MASTER

The impact of physical retail locations on omnichannel shopping behavior

A transactional data analysis into shopping channel behavior of sports retailing consumers

Frieling, Rens E.J.

Award date:
2021

Eindhoven, 14-6-2021

The impact of physical retail locations on omnichannel shopping behavior

A transactional data analysis into shopping channel behavior of sports retailing consumers

R.E.J. Frieling

Master Thesis

Urban Systems & Real Estate

Eindhoven University of Technology

Eindhoven, June 2021

The impact of physical retail locations on omnichannel shopping behavior

A transactional data analysis into shopping channel behavior of sports retailing consumers

Author	R.E.J. Frieling
Student number	0956239
Master's program	Architecture, Building and Planning Track Urban Systems & Real Estate Master Thesis (45 ECTS)
Graduation Committee	dr. ir. A.D.A.M. (Astrid) Kemperman ir. A.W.J. (Aloys) Borgers dr. ir. P.E.W. (Pauline) van den Berg
Date	June 14, 2021
Place	Eindhoven

This Master's thesis has been carried out in accordance with the rules of the TU/e Code of Scientific Integrity.

TU/e EINDHOVEN UNIVERSITY OF TECHNOLOGY

Preface

This master thesis is the result of my graduation research for the master Urban Systems and Real Estate at the Eindhoven University of Technology and was completed during an internship period at sports retailer Decathlon.

The subject of this report, the impact of physical retail locations on omnichannel shopping behavior, turned out to be a challenging topic. Shopping behavior, in general, is well researched, but there was little scientific research conducted on adding shopping channels. Especially during these COVID-19 times, it is important to investigate the impact of physical retail locations on omnichannel shopping behavior and I was able to find very interesting results.

First, I would like to thank Astrid Kemperman, Aloys Borgers, and Pauline van den Berg of the Eindhoven University of Technology for their supervision, guidance, feedback, and suggestions during my research process. Without their help and commitment, this thesis would not have been established the way it is now.

Furthermore, I would like to thank Decathlon for the possibility of an internship during these extraordinary times. My word of thanks goes out to Kaj Slenter for his directions and guidance during this research. I would also like to thank Sezer Kumru for providing me with the data to complete my research.

In the end, I would like to thank my family and friends for their support, understanding, and necessary distraction during my research process. Without them, I would not have had nearly so much fun during this research.

Rens Frieling

Eindhoven, June 2021

Summary

Over the last decade, the transformation of the retail industry is impacted by the increasing digitization, recently reinforced by the Covid-19 pandemic. Retailers face more challenges to compete within a complex market and they are challenged to perform simultaneously within different retail channels. The popularity of online shopping is booming because of the possibilities to compare price and quality and the increasing information that is available and easily accessible, while the number of physical stores is decreasing in the Netherlands. The role of the physical store is changing with consumer expectations continuing to grow. Physical stores need to focus on offering unique consumer experiences rather than simply serving as a traditional sales channel to still be relevant. Opening an additional physical store has an impact on both online and offline channel performance. The objective of this research is to gain insight into the effect of opening additional physical shopping channels on omnichannel shopping and to gain insight into individual decision-making when a physical store is added. This information can be used by retailers to predict consumer behavior when the retail environment changes. The main research question is formulated as follows:

- What is the impact of opening an additional physical store on omnichannel shopping behavior in a region?

The increasing use of the internet changes the shopping behavior of consumers and increases the number of channels that consumers can use. Moving between different channels is becoming the norm for consumers and they expect a seamless experience through all channels. Omnichannel retailers need to understand their consumers and they must keep adapting to new channels with the increasing digitization in the retail industry. Retailers can work with customer cards that consumers may use to save for discounts or to keep track of all their purchases. The customer card can be used to collect widespread transaction- and socio-demographic data from consumers. This data can be used to analyze changing consumer behavior trends over time. There is still a limited number of studies that use this transactional data, while it gives opportunities to get a better picture of consumers, as consumers differ in characteristics and their shopping behavior. The literature review shows that there are contradicting results when looking at socio-demographic and spatial characteristics of consumers and their shopping channel behavior. Retailers need to have a better understanding of the behavior of different consumers to create strategies in areas with low market share and it is important to analyze these characteristics further to gain more insights into consumer behavior.

This research uses data that has been collected with customer cards from the sports retailer Decathlon. The research uses revealed transaction data to investigate the impact of opening a physical store in catchment areas where, in the last decade, a limited number of physical stores have been opened. There are four regions where two or three physical stores have been opened in the Netherlands, namely Eindhoven, Amsterdam, Utrecht, and Rotterdam. This research will investigate the effect on consumer shopping behavior when a physical store is opened in a region. There are also two regions, Kerkrade and Apeldoorn, with only one physical store. These regions will be used to investigate consumer shopping behavior when there are no additional physical store openings in the region that could have affected their performance. Customer cards of consumers living in the six regions (catchment areas) in the Netherlands will be analyzed using their transactions in the different shopping channels. These customer cards include data about the age, gender, and home address of the consumer. The main shopping channels are the physical channel, the online channel with different delivery methods, and online orders in a physical channel. In total, 166,212 customer cards including 1,256,696 transactions are analyzed using the mixed logit (ML) model. The ML model is a

random utility model and an extension of the multinomial logit (MNL) model. The model was specified to measure substitution between the physical shopping alternatives. This was useful in this research, as the ML model can take uneven competition between online and offline shopping alternatives in a region into account.

ML models are estimated for all catchment areas separately and there is one ML model that is estimated with data from all catchment areas. The results of these models show that the models perform well and can be used to predict changing shopping behavior in a catchment area. This study confirms previous findings and contributes additional evidence that suggests relations between omnichannel shopping behavior and personal characteristics. The descriptive analysis shows that there are some trends among the analyzed consumers and the transactions that they have made in the six catchment areas. The most important shopping channel is the physical channel where most of the transactions have taken place. However, the total number of transactions in the physical channel shows a small decrease in the last few years. The transactions in the online channel show an increasing trend since the start of the data analysis and even overtakes the physical channel transactions during the Covid-19 pandemic. This indicates that the online channel is becoming more important and that this trend should be followed closely in the future. In total, the number of transactions is increasing due to physical store openings and the increasing popularity of Decathlon in the Netherlands.

The model estimation has given some interesting insights into the shopping behavior of consumers. Every attribute that is included in the model estimation is significant which indicates that all attributes contribute to the utility of the shopping alternatives to a statistically significant extent. Considering the overall utility of the different shopping channels, the results of the model estimation follow the same results found in the descriptive analysis. Furthermore, the travel distance to a store has a negative utility indicating that the further the consumer must travel to a physical store, the lower the overall utility of that store. The higher the travel distance to the physical store, the more likely the consumer is to adopt online shopping. This is in line with the efficiency hypothesis presented by Farag et al. (2006). The size of the store has a positive utility value indicating that the larger the physical store, the higher the overall utility of that store. This is in line with the findings of Briesch et al. (2009). Additionally, the combination of the travel distance to a physical store and the size of the physical store is in line with the law of retail gravitation that was formulated by Reilly (1931). Additionally, the age and the gender of the individual consumers influence choice probabilities. Consumers in the age group >55 years tend to have a higher preference for purchasing products in a physical store than the in-between age group. Consumers in the age group 18 - 35 years old tend to have a lower preference for purchasing products in a physical store than the in-between age group which is in line with the findings of Farag et al. (2007) and Irawan (2015). For the gender of the consumer, men tend to have higher preferences for shopping in a physical store, and women have lower preferences for shopping in a physical store which is in line with the findings of Maat and Konings (2018).

The ML model can be used to predict the probability that a consumer chooses a shopping channel given a possible set of alternatives. The average predicted probability of the shopping alternatives has been calculated for the different catchment areas depending on the number of available shopping alternatives and the residential location of the consumer. Purchasing products online with home delivery has the highest probability in every catchment area indicating that online shopping is preferred when no store is available. The opening of the first physical store in a catchment area has a substitution effect on the other shopping alternatives which is in line with the findings of Joewono et al. (2020). When an additional physical store is added to a catchment area, the combined probability of shopping in a physical store increases, while the individual probabilities of shopping in the existing physical stores decrease indicating a substitution effect of shopping trips between physical stores.

The results of the different model estimations can be used to predict the omnichannel shopping behavior of consumers when the retail environment changes. A small case study is conducted in the catchment area of Eindhoven to predict the changing shopping behavior when a physical store is added in the city center of Helmond. The case study includes a synthetic population of 984 consumers where every 4-digit postal code included six consumers with a different age and gender combination. Using two different models – one specifically estimated for the Eindhoven catchment area and one more general model estimated for all catchment areas – the probabilities of consumers choosing alternative shopping channels were predicted for different scenarios. Probabilities at the postcode level were aggregated into overall probabilities for the entire catchment area by weighting the probabilities at the postcode level with the number of inhabitants in the postcode areas. The overall results indicate that the opening of the additional physical store in Helmond follows the same trends that were seen in the model estimation. The models show that the larger the size of the physical store in Helmond, the larger the competition with the existing physical store in the catchment area. Thus, there is a trade-off between the cannibalization of the physical stores in Best and Eindhoven and the number of consumers that will be attracted to the physical store in Helmond that would have otherwise chosen to purchase a product online or in a physical store outside the catchment area.

Research regarding changing omnichannel shopping behavior when opening physical stores is scarce. This research adds some further insights into the preferences for shopping alternatives. However, this research has some limitations which provide opportunities for future research. One of the main issues that need attention in predicting future store openings is the number of new consumers and the increase in the number of transactions that the opening of a physical store creates. The current research focuses on the shopping behavior of consumers already being a consumer taking into account the competition between shopping alternatives. It is unclear how many new consumers that have never shopped at Decathlon will be attracted by opening new channels. It would be interesting to research the increase in consumers and transactions when a new physical store opens in a catchment area.

Additionally, an issue that needs attention is the Covid-19 period that has a huge impact on omnichannel shopping behavior and might change the way that consumers shop in the future. The number of online channel transactions overtook the number of physical channel transactions for the first time in 2020 due to Covid-19. Covid-19 will still have an impact on the omnichannel shopping behavior of consumers in the future and it is, therefore, important to conduct this research again in the near future to identify whether the results in this research are still relevant.

Despite these limitations, this research has contributed to the understanding of omnichannel shopping behavior. It shows how omnichannel shopping behavior changes when additional physical stores open and how these results can be used to predict the effects of future store openings.

Table of Contents

1. INTRODUCTION

1.1 BACKGROUND AND RELEVANCE

In recent years, the retail landscape is changing with physical stores being under pressure reinforced by the Covid-19 pandemic creating challenges for researchers and professionals involved in the retailing industry. As the industry tries to create a seamless experience for consumers, the distinctions between physical stores and online shopping will vanish, turning the world into a showroom without walls (Brynjolfsson et al., 2013). Therefore, it is important to understand the impact of physical and online channels to target and position these channels better and to be able to reach new consumers and serve current consumers.

1.1.1 OMNICHANNEL

Consumers can choose between different shopping methods when they want to purchase a product. Products can be purchased in brick-and-mortar stores and online via social media, websites, and applications. These shopping methods are known as shopping channels and integrating these channels is known as omnichannel shopping. Omnichannel shopping is formally defined by Verhoef et al., (2015) as the "synergetic management of the numerous available channels and consumer touchpoints, in such a way that the consumer experience across channels and the performance over channels are optimized". Over the last decade, the transformation of the retail industry is impacted by the internet and the emergence of new online channels. The future of retailing will have to focus on integrating shopping channels into a seamless omnichannel environment for consumers (Briel, 2018). Integrating these channels requires a thorough network to link and coordinate processes, technologies, and businesses throughout all shopping channels for each product to achieve a reliable and consistent flow of material and information in multiple channels (Mirzabeiki & Saghiri, 2020). When this network performs well, consumers will have a seamless experience in both offline and online channels. Omnichannel retailing is constantly evolving due to the changing shopping environment and the emergence of new shopping channels.

1.1.2 THE CUSTOMER JOURNEY

The popularity of online shopping is booming because of the possibilities to compare price and quality and the increasing information that is available and easily accessible. Online shopping is very convenient and changes the decision-making process of the consumer, which in this research will be called 'the customer journey'. Consumers seek more than just the purchase, delivery, and consumption of products and services and with the upsurge of the internet, they have more power than ever to search for the best and unique experiences with their products and services. Consumers have experiences whenever they 'touch' any part of the product, service, brand, or organization across multiple channels and at various points in time and these are called touch points (Stein & Ramaseshan, 2010; Pantano & Viassone, 2015). The journey that a consumer goes through is formed by many touch points and every experience influences the consumers' overall perceptions of the retailer and their choices. The customer journey follows different stages and consumers act with companies through many different touch points in multiple channels and media, creating more complex customer journeys (Lemon & Verhoef, 2016).

The rise of the internet and smartphones has created a new type of consumer. Consumers are always online and have fast access to information using different channels to compare prices and evaluate products. This shopper searches, purchases, and shares products online as well as offline and they have increasing expectations for meeting their needs and desires. Mobile devices and laptops have seen an increase in popularity in recent years. In October 2019, 92% of the Dutch population owned a smartphone and 71% owned a tablet (CBS, 2019a). The tablet and smartphone ownership both show an increase of 26% since 2013 (CBS, 2019a). Furthermore, the laptop is still a very popular product, as almost 83% of the Dutch population owns a laptop (CBS, 2019a). This shows that the Dutch population is more connected than ever before which is expected to increase even more in the coming years giving opportunities for retailers to expand their digital presence.

1.1.3 RETAIL CHALLENGES

The number of channels where consumers can compare, choose from and purchase products increases rapidly and consumers have access to products anytime and anywhere (Wagner et al., 2013). Retailers face new challenges to compete within a complex market and they are challenged to perform simultaneously within different retail channels. Retailers must choose which channels they would like to use and invest in and how they integrate the different channels to create a seamless experience for the consumer. Furthermore, the increase in online channels causes a decrease in the performance of physical channels. Physical channels in the category of recreation and sports experienced a decrease in revenues of more than 30% in the period 2009 to 2019 in the Netherlands (CBS, 2020a). Within the Netherlands, the amount of non-food retail stores experienced a decrease of 15% from 2010 until 2019 (CBS, 2019b). This shows that the role of the physical store is changing with consumer expectations continuing to grow. Stores need to focus on offering unique consumer experiences rather than simply serving as a traditional sales channel to still be relevant. Due to the increasing sales in online channels, retailers must implement new methods and technologies to keep physical stores relevant and interesting and to gain insight into the shopping behavior of consumers. One of the methods that can be used to investigate consumer behavior is big data. Big data is a large volume of data that can be analyzed computationally to reveal trends and patterns. Big data can be used in the retail industry to predict spending behavior, personalize consumer experiences, forecast demand, and for many other applications. Customer cards are widely used by retailers to collect widespread transaction- and socio-demographic data from consumers. These data can be used to investigate consumer segments, understand their behavior, and offer targeted incentives to individual consumers.

Retailers also face the challenges of the Covid-19 outbreak which impacts the entire global economy since March 2020 with the closure of many retail locations due to lockdowns. Retail businesses are operating in difficult conditions, including labor supply shortages, major disruptions in supply chains and working conditions, and sometimes large spikes in demand for specific items (OECD, 2020). The spikes in demand for specific items can be seen clearly within different retail sectors showing that some sectors are affected more than others. In the EU, while the sales of non-food products dropped by 23.8%, the sales of food, beverages, and tobacco still increased by 1.2% (Eurostat, 2020).

1.2 RESEARCH GOALS

Following the background, relevance and scope, objectives, and research questions can be formulated.

1.2.1 OBJECTIVES

There is a need for research that gives insight into the consumer behavior of Dutch consumers within different shopping channels. Omnipresent retailers need to understand their consumers and they must keep adapting to new channels with the increasing digitization in the retail industry. Physical retail store performance is decreasing in the last few years in the Netherlands, while online shopping is increasing rapidly. The decrease in the number of physical stores in the Netherlands has an impact on the retail landscape. It is therefore important to analyze the impact of opening an additional physical store on the shopping region. Opening an additional store has an impact on both online and offline channel performance and it is important to gain insight into consumer behavior in different shopping channels. This research will try to predict how consumers use different shopping channels and how they integrate an additional physical store in their shopping behavior. This can be used by retailers to predict the changes in consumer behavior when adding or removing physical stores. The main objectives of this research are as follows:

1. Describe omnichannel shopping behavior throughout the customer journey.
2. Gain insight into the effect of opening an additional shopping channel on omnichannel shopping behavior.
3. Gain insight into individual decision-making when a physical channel is added.

1.2.2 RESEARCH QUESTIONS

For this research, the following research question can be derived from the objectives:

- What is the impact of opening an additional physical store on omnichannel shopping behavior in a region?

To be able to answer the main research question, the following sub-questions must be answered:

1. Which different channels can be identified within the customer journey of consumers?
2. How can the shopping behavior of consumers be measured throughout the customer journey?
3. How can the parameters that show the impact of opening an additional store be predicted based on revealed transactional data?

1.3 RESEARCH STRUCTURE

This research focuses on changing omnichannel shopping behavior when an additional physical store opens. A quantitative study approach is used which involves revealed transactional data collected in six different regions within the Netherlands using the customer cards of consumers. The customer card data is collected from the period 2014 until 2020. The research is divided into five different chapters. The second chapter starts with the literature review. The literature review is conducted on consumer behavior, the customer journey, and the impact of adding shopping channels. The third chapter introduces the methodology and includes the research design which clarifies the data collection methods and the methods that will be used for the model estimation. The fourth chapter discusses the descriptive analysis of the data and the results of the model estimation and this research ends with a conclusion in the fifth chapter describing the managerial implications, limitations, and further research.

2. LITERATURE REVIEW

This chapter gives an overview of the literature on omnichannel shopping behavior and related topics. Section 2.1 introduces the different types of consumer shopping channels and describes the increasing usage of omnichannel shopping. Section 2.2 introduces the customer journey explaining the different stages within the journey and explaining individual customer journeys. In section 2.3, the concept of consumer behavior is explained including factors that can influence consumer behavior. Section 2.4 describes the effects of opening online or offline channels on the decision-making of consumers. In section 2.5, shopping trip behavior is illustrated using the concepts of consumer choice modeling and section 2.6 gives some concluding remarks to this chapter.

2.1 CONSUMER SHOPPING CHANNELS

Single-channel shopping has been the traditional retail model where the focus is on one channel where retailers have a brick-and-mortar store (offline) or a website (online). This model is adopted the most in the retail industry, as the model focuses on minimizing expenses and increasing the sales growth. With the digitization process of consumer retail and increasing expectancy of a more convenient experience during shopping, the popularity of retail models with multiple channels increases, and it will transform the consumer retail sector (Reynolds & Sundström, 2014, Hagberg et al., 2016). Multi-channel shopping emerges with the rise of the digital revolution and this retail model focuses on offering consumers multiple ways to purchase products in both online and offline channels. This retail model increases flexibility and reaches significantly more consumers, and the consumers engage in more purchases when compared to single-channel shoppers (Pantano & Viassone, 2015). However, multi-channel shopping lacks the integration of the different channels as every shopping channel has a separate strategy. Omnichannel shopