the buy-in prices from them. Thus ending up with the amount each sold product creates in profit.

Of course things like employee salaries, other taxes in addition to the value added tax, warehousing costs and rent payments affect the actual, ultimate profit for the company. These factors are left to another research though.

In the equation group (24), the last constraint (GM) is the gross margin constraint. Gross margin is the relation between the difference of total revenue and the cost of goods sold, and the total revenue. This relation is usually presented as a percentage of total revenue. The equation (25) shows how gross margin is calculated:

$$GM\% = \frac{R - COGS}{R} * 100\%, \tag{25}$$

where R stands for total revenue created from the sales of products under observation, whereas $COGS$ stands for cost of goods sold which in this case means the buy-in prices of selected products. The gross margin constraint is used as an optional constraint, meaning that it is left out on some computations to see how the prices are affected and if the constraint is even necessary.

In conclusion, we create two objective functions for the optimization phase. These objective functions include results from the cross-price elasticity computations, in the form of the regression model for quantity demanded. Additionally, some of the most important constraining factors included in the optimization were explained. The results from the different phases of this research are presented in the following section.

4 Results

This section introduces the results obtained from our computations. We start by illustrating the own-price elasticities using graphs and evaluating the fit of the model by referencing the R^2-values. We continue by presenting the results for cross-price elasticities. The tables presented have only cross-price elasticity values, but no product specific information due to confidentiality. Here we will show how the sample size might affect the accuracy of the results. These cross-price elasticity values are our primary research goal. They will be subject to analysis in this section.

The conclusion to this section includes the results from the price optimization. We compare results from different objective functions, in essence we compare the results when optimizing revenue and profit. We cannot give any exact prices, but we will describe the theoretical implications obtained from the optimization. We will use percentages and proportions to give the reader some idea on the effect this research could have. The quality of these results is the other point of interest for this research, and it will be addressed in this section as well.

4.1 Elasticities

Our goal was to enhance the quality of the own-price elasticity computations and to run these computations on as many products as possible. The constraining factors included low sales numbers and unchanging prices. In the end, we were able to obtain thousands of individual elasticity values for products. The following section will introduce some of those results and provide some interpretation to those as well.

More importantly, we aimed at creating cross-price elasticity tables for product groups. We will present some of those tables without giving away any product specific information. We will show the cross-price elasticity values for different sizes of product groups. The accuracy of the model is addressed through statistical variables, such as the R^2-value.

Finally, we will give the optimization results for the product groups that were presented in the cross-price elasticity section. The optimization results are derived from several different optimization methods and the quality of the results is addressed. We will analyze the obtained results and implications derived from the optimization. Further discussion of the thesis is left for the final section of the work.

4.1.1 Own-price elasticity

We had to develop the script from the base work further in order to be able to handle more data from an uncontrolled sales setting. The difficulties arose from the fact that the products are subject to a large scale of different promotions, discounts and special offers that are impossible to account for at this stage and hardware. The base work obtained results from controlled store level tests, which is not the case in our research. However, the results we obtained for own-price elasticities are reasonable and now we will present different cases with graphs and analysis thereof.

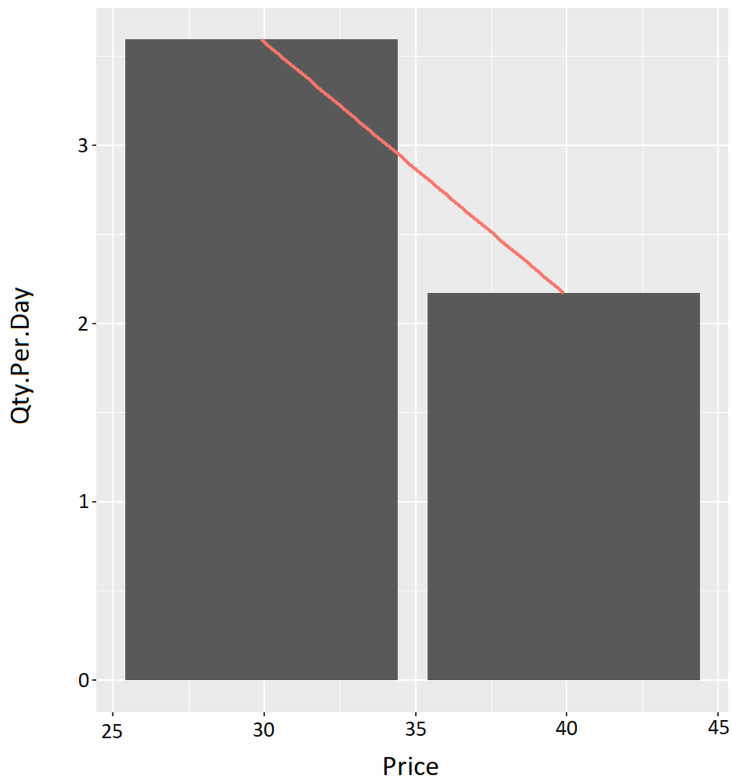

Figure 5: An example from functional clothes with negative own-price elasticity.

For fashion retail products, the assumption is that most of the products would be so called normal products with negative own-price elasticity. A large portion of the items showed this kind behavior, namely the demand for a product seemed to increase when the price decreased. This kind of consumer behavior would seem rational. The figure 5 presents a very simple case, where a product has been for sale for two distinct prices and the average daily demand for that product has been higher when the price has been lower.

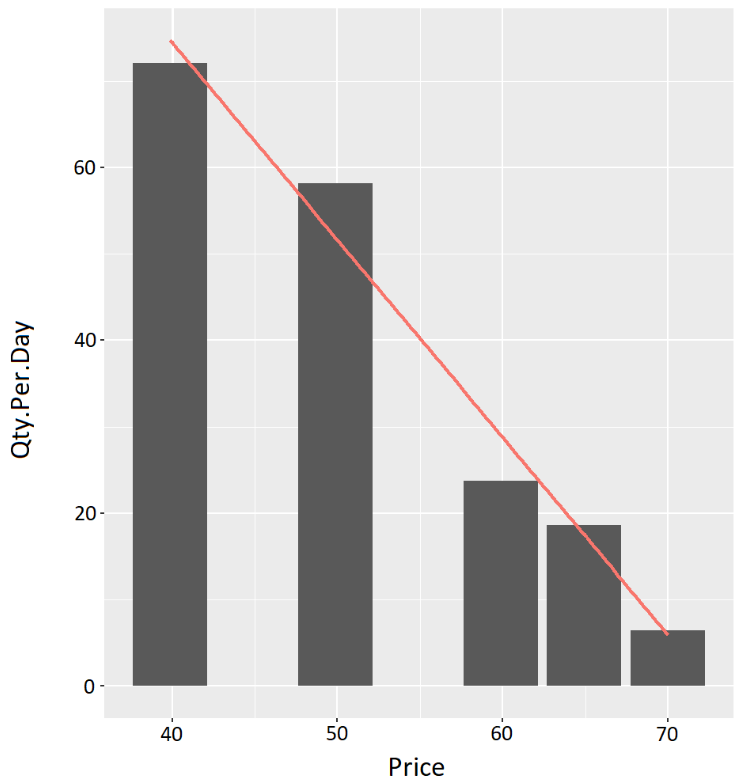

Figure 6: An example from footwear items with several distinct prices and a negative own-price elasticity.

The product in figure 5 is from the group of functional clothing. The negative own-price elasticity value can be derived from the slope of the fitted linear regression model that is also presented in the figure 5. This product was a simple example of the results we obtained in this phase. Similar results were obtained for products with more discrete prices and from other product groups, such as is presented in figure 6. This product is from the fashion footwear category and is a product that has been at the market for several years. Thus it can be assumed that the consumers are somewhat familiar with the product and its' pricing. Consequently, it can also be assumed that the consumers recognize the good, low prices and are encouraged to make the purchase. The very high demand for the product in figure 6 on low prices would suggest this as well. The increase in the amount of discrete prices for an individual product affects the model's goodness. Naturally, having more prices causes the linear regression model to account for less of the variation in the data. Thus the R^2-value does not reach as high levels. For example the product in figure 6 has $R^2 \approx 0,98$, which is still very high. Additionally, all the statistical tests, such as the F-test for overall significance, provide good values. However, looking at the residual plot for this product in figure 7, we can see how the residuals have a rather large variation even though the model seems to fit well. Of course, the residual plot has to be read with the scale in mind. The conclusion is that, even though the model seems to fit quite well, there would be room for enhancement. The model could be underspecified and could thus benefit from additional predictors or nonlinear regression instead of linear. Such enhancements will be left for the future development of this project.

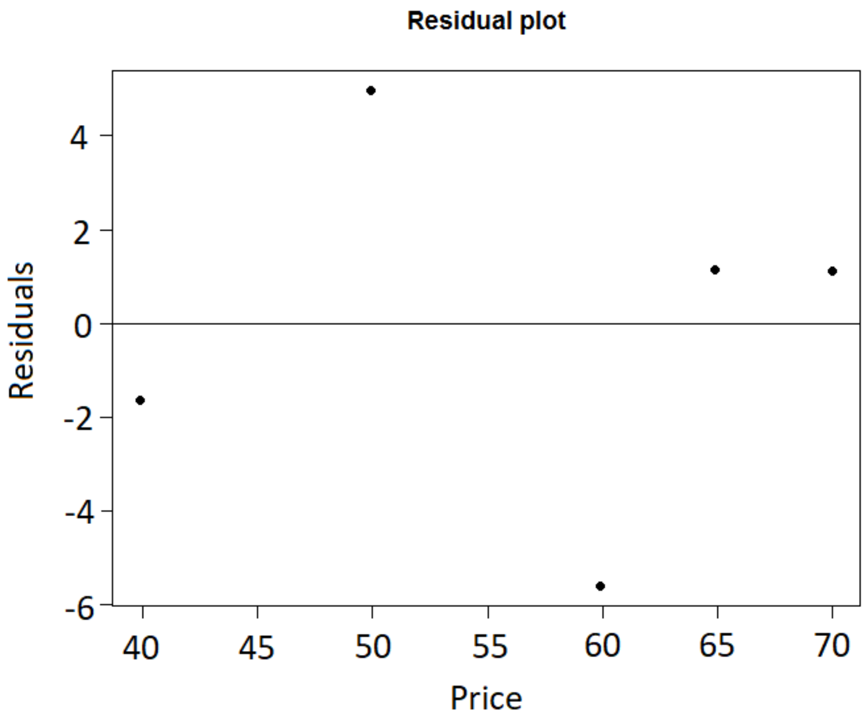

Figure 7: Residual plot for the product in figure 6.

Apart from the "normal" products as described above, there were some products that, for some reason, showed a positive own-price elasticity. Usually positive own-price elasticity means that the demand for the product increases when the price increases. Such items are somewhat rare and were not expected to be included in the product pool of the Company X. The figures 8 and 9 present two of such products from different product categories and significantly different demand volumes.

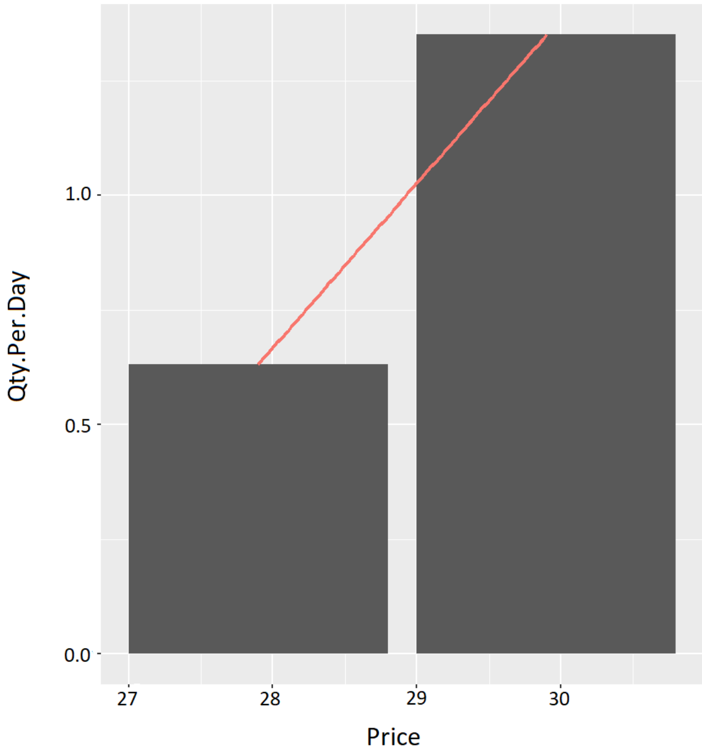

Figure 8: An example of positive own-price elasticity from training equipment.

The sales data for the products in figures 8 and 9 was run through a script that removed the effect of running out of stock and long periods of no sales. This procedure had no effect to their elasticity values however. While the regression models fit the data very well ($R^2 \approx 1$), the reason for such phenomenon is not clear. The most probable reason for this kind of anomaly is that the products have been subject to a price change when their volumes have been reduced significantly. Such situations could arise for example in

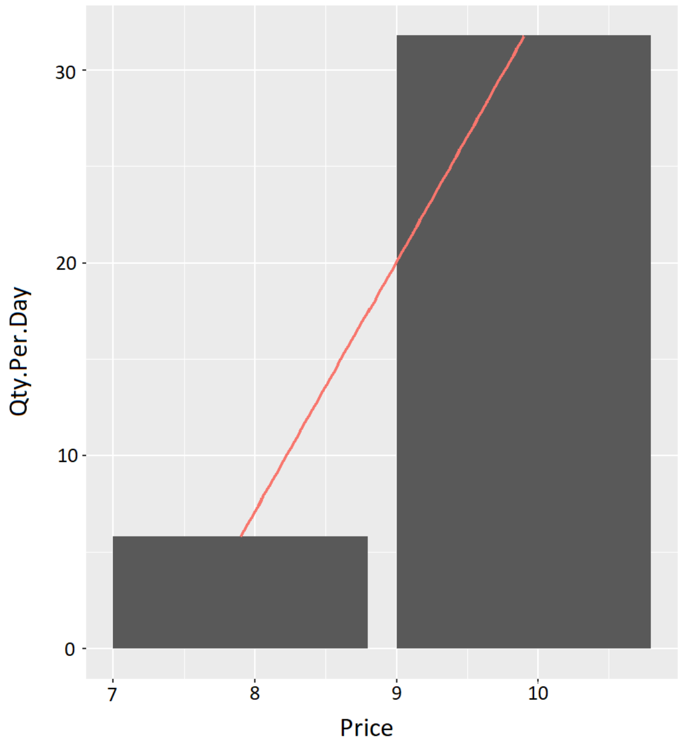

Figure 9: An example from fashion accessories with a positive own-price elasticity.

end of season sales when the products running out of season get a discount to increase their sell through. Furthermore, a release of an updated model of the same product (with a new product ID) could affect the demand for these "outdated" products. In relation to the updated model being released, the older model, especially if low on volume, might not have the best location in the store, which could lower the sell-through even further.

Additionally, there were a small number of products that showed odd demand behavior. These products seem like they have characteristics from both of the previously presented examples but as a result, the linear regression model does not fit the data at all. The common aspect for these products is that there is no clear trend in the data and thus, the fitted model does not represent the data well. The R^2-values of these models are near zero as a consequence. Two examples are given in the figures 10 and 11.

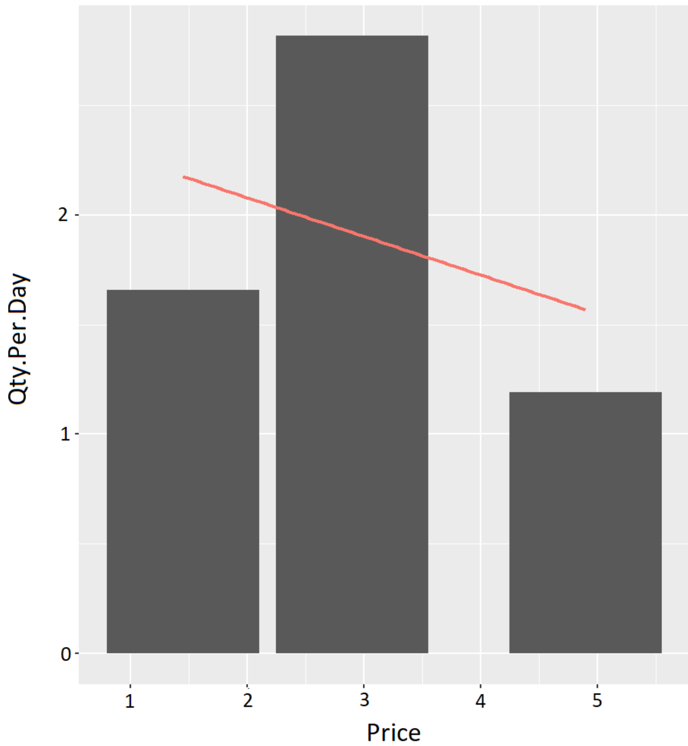

Figure 10: An example from training accessories presenting an odd demand profile.

The products showing this kind of demand profiles were run through the script that removes the effect of running out of stock, but it had little effect on most of them. The products kept showing similar demand profiles and the elasticity values were still unreliable. An example of a product that had a moderate change in the demand profile when run through the additional refinement is presented in figures 12 and 13.

As can be seen when comparing the figures 12 and 13, the demand profile changes after the refinement. However, the graph 12 shows a clear trend if the second lowest price was to be disregarded. This trend is lost in the out-of-stock refinement. The statistical tests give very poor results of statistical significance; the R^2-value equals 0.50 and the p-values are rather large (≥ 0.1) on every statistical test, suggesting poor statistical significance, thus it can be stated that the linear regression model fails to explain the variations in the data. A confirmation to that statement can be seen in the residual plot of the data in figure 14, which represents the residual plot of the model

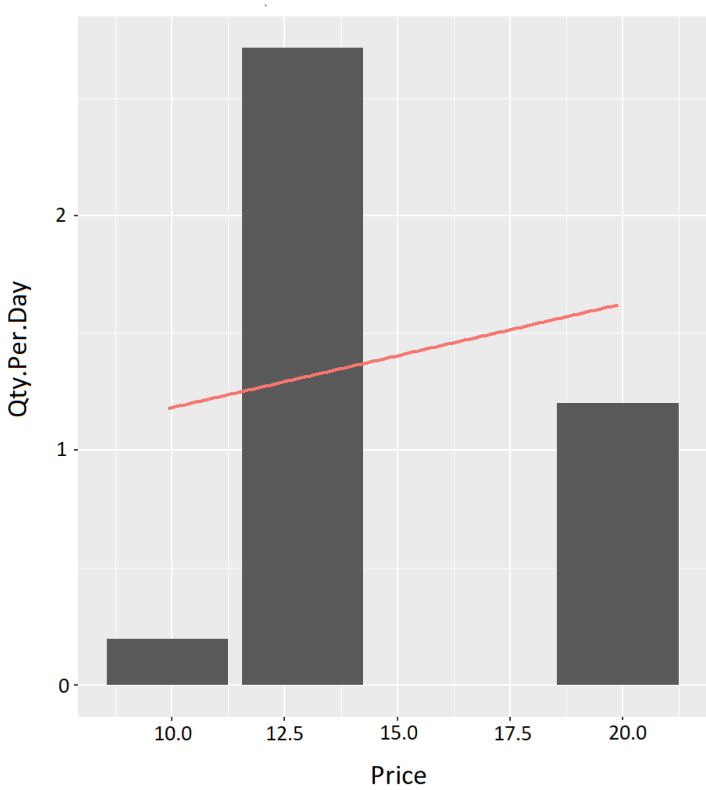

Figure 11: An example from fashion accessories presenting an odd demand profile.

from figure 12. This plot seems similar to that in figure 7 but the scale of the sold quantity must be taken into account. This latter example has residuals that extend over the actual sales quantities, whereas the residuals in the first example are significantly smaller compared to the actual observed quantities.

As noted earlier, some of the items produced odd demand profiles, to which it is not possible to get a well fitting linear regression model. This notion is supported by the R^2-values derived from those regression models. These values, depicting how well the model fits the data, range as low as 10^{-4}, which would suggest the model does not fit the data at all. The R^2 value ranges between $[0, 1]$, where value 1 means complete fit. Another interpretation for the R^2-value is the regression models ability to explain the variations in the data. However, only high R^2-values are not enough to validate the model to fit well. The residual plots need to be consulted to ensure the goodness of fit. Additionally, adjusted and predicted R^2-values provide additional insight to

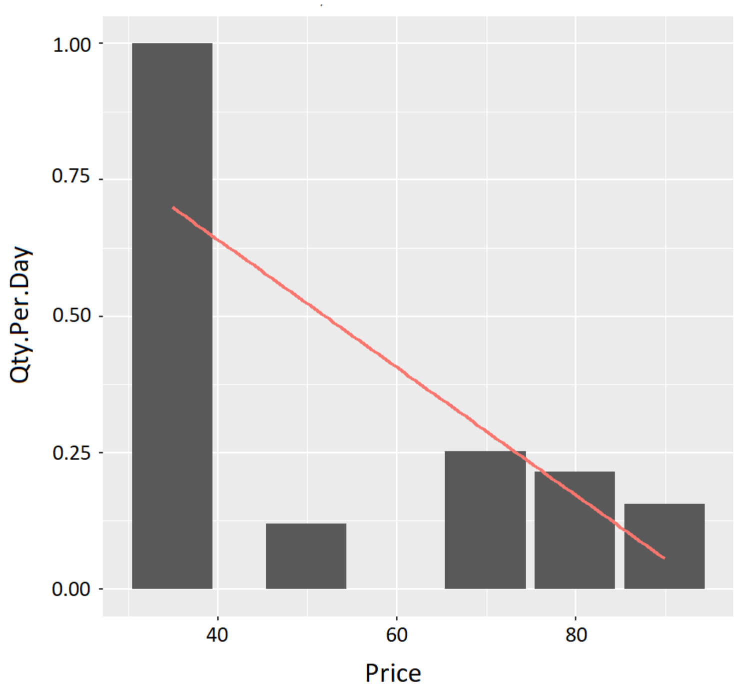

Figure 12: Badly fitting regression model from footwear category.

the goodness of fit. The latter ones are especially useful later in the multiple regression analysis. The odd graphs produced by the simple linear regression analysis in this section can be stated to have a poor fit, since the fitted line does not follow the data points (or histogram bars) at all.

The poor fit could be a consequence of several issues. For example, the model could be underspecified, missing data for the predictors or some predictors altogether, which is called a selection bias (Dubin & Rivers, 1989). This is likely at least partly the reason for our anomalies. Additionally, the data could be of poor quality or too aggregated, which would cause some information to be lost in the computation. The data refinement could also be to blame for some issues revealed by the results. If the data is refined too eagerly, some relevant data points could be lost and thus result in distortion in the final results. These results should and will be consulted during the cross-price elasticity computations. If those computations produce unexpected numbers, then the cause might be visible here.

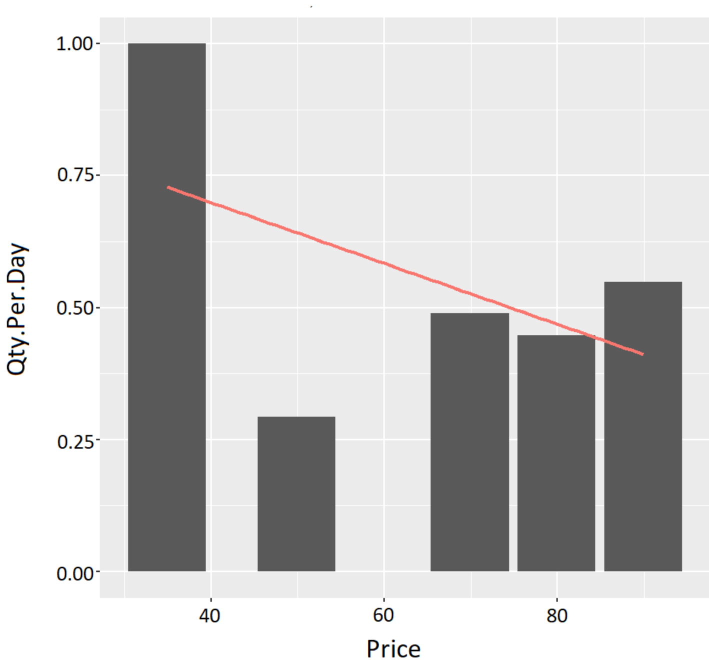

Figure 13: The same product from figure 12 with an out of stock refinement.

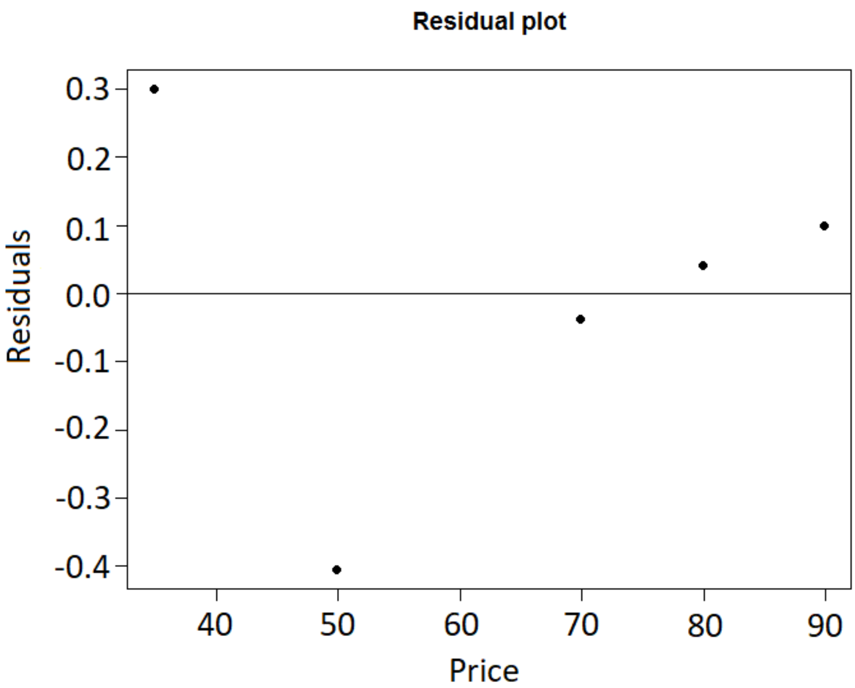

Figure 14: Residual plot for the product in figure 12.

4.1.2 Cross-price elasticity

For computing cross-price elasticity values for product groups, we needed to produce a completely new script, which would handle data refinement and relevant computations to form a multiple regression model from which the elasticity values could be extracted. Such a script proved to be quite challenging due to the quality of data and the desired format of outputs. This section will present some of the obtained cross-price elasticity results and also address the computation settings, including the observation period and the size of the product groups, as well as interpretation of the obtained values.

We started out testing small product groups and determined their cross-price elasticity values before building up to larger groups. The first product group is a small pool of functional footwear items. These items were considered substitutes and only the basic variants with highest volumes were chosen to this computation. The elasticity values and corresponding R^2-values are presented below in the tables 1 and 2.

Table 1: A table of cross-price elasticity values for a group of footwear with aggregated data.

	Shoe 1	Shoe 2	Shoe 3	R^2	Adj.R^2
Shoe 1	-12.60	-0.72	4.10	0.99	0.93
Shoe 2	-1.06	-0.39	2.32	0.94	0.86
Shoe 3	-1.96	1.20	3.18	0.94	0.84

The table 1 includes cross-price elasticity values derived from three different regression models, the R^2-values of which are presented in the table 1 as well. Each model is filled in the table on it's own row, as will be done from here on. This means that the regression model for the demanded quantity of Shoe 1 is presented by the row named "Shoe 1". The elasticity values present the effect a change in the price of products in the columns has on the quantity demanded of a product in the rows as demonstrated earlier in equation (6). In this table, the aggregated values were used when fitting the regression model. The values in table 2 were computed without aggregation for the exact same items and time period.

Let's first compare the results in the tables 1 and 2. The elasticity values are somewhat similar in both examples. The crucial difference can be found in the R^2-values. The aggregated data yields significantly better R^2-values and

Table 2: A table of cross-price elasticity values for a group of footwear without aggregation in the data.

	Shoe 1	Shoe 2	Shoe 3	R^2	Adj.R^2
Shoe 1	-12.27	-0.78	3.96	0.23	0.22
Shoe 2	-0.86	-0.52	2.26	0.02	0.01
Shoe 3	-0.50	1.78	2.08	0.26	0.25

the adjusted R^2-values remain rather high as well. On the other hand, the latter table without aggregation shows very poor values. This is caused by the deviations in the daily sales quantities, namely an item can experience a demand of no items on one day, and dozens of items a day after without any price change. The aggregation removes the effect of such deviations. On the other hand, too much aggregation leads to loss of data points and thus can make the results unreliable and distorted.

The values in the table 1 differ in magnitude, the highest absolute value being ≥ 12 and the smallest being close to zero. The surprising find is the high own-price elasticity value for the Shoe 1. A value this high would suggest that a 10% change in the price of Shoe 1 creates an opposite 126% change in its' demand. Even though in this case we used average daily demand numbers, the change would be extremely high. Either the pricing for Shoe 1 is far from the optimal level or the demand profile is very steep. Furthermore, the own-price elasticity value (on the diagonal) for Shoe 3 is a positive value, which is usually not the case. This would suggest that the demand for Shoe 3 increases when it's price is increased. This anomaly can be caused by simultaneous markdowns in the product group or some promotion that is only applicable on Shoe 3 and could not be accounted for in the data refinement. Further, this phenomenon could be due to the Shoe 3 getting a price adjustment only when the volumes were low, or when the observation period was about to end.

Finally, it is interesting that the Shoe 2 has a rather small own-price elasticity ($= -0.39$), whereas the respective cross-price elasticities are significantly higher in magnitude ($-1.06; 2.32$). This would suggest that the Shoe 2 has a stable (inelastic) demand respective to it's own price, while it serves as a substitute to the Shoe 3. Namely, decreasing the price of Shoe 3 leads to a decrease in demand of Shoe 2 and vice versa. This interpretation follows the cross-price elasticity values of the two items with respect to each other. They are both positive and ≥ 1 in magnitude. Overall, the demand for Shoe

2 seems to be less sensitive to price changes than that of the other two items. Shoe 2 has a weak cross-price effect on the demand of items Shoe 1 and Shoe 3, but it also has a small own-price elasticity value.

The table 1 tells us about the dynamics behind the price adjustments and more precisely how the demand for products is affected by the change in price. If we were to take a look at the first of the three rows in table 1, we can see that the own price elasticity value is highest in magnitude. Namely, it is higher than the two other elasticities combined. This suggests that, given a simultaneous 10% reduction in price for all the products in the group, the demand for Shoe 1 would still increase by about 90%, which is a very large reaction on small price changes. The other two items do not behave as radically though. With a similar example of simultaneous markdowns, the demand for Shoe 2 would decrease by about 9% and the demand for Shoe 3 would decrease by about 24%. The decrease in demand for the Shoe 3 is caused by the strong positive own-price elasticity.

The computations were run on the same product group, with the extension of including all product variants, in essence all colors, in the same product group. The resulting tables C.1 and C.2 can be found in the Appendices section under Tables. The table C.1 has been derived by using aggregated data, whereas table C.2 has no aggregation in the data. The addition of many products causes the regression models from aggregated data to include singularities, which shows on the table C.1 as the $R^2 = 1$ while the adjusted R^2 is not applicable (NA). These singularities can be caused by too much aggregation, no price changes in the observation period or a simultaneous markdown. Additionally, the problems could be caused by the method chosen to run the multiple regression. Furthermore, the R^2-values seem to decrease for those items that were chosen for the previous computation presented in table 1. For clarity, the corresponding rows and columns were named similarly to the smaller tables. As we take a look at table C.2, we note that the R^2 values for items Shoe 1, Shoe 2 and Shoe 3 have increased compared to those in the previous computation. Still, the values are very low for all the products. However, the computation without aggregation seems to avoid any singularities and returns a multiple regression model for all the items, although one with poor R^2-values. Alone, the R^2-values are not enough to make a definitive statement of the fit of the model, but values this poor suggest that the model is unable to explain the variations in the data and should be developed further. Applying more extensive aggregation to the data could lead to better fit while sacrificing a portion of the data points. This sacrifice means less data points and thus a less reliable model. The next step could be to try finding the best predictor variables to each item individually, or trying to develop the model further to account for more of the variations in the data.

Both the tables C.1 and C.2 include cross-price elasticity values that equal 0. This would mean that a price change in the corresponding item (in the columns) has no effect on the demand of the item in the rows. For example, Shoe 05 has a cross-price elasticity value of 0 for every other item in the group (column Shoe 05). The last value on the column is the own-price elasticity of Shoe 05. In essence, the price changes in Shoe 05 have no effect on the demand of other products in the group, but some of the others have a cross-price effect on the demand of Shoe 05. This phenomenon might be due to Shoe 05 having no price changes during the observation periods for the other items, or simply due to a price change in Shoe 05 having absolutely no effect on the other items' demands. The same logic can be applied to the other examples, where the cross-price elasticity value equals 0. If the individual demand models were to be developed for each item, we would try excluding the items with a cross-price elasticity value of 0 from the respective models. Thus we could obtain better overall models for each item. It needs to be noted that even if an item has a cross-price effect on another item, that does not necessarily mean that the effect would work both ways. For example, in table C.2, Shoe 01 has no cross-price effect on Shoe 0, but Shoe 0 has a cross-price effect of -22.06 on Shoe 01.

Another interesting notion was found when running pairwise comparisons on similar footwear items. These items were not included in the previous computations and they were considered direct substitutes in the sense of their attributes and the pool of consumers who require these items. The result of the pairwise comparison is presented in the table 3 below. We see that according to the first row, the assumption of the products being substitutes is supported by the cross-price elasticity values. However, the values are not necessarily the most reliable ones due to a very low adjusted R^2-value. Having a negative value could be considered a failure with the fitted model even though the normal R^2-value is significantly better. On the other hand, the second row suggests a very good model fit when the roles of the items are changed. This model would also suggest the items not to be substitutes due to both elasticity values being negative. This is not necessarily a wrong interpretation, since Shoe 5 can be considered a substitute for Shoe 4 while Shoe 4 might not be one for Shoe 5. As stated earlier, such values could also be caused by simultaneous markdowns and the resulting distortion in the data analysis. Simultaneous price changes would disturb the analysis in the sense that we want to find an effect of a price change in one item on the demand of another when all other factors stay unchanged. Simultaneous price changes cause inaccuracy since, with the resources at our disposal, it cannot be determined how much a single price change affected the outcome if there were more than one. Unfortunately, in fashion retail there are often times, such as summer or Christmas sale periods, when items are subject to simultaneous price adjustments.

Table 3: A table of two footwear items assumed to be substitutes with aggregation used in the data.

	Shoe 4	Shoe 5	R^2	Adj.R^2
Shoe 4	-1.98	4.12	0.61	-0.17
Shoe 5	-2.23	-0.32	0.999	0.997

The following item group includes three highly seasonal products that are designed for the same purpose. These items are from two market leader brands and considered similar in image. Two of the items belong to the same collection, one being the "high quality" model while the other is the downgraded and more affordable version. The elasticity values for these items are presented in the table 4 below.

Table 4: Elasticity values for a small group of seasonal products with aggregated data.

	Item 1	Item 20	Item 21	R^2	Adj.R^2
Item 1	-3.50	-1.43	-1.59	0.58	0.16
Item 20	-4.25	-5.07	0.74	0.79	0.63
Item 21	1.40	1.23	3.55	0.86	0.73

Again, there are some differences in the R^2-values between the models on rows. Nonetheless, the first two own-price elasticity values seem rather normal. The third however, is a positive value which is not usual for these products. The explanation behind this value could be found in the seasonality of the product. The Item 21 is the downgraded version of Item 20 and is sold for a lower price. Lowering the price of Item 21 could present lower demand due to the end of the season or lower quantities of inventory. If there are simply less items left to sell, the demand becomes lower with the reduced price. Further, it could suggest that the lowered quantities in inventory have actually motivated the price reduction to make room for newer in-season items. The interpretation of lowering the price of Item 21 in the end of season is also supported by a notion of its' negative cross-price effect on Item 1. Item 1 is a typical example of a product that has been on the market for years and usually gets a price adjustment for the end of season sales period if at all. Thus, the value can be caused by simultaneous price adjustments, of which the price adjustment on Item 1 carries a stronger effect on its' demand than the adjustment on Item 21. On the other hand, if the Item 21 gets a price adjustment when the inventory is already low, it is easier if not necessary for customers to choose the more expensive product instead to find the right purchase for them. This would also facilitate the cross-price elasticity value to become negative.

The products Item 1 and Item 20 belong to the category of more expensive products in their respective group. They also have negative cross elasticity values on each other's demands. This could be at least partly due to the same reason as discussed above. Consumers could choose the more expensive product instead of the basic model if they get a better price for it than normally. This could lead to a situation where reducing a price of such an expensive product could tempt consumers, who would otherwise buy the less expensive basic model, into buying the "better" alternative. In essence, the top line items would steal or cannibalize a part of the demand that would otherwise be directed on the basic line products. A combination of these reasons could present negative cross elasticity values for items that are assumed to be substitutes. Thus also leaving a question if there was a selection bias present when choosing the predictors to the model.

Further problems may occur from poor choice of items in the observation group. Correlation analysis could help to identify the best suited product groupings, although that approach works best when determining the model predictors for the demand of a single item instead of a group of items. The following table 5 presents a poorly chosen product group.

What indicates the bad choices for products is the amount of zeros in the table. For instance, items Acc21, Acc31 and Acc32 yield only an own-price elasticity value. This means that they have no cross-price effect on any other item in the group. This could be due to, for example, these items having no price adjustments during the respective sales periods. It could also be that the price adjustments made to these three products simply had no effect on the demand of the others. Such a situation could arise from the items being of different "tiers" or quality levels. Given that the three previously mentioned items were of different levels of quality, changing their prices might have no effect on the items in higher tiers that would in turn keep demonstrating sta-

Table 5: A table of cross-price elasticity values for a group of poorly chosen accessory items.

	Acc11	Acc21	Acc31	Acc32	Acc12	Acc13	R^2	Adj.R^2
Acc11	-2.99	0.00	0.00	0.00	-6.19	4.82	0.35	-0.63
Acc21	-1.90	-0.86	0.00	0.00	-1.66	2.46	0.85	0.62
Acc31	-0.91	0.00	-0.40	0.00	-2.62	2.74	0.29	-0.76
Acc32	-1.36	0.00	0.00	-0.59	-2.42	0.79	0.35	-0.63
Acc12	1.88	0.00	0.00	0.00	0.35	0.24	0.21	-0.97
Acc13	-4.26	0.00	0.00	0.00	0.00	-9.12	0.65	-0.05

ble demand levels. On the other hand, the demand for the three previously mentioned items Acc21, Acc31 and Acc32 is affected by price changes for the other three items, which would suggest simultaneous price changes or the top tier items cannibalizing demand from the other products when subject to a price adjustment. Furthermore, this item group has very low sales quantities, averaging well below an item per day, which would mean that the demand graph would present spikes on some days while being zero on others. Such a demand curve would become distorted when being subject to aggregation. Thus causing loss accuracy on the calculations. Even running the correlation analysis on such a group of items could prove inadequate to avoid these issues, since the cross-price effects may differ when the roles of items are reversed, or when the observed variable becomes a predictor instead.

All in all, the elasticity values in table 5 are not the most reliable ones due to poor fit to the data, according to the R^2-values. This group of items should be run through a correlation analysis and their demand analyzed item by item with the chosen predictors having the highest correlation between their prices and the observed demand. As is, the modeling practice is inadequate to produce accurate and reliable readings on such item groups, although performing well on some individual items.

Similar observations were made of several product groups including clothes, accessories and footwear. The poor values could be caused by simultaneous price adjustments, or the lack thereof as in table C.3 in the appendices. The lack of data points and simultaneous price changes produces singularities in the calculations and leads to a failure in modeling the price-demand dynamics. These poor values provide little information and the reliability of the interpretation is questionable. Further development of the model and computations is needed to obtain better quality results. It can be stated that at this stage, the project has failed in providing a high quality insight to pricing on a universal level across product groups. Although, some individual products seem to produce a well fitting model. Thus it would most likely be best, to run correlation analysis on individual items and analyze each product of interest separately. Running the computation for larger item groups suggests that a correlation analysis on individual items would facilitate a more efficient choice of products to include in each model. Large groups produce cross-elasticity tables with many items having no cross-price effects on others, or only on a few other products. Running the cross-elasticity computations on individual items would however, make it more difficult to optimize entire product groups simultaneously. The next phase provides some optimization results while keeping in mind that the results are derived from the poor quality cross-elasticity values.

4.2 Price optimization

Although the cross-price elasticity results were unreliable and poor in quality, we ran the optimization scripts to see if they could handle the desired form of elasticity values. The regression model coefficients were run through several optimization routines and the following section presents the results thereof. The positive notion is that the optimization routines seem to work and the results are reasonable baring in mind that the cross-price elasticity computations are the basis for the models entered in to the optimization scripts.

We present optimization results for the same item groups that were observed in the previous section in order to have some data that we can reflect on when analyzing these results. It needs to be noted that these results are only to get some insight on pricing. For more accurate and actionable results, the regression model should be developed further in order to produce more reliable and realistic cross-price elasticity values. Additionally, we used several different optimization methods as explained earlier in this work. This section will address the differences between the results obtained from each of them as well as the performance of the methods. The optimization methods utilized in our research are a conjugate gradient method, Nelder-Mead algorithm and a limited memory BFGS-method with bound constraints (L-BFGS-B). The methods were briefly introduced earlier in this work. Each product group was run through each of these optimization methods and the obtained results are now compared and analyzed.

The first item group presented in the cross-price elasticity results (table 1) is a footwear group with three individual products that were considered substitutes and yielded a model with reasonably high statistical significance. The three optimization methods were run on the group for both maximization of revenue and profit. Naturally, the objective function value for profit is always lower than that of revenue since profit takes into account the cost of the goods sold. The higher objective function values were obtained from Nelder-Mead and L-BFGS-B methods. Those two methods yielded the exact same objective function and decision variable values. The issue with these results is that two of the items in the group (Shoe 2 and Shoe 3) were given the highest possible price (160€) that acted as an upper bound in the optimization. That price is approximately 24% higher than the original base price for the products (129€). The third item (Shoe 1) gets a price (100€) that is approximately 24% lower than the original base price (132€). Still, the price for Shoe 1 is rather realistic and the item has actually been sold for that price during some special promotion periods. All in all, Nelder-Mead and L-BFGS-B methods provide high objective function values but the decision variable values leave room for criticism. Having very high prices for the items in question is not realistic due to high competition in the field, and such high prices would steer consumers hoping for lower prices to the competition.

However, the conjugate gradient method provides rather realistic decision variable values while the objective function values are significantly lower than with the other two methods. The obtained prices are 75€, 121€ and 135€ for products Shoe 1, Shoe 2 and Shoe 3 respectively. The price for Shoe 1 clearly reflects the high own-price elasticity value while the others have only had a small adjustment due to lower elasticities. The cross-price elasticity values explain why the price for Shoe 3 actually increases while the other two are lowered. Shoe 3 has only positive cross-elasticity effects, which means that increasing the price for Shoe 3 would increase the demand for each of the products in the group. The opposite stands for Shoe 1. Since it only has negative cross-elasticity values, lowering the price for Shoe 1 would cause an increase in demand for all three products. This would also cancel the loss of profits caused by lowering the price of Shoe 1 with such proportion (-43%) nearly to the lowest possible price. The objective function value for profit maximization is approximately 35% lower than those obtained from the other optimization methods. However, the prices given by the conjugate gradient method are more realistic than those from the other two methods even though the price for Shoe 1 is significantly, and almost unrealistically, lowered. In addition, even the conjugate gradient method that produced the lowest objective function value for profit, yields a value that is 38% higher than that which would theoretically be obtained with the base prices.

Following the first group of footwear items, we presented a pair of assumed substitute items, also from the footwear assortment (table 3). For this product pair, we noted that the conjugate gradient method for profit maximization did not converge on any optimum even though the initial conditions were altered. For revenue maximization however, there is a result which gives Shoe 4 the highest allowed price while Shoe 5 gets a significant price reduction taking the price only 2% above the lowest allowed price. The other two optimization methods yield exactly similar results between them again. Still, the decision variable values are not realistic since Shoe 4 is suggested to be priced as high as possible regardless of the objective function. The highest allowed price in this optimization run was 220€, which is 43% above the base price 154€. As for the Shoe 5, profit maximization would suggest lowering the price to 153€ from the base price 209€. This seems rather realistic, since Shoe 5 has a negative own-price elasticity and would thus experience an increase in demand. If we were to follow the pricing suggestions derived from L-BFGS-B method, the profit levels for this item pairing would theoretically increase from the base price driven profits by approximately 18%. Of course, it needs to be noted that these values are still computed with the model we created earlier in this work. Thus the results cannot be considered as accurate but merely estimates.

These price changes suggested by the optimization are not completely realistic and the resulting changes in demand would be curious if reflected on reality. Having experience in retail, Shoe 4 would most likely have no demand at all due to a very high price compared to the competition. Shoe 5 would sell significantly better than with the base price until the competition reacted to the price reduction with an adjustment of their own. In conclusion, the optimization results for this product pair would suggest that the price for Shoe 4 should be increased, while the price for Shoe 5 should be decreased. The magnitude of the adjustments should be considered however, since these results suggest very high alterations. Additionally, the model that was used to compute the cross-price elasticity values, and thus resulted in the regression model coefficients for the optimization, was not a very reliable one due to poor fit to the data. That is why the optimization phase would also benefit from further development of the regression model and data analytics behind the obtained results.

The item group we observed after the footwear items was one that included three highly seasonal products, two of which were from the same product line. The price optimization returned prices that were quite different with the three optimization methods. We will not go through the revenue maximization since most prices were on the lowest possible price level on almost every method and item. Only Item 1 had a price that was 41% higher than the lowest price allowed from the Nelder-Mead and L-BFGS-B methods. For profit maximization, L-BFGS-B method yields prices that are still on the lowest allowed price, 79€ and 36€ for Item 20 and Item 21 respectively, while the price for Item 1 is 121€. However, Nelder-Mead yields very different prices, since it suggests that Item 1 would be assigned a price of 62€, which is just above the lowest allowed price. While L-BFGS-B gave Items 20 and Item 21 the lowest allowed prices, Nelder-Mead returns the highest allowed price 150€ for Item 21 and the price for Item 20 is 205€, which is 7% below the base price. Again, the conjugate gradient method for profit optimization provides the most realistic price suggestions when reflected on reality. Item 1 receives a price 83€, Item 20 has a price 207€ and Item 21 has a price 71€. The curious thing though, is that the highest profit would be obtained by L-BFGS-B method, that gave lowest allowed prices to Item 20 and Item 21. That profit would be nearly twice as high as with the other methods.

The only possible explanation to that result is that even though two of the prices are significantly lower than with the other two methods, the demand quantities are high enough to produce higher profits. Furthermore, the suggested pricing would give most importance on the sales of Item 1 that would actually bring in all the profit with this pricing choice. Unfortunately, this example brings to light the shortcomings of our model, since the theoretical profit with base prices becomes a negative figure. This would mean that the items would not sell at all and buying them from manufacturers would just create losses to the retailer. The optimization with L-BFGS-B would turn the profits to a positive amount. To give the reader an idea about the magnitude of the difference, let's set the base price driven profits as -100%. Then the pricing suggested by L-BFGS-B method would yield a profit of $+70\%$. This difference is very significant, but still the reality is that the prices that the L-BFGS-B suggests would not be applicable. The most realistic price suggestions were given by the conjugate gradient method, which would still change the profit level to $+35\%$.

We gave an example of an item group for which there were several items that had no cross-price effects on other items in the group. The cross-price elasticity values were presented in table C.3 in the appendices. The optimization for this group tends to converge on the highest or lowest prices allowed depending on the initial conditions. The optimization results vary a lot as well. Some profit levels drop below the theoretical level yielded by the base prices, while at best the profit increases by 182%. However, again the highest profits yielded (L-BFGS-B) suggest prices that are not reasonable. All those prices are on the highest possible level and thus would most likely result in very low demand levels because the cost-conscious consumers would find the same items for a lower price elsewhere. Namely, the highest allowed price levels are now chosen arbitrarily for optimization purposes. In reality, we would need to take into account the price level available elsewhere in the market. The poor quality results from the cross-price elasticity computations are the cause for the optimization results being so inconsistent.

The final group we chose to present in the earlier section is a group of accessories in table 5. This group served as another example of poorly chosen items with a few items that had no cross-price effects on other items in the group. Quite similarly to the previous example, the optimization results varied a lot depending on the initial conditions. Now the highest objective function values were obtained from both Nelder-Mead and L-BFGS-B methods. They suggested the exact same decision variable values, which would result in a theoretical increase in profit of 44% when compared to that from the base prices. In reality though, the increase in profit would likely be a lot lower since the underlying model was stated to be inadequate to account for the variations in the data. Additionally, the highest profit levels were often obtained by the highest allowed prices. Setting prices that high would not be rational due to the nature of fashion retail market where consumers are highly cost-conscious and have a low threshold to shift their alliance from one retailer to another in hopes of better prices.

After running the optimization scripts on various product groups, it can be stated that the highest and best profit values are obtained from the L-BFGS-B method. Many times the Nelder-Mead method converges on the same values as L-BFGS-B method but does not perform as consistently. The poorest objective function values were obtained from the conjugate gradient method, while on the other hand the prices suggested by this method were on most cases more realistic when compared to the other methods. To make any definitive statement of the goodness of the optimization methods, further development of the model is required. To conclude the optimization chapter, we note that the obtained results provide some insight into pricing and can be referenced during pricing decisions. The potential for enhancement in profits is there but would require further development of the model to be utilized. Namely, the objective function is based on the model that was used in the cross-price elasticity computations. Since the model was considered inadequate, the results from the optimization are lacking in accuracy as well. That notion is obvious in the pricing suggestions derived from the optimization, since a large portion of the prices converge on either the highest or lowest allowed level. Those levels were chosen to present prices that were not rational, nor profitable to use in reality. Either the gross margin would fall to zero or the price would be higher than that of the competition and thus would lead to decrease in demand. Unfortunately, the model does not act as was expected and as would be rational. The reasons for the models shortcomings were already analyzed earlier. The following section will address the project as a whole and discuss the future of the project.

5 Discussion

This project aimed at computing cross-price elasticity values for product groups and to use these values in price optimization with objective functions in revenue and profit. The goal was not to obtain accurate prices since the elasticity values are always estimates, but to gain additional insight to product pricing and inter-product relationships. The results provide that insight and the optimization phase yielded price suggestions that are more or less reasonable and result in significant enhancements in created profit. However in many cases, especially with large product groups, the model fails to account for the variations in the data. Consequently, the optimization results are also poor in accuracy since they are based on the cross-price elasticity estimates. The shortcomings and development ideas for the project are discussed in the following section. The project could be continued by developing the model further to obtain more accurate estimates from the cross-price elasticity computations. The decision on the future of this project will be made by the Company X.

5.1 Development ideas

As stated on several occasions throughout this thesis, the model that was chosen to explain the data and produce the elasticity coefficients was not good enough to account for the variations in the data, especially when handling larger product groups. For individual items, the model seemed to work well enough to produce demand graphs and own-price elasticity values. This section will include thoughts and development ideas, as well as discussion on factors that had to be left out of scope in this project.

First of all, the choice of using linear regression in the first part of the project worked well for individual items, whereas expanding to multiple regression when handling product groups proved to work quite poorly. As a development idea for the individual items, it could be beneficial to trade linear regression model into a more complex model to better fit the data. For instance, developing a nonlinear demand function (e.g. utilizing cumulative beta distribution function) that could account for several items as predictors and had a better weight coefficient structure could be beneficial. That approach however might require tailoring the model coefficients for each case separately. Additionally, the model might benefit from including other external predictors, such as special promotions, marketing inclusion (e.g. digital and paper advertisement) or seasonality (e.g. Christmas). With these notions, a time series approach could be tempting but would not work well for items with very low and inconsistent demand. Thus, creating a dynamic pricing model that includes the proper predictors and company specific constraints and constant factors should be considered. Tailoring a model for specifically Company X's product pool is a challenging and complex task and would require further resources and knowledge as well as sophisticated mathematical programming. Although, such models are already utilized in other markets. The dynamic pricing models could update the optimal prices on the fly to react on fast and abrupt changes in the fashion retail market and provide up to date elasticity estimates. The dynamic pricing models have been discussed by for example Hall, Kopalle and Krishna (2010) as well as Grewal and his colleagues (2011). Typically, such dynamic pricing models that are provided by specialized companies and tailored to the needs of the company in question are extremely expensive and thus not the primary option for Company X.

An interesting notion on the cross-price elasticity computations was that it could be more beneficial to observe the products one at a time, especially if there was no investment on the development of a dynamic pricing model. This practice would require more time, but allow us to find the items carrying most meaning across all item groups through for example correlation analysis, thus providing further insight on product complementaries and substitute products. Such approach would also allow us to determine cross-price elasticity estimates with potentially higher accuracy and statistical significance than the approach of this research. As running the computations for product groups proved to be challenging and inaccurate, it could be better to take the items individually and determine an optimal price for them one at a time instead of an entire product group at once.

In addition to developing the model further, more computational power is needed to run the analyses efficiently. In this work, the largest product groups, that included approximately 20 items took more than two hours to run the computations. Having more data and more complex models would require updating the hardware to another level. Furthermore, the possibility of transitioning to an R-SQL server environment would allow more random access memory (RAM) and thus relieve some of the strain the computations will cause.

Finally, due to a strict scope that we were forced to set on this project, some interesting and rather meaningful factors had to be left outside the research or worked around in the data refinement phase. Some products were sold for dozens of different prices during the observation periods, which was caused by various different promotions, special offers and coupons to name a few reasons. We managed to take some of these into account in the data refinement, but more work on this issue is required to provide a more realistic and less aggregated data input. Additionally, we were unable to handle simultaneous price adjustments well enough. Such adjustments caused distortion in our cross-price elasticity estimates and thus later in the optimization results. In the future research on this topic, it would be interesting to include marketing inclusion in the modeling phase. Thus we would be able to gain insight into the extent a product's demand is affected by that product being shown in different marketing media. Another related point of interest would be to make a comparison between different promotion and markdown types. This would require further development in the data refinement to be able to recognize different types of markdowns, such as loyalty program prices, package prices (e.g. buy 3 pay for 2), percentage offers (e.g. -20% off all jackets) or normal markdowns. The difference between the loyalty program offers and other markdowns carries a special interest for Company X, since their loyalty program members are extremely active.

This thesis did not separate different demographic (Hoch et al. 1995) or geographic areas when running the computations. Doing so would provide insight into market differences between different areas of business. Combining this approach with the dynamic pricing model could provide Company X with an opportunity to tailor prices (Levy et al. 2004) and product supply according to the specific local market. Different pricing in different market areas could prove problematic though, especially if market areas with close proximity have significant deviations in their prices. Maybe running the computations on individual stores or market areas could provide some insight on local demand profiles and thus motivate special promotions or marketing decisions as well as assortment management decisions.

In order to get more realistic and accurate numbers on the total profitability, we would require more data on costs. The dynamic pricing model could include factors such as warehousing costs for the store inventory and the distribution centers as well as the cost of logistics both within and between the Company X's establishments. These costs are usually affected by the sell-through of items in inventory and thus could be included in the model. However, many allocate those costs somewhere else, like with the costs of the supply chain or logistics which is a totally different cost account. A choice would have to be made whether or not these costs were included in the pricing computations.

Another aspect that could be utilized in the optimization phase is further research on the price level in the market. Including extensive intelligence on competitors' prices could provide some additional constraining factors for the price optimization. Gathering such intelligence could prove very time consuming though. Furthermore, a related issue would be a manufacturer specified sales price. Some brands want to enhance their image in the market by setting restrictions on the pricing of their products. Such restrictions could be included in the pricing model as further constraints. These issues would require additional data entries however, which means that even more computational power is needed to reference all required data when running the optimization.

The approach in this thesis was to gain insight on pricing decisions through data analysis. It was a conscious decision to leave any aspects of psychological origin outside the scope of this thesis. These issues were only addressed in the analysis of the obtained results to some extent. The assumption when observing price elasticity of demand is that the consumer behaves at least with some rationality and that things like brand or store preferences, be it the consumers' or the salesman's, do not have any effect on the purchase decisions. In reality things are different however. Many people choose to pay a little more for their purchase if the customer service is better in another store while some consumers as well as salesmen only support a specific brand. These kinds of personal preferences and psychological aspects could not be included in our research. Additionally, it was mentioned that consumers are becoming increasingly aware of the prices in the market. This includes the notion that consumers are learning the timing of various sales periods, such as summer sale, Black Friday and Christmas sales, when a large pool of products will be given a markdown. We did not address the deviations that could have been caused by consumers postponing their purchases due to their belief of upcoming discounts. Such psychological and behavioral aspects could even be included in an interesting thesis topic of their own.

5.2 Future

The future of this project is in the hands of the Company X. Provided that an investment was made into more extensive data analytics, this tool could be developed further and the quality of the computations could be enhanced. Several issues could be resolved with additional resources, such as transitioning to an R-SQL server for which the Company X has had plans in place. The quality of the obtained results is rather disappointing, but the resources needed to develop the model further were lacking at this stage. The project was performed according to the project team's initial wishes and the modeling practice was not altered during the work. In conclusion, continuing this project is an interesting prospect with serious potential in pricing analytics due to the high value of all data that is available. The possibility of future projects and development of a more complex pricing model needs to be discussed further with the Company X.

6 References

Andreyeva, T., Long, M. W., & Brownell, K. D. (2010). The impact of food prices on consumption: a systematic review of research on the price elasticity of demand for food. American journal of public health, 100(2), 216-222.

Cooper, J. C. (2003). Price elasticity of demand for crude oil: estimates for 23 countries. OPEC Energy review, 27(1), 1-8.

Cross, R. G., Higbie, J. A., & Cross, Z. N. (2011). Milestones in the application of analytical pricing and revenue management. Journal of Revenue and Pricing Management, 10(1), 8-18.

Dodds, W. B., Monroe, K. B., & Grewal, D. (1991). Effects of price, brand, and store information on buyers' product evaluations. Journal of marketing research, 307-319.

Dubin, J. A., & Rivers, D. (1989). Selection bias in linear regression, logit and probit models. Sociological Methods & Research, 18(2-3), 360-390.

Ermisch, J. F., Findlay, J., & Gibb, K. (1996). The price elasticity of housing demand in Britain: issues of sample selection. Journal of Housing Economics, 5(1), 64-86.

Espey, M., Espey, J., & Shaw, W. D. (1997). Price elasticity of residential demand for water: a meta-analysis. Water resources research, 33(6), 1369-1374.

Fox, E. J., Postrel, S., & Semple, J. H. (2009). Optimal category pricing with endogenous store traffic. Marketing Science, 28(4), 709-720.

Genchev, E., & Yarkova, J. (2010). Analysis of price elasticity of food products (for the period 1999-2009). Trakia Journal of Sciences, 8(3), 191-194.

Grewal, D., Levy, M., & Kumar, V. (2009). Customer experience management in retailing: An organizing framework. Journal of retailing, 85(1), 1-14.

Grewal, D., Ailawadi, K. L., Gauri, D., Hall, K., Kopalle, P., & Robertson, J. R. (2011). Innovations in retail pricing and promotions. Journal of Retailing, 87, S43-S52.

Hall, J. M., Kopalle, P. K., & Krishna, A. (2010). Retailer dynamic pricing and ordering decisions: category management versus brand-by-brand approaches. Journal of Retailing, 86(2), 172-183.

Hanushek, E. A., & Quigley, J. M. (1980). What is the price elasticity of housing demand?. The Review of Economics and Statistics, 449-454.

Hoch, S. J., Kim, B. D., Montgomery, A. L., & Rossi, P. E. (1995). Determinants of store-level price elasticity. Journal of marketing Research, 17-29.

Kelley, C. T. (1999). Iterative methods for optimization. Society for Industrial and Applied Mathematics.

Kirschen, D. S., Strbac, G., Cumperayot, P., & de Paiva Mendes, D. (2000). Factoring the elasticity of demand in electricity prices. IEEE Transactions on Power Systems, 15(2), 612-617.

Hughes, J. E., Knittel, C. R., & Sperling, D. (2006). Evidence of a shift in the short-run price elasticity of gasoline demand (No. w12530). National Bureau of Economic Research.

Levy, M., Grewal, D., Kopalle, P. K., & Hess, J. D. (2004). Emerging trends in retail pricing practice: implications for research. Journal of Retailing, 80(3), xiii-xxi.

Marshall, A. (2009). Principles of economics: unabridged eighth edition. Cosimo, Inc.. (The first edition from 1890, MacMillan & Co. Ltd.)

Mulugeta, D., Greenfield, J., Bolen, T., Conley, L., Cardinal Health, Pricing Analytics Team, Dublin, Ohio 43017, USA (2013). Price- and Cross-Price Elasticity Estimation using SAS. SAS Global Forum 2013, Statistics and Data Analysis, Paper 425-2013.

Nachbar, J. H. (1998). The last word on Giffen goods?. Economic Theory, 11(2), 403-412.

Niemeier, S., Zocchi, A., & Catena, M. (2013). Reshaping retail: why technology is transforming the industry and how to win in the new consumer driven world. John Wiley & Sons.

Owen, G. W. (2012). Applying point elasticity of demand principles to optimal pricing in management accounting. The International Journal of Applied Economics and Finance, 6(3), 89-99.

Schoengold, K., Sunding, D. L., & Moreno, G. (2006). Price elasticity reconsidered: Panel estimation of an agricultural water demand function. Water Resources Research, 42(9).

Sorescu, A., Frambach, R. T., Singh, J., Rangaswamy, A., & Bridges, C. (2011). Innovations in retail business models. Journal of Retailing, 87, S3-S16.

Thimmapuram, P. R., & Kim, J. (2013). Consumers' price elasticity of demand modeling with economic effects on electricity markets using an agent-based model. IEEE Transactions on Smart Grid, 4(1), 390-397.

Verhoef, P. C., Lemon, K. N., Parasuraman, A., Roggeveen, A., Tsiros, M., & Schlesinger, L. A. (2009). Customer experience creation: Determinants, dynamics and management strategies. Journal of retailing, 85(1), 31-41.

7 Appendices

A Linear regression

Linear regression is a common method in modeling the relationship between two observed variables through fitting a linear equation to the acquired data. One of the variables works as an explanatory variable, whereas the other is dependent on the explanatory one. Before attempting to fit a linear model to the observed data, it is beneficial to run statistical tests to determine whether or not there is any association between the variables to begin with. Such tests could include generating a scatter plot showing all data point pairings, or calculating a correlation coefficient. In the sense of this thesis, figure 3 shows two variables and linear regression model fitted to their observed values. In this example, the dependent variable is "Qty.Per.Day" and the explanatory variable is "Price".

The general formula of the linear regression model is presented below:

$$y_i = b_0 + b_1 x_i, \tag{A.1}$$

where y_i is the dependent variable, x_i is the explanatory variable, b_0 is a constant which can be calculated from the y-axis intersect by setting $x_i = 0$, and b_1 is the slope of the fitted linear model.

This thesis uses one of the most common methods to fit a linear regression model, which is called the least-squares regression. The name of the method reveals a lot, since the best fitting line is determined by minimizing the sum of squared vertical deviations from each data point to the fitted line. Given that our data is of the form $(x_1, y_1), ..., (x_N, y_N)$, the deviation is defined as:

$$E(b_1, b_0) = \sum_{n=1}^{N} (y_n - (b_1 x_n + b_0))^2. \tag{A.2}$$

In order to obtain the best fitting linear regression model, we need to minimize the deviation in equation (A.2). The deviation is minimized when b_0

and b_1 satisfy

$$\frac{\partial E}{\partial b_1} = 0, \; \frac{\partial E}{\partial b_0} = 0. \tag{A.3}$$

Setting equation (A.3) as $\partial E/\partial b_1 = \partial E/\partial b_0 = 0$, and solving, yields a linear system of equations. The solution for which is the pair of coefficients $b_1, \; b_0$, that produces the optimal linear equation for the data.

The data can include points that do not lie on or near the best fitting line. These points are called outliers, and can be observed after the regression model has been chosen. A residual plot shows the extent to which the observed data points deviate from the fitted model. This is a further tool to assess the linear relationship between the observations. Large deviations could indicate a possibility of non-linear relationship instead of a linear one.

B Multiple regression

Multiple regression is an advanced statistical tool, that depicts the relationship between a dependent variable and multiple explanatory, or independent, variables. Determining the relationships between the included variables will facilitate making more powerful and accurate predictions about the observed variables. The general presentation for multiple regression model with n independent variables is presented below:

$$y_i = b_0 + b_1 x_1 + b_2 x_2 + \ldots + b_n x_n + \varepsilon. \tag{B.1}$$

Here, y_i is the dependent variable that is to be predicted with the dependent variables $x_1, ..., x_n$. The coefficients related to the dependent variables are called partial regression coefficients $b_1, ..., b_n$, whereas b_0 is the intercept. The final term ε is an error term which is usually assumed to be normally distributed with a 0-mean. The multiple regression model assumes that there is no linear relationships between different independent variables and that no independent variable is constant. Furthermore, the partial regression coefficients b_n are assumed to be linear. With these assumptions, the model coefficients can be estimated with an ordinary least-squares method.

The goodness of the fitted model can be examined by running statistical tests as with simple linear regression. We have a measure called R-squared (R^2) which equals to the squared correlation between the fitted prediction of y_i and the actual observed data. The R^2-value can be computed from the deviations of the fitted model and the observed values:

$$R^2 = \frac{SSE}{SST} = 1 - \frac{SSR}{SST}. \tag{B.2}$$

In equation (B.2), the terms SSE, SST and SSR stand for different sums of squares. The following equations present how each of these values is computed:

$$SST \equiv \sum_{i=1}^{n}(y_i - \bar{y})^2; \tag{B.3}$$

$$SSE \equiv \sum_{i=1}^{n}(\hat{y}_i - \bar{y})^2; \tag{B.4}$$

$$SSR \equiv \sum_{i=1}^{n}\hat{\varepsilon}_i^2. \tag{B.5}$$

In equation (B.3), we compute the sum of squared differences between observations y_i and the total mean of observations $\bar{y}$. This is called total sum of squares. For equation (B.4), the squared difference is between the predictions $\hat{y}_i$ and the total mean of observations $\bar{y}$ and is called the explained sum of squares. The equation (B.5) simply computes a sum of squared differences between the actual data and the values predicted with the fitted model. The final equation (B.5) is called residual sum of squares, or sometimes the sum of squared errors of prediction. Inserting these into equation (B.2) yields the R^2-value for the fitted regression model.

The R^2-value does not decrease with the addition of new independent variables, which is why an adjusted R^2 value is often used. The adjusted R^2 value decreases if the added independent variable does not enhance the fitted model. The R^2-values are additional statistical coefficients that can be utilized when estimating the adequacy of the fitted model in explaining the variations in the observed data. It does not however indicate whether or not the chosen regression was the best alternative, whether or not the best combination of independent variables was chosen or whether or not there were enough data points for a reliable conclusion.

C Tables

The tables C.1 and C.2 include a large group of footwear products, from which the items in tables 1 and 2 are singled out. Table C.3 presents a collection of clothing items that yield poor cross-price elasticity values and the statistical tests suggest poor fit to the data. The model is not good enough and singularities are formed when running the computations. These tables are referenced and analyzed further in the Results section.

Table C.1: A table with a large group of footwear items' cross-price elasticities from aggregated data.

	Shoe 0	Shoe 1	Shoe 2	Shoe 3	Shoe 01	Shoe 02	Shoe 03	Shoe 04	Shoe 05	R^2	Adj.R^2
Shoe 0	-7.33	-8.45	-6.86	2.79	0.00	4.77	9.63	9.89	0.00	1.00	NA
Shoe 1	0.44	1.48	-0.50	3.49	-1.97	0.80	1.73	-9.69	0.00	0.88	0.34
Shoe 2	0.95	-6.95	-9.09	-1.44	-0.79	-4.22	13.68	10.26	0.00	0.92	0.60
Shoe 3	-6.30	-0.36	1.51	3.64	0.18	5.14	0.68	-0.53	0.00	0.84	0.40
Shoe 01	-22.06	-8.21	-0.51	-5.63	17.04	2.46	6.61	10.88	0.00	1.00	NA
Shoe 02	0.00	19.95	-0.58	4.45	0.00	-17.36	0.74	0.00	0.00	1.00	NA
Shoe 03	-18.59	-14.73	1.90	1.36	0.00	15.12	-1.40	17.99	0.00	1.00	NA
Shoe 04	5.99	-0.31	0.00	0.00	-2.48	0.00	0.05	-5.96	0.00	1.00	NA
Shoe 05	10.29	0.33	0.00	0.00	-0.92	0.00	-0.44	0.00	-7.99	1.00	NA

Table C.2: A table with a large group of footwear items' cross-price elasticity values from unaggregated data.

	Shoe 0	Shoe 1	Shoe 2	Shoe 3	Shoe 01	Shoe 02	Shoe 03	Shoe 04	Shoe 05	R^2	Adj.R^2
Shoe 0	-7.33	-8.45	-6.86	2.79	0.00	4.77	9.63	9.89	0.00	0.17	0.15
Shoe 1	0.03	4.27	-0.94	2.78	-2.65	0.86	3.49	-9.92	0.00	0.37	0.36
Shoe 2	1.99	-8.00	-10.45	-1.83	-0.47	-2.88	12.76	11.74	0.00	0.14	0.11
Shoe 3	-2.94	-1.25	-3.34	3.29	-0.38	3.27	4.68	3.54	0.00	0.32	0.30
Shoe 01	-22.06	-8.21	-0.51	-5.63	17.04	2.46	6.61	10.88	0.00	0.11	0.08
Shoe 02	0.00	19.95	-0.58	4.45	0.00	-17.36	0.73	0.00	0.00	0.05	0.02
Shoe 03	-18.59	-14.73	1.90	1.36	0.00	15.12	-1.40	17.99	0.00	0.24	0.21
Shoe 04	5.99	-0.31	0.00	0.00	-2.48	0.00	0.05	-5.96	0.00	0.42	0.40
Shoe 05	10.29	0.33	0.00	0.00	-0.92	0.00	-0.44	0.00	-7.99	0.33	0.30

Table C.3: Cross-price elasticity values for a group of clothing. Note a very poor fit and many items with no cross-price effect on others.

	SP1	SP2	SP3	SP4	SP5	SP6	SP7	R^2	Adj.R^2
SP1	4.48	0.00	0.00	0.00	0.00	0.00	0.00	1.00	NA
SP2	5.26	0.88	0.00	0.00	0.00	0.00	0.00	0.98	0.93
SP3	1.73	0.00	1.44	0.00	0.00	0.00	0.00	0.92	0.76
SP4	-0.27	-0.25	0.00	-4.08	0.00	0.00	0.00	0.99	0.98
SP5	0.00	0.00	0.00	0.00	2.04	0.00	0.00	1.00	NA
SP6	0.00	0.00	0.00	-0.50	0.00	-0.65	0.00	1.00	NA
SP7	0.00	0.00	0.00	-0.90	0.00	0.00	-1.18	1.00	NA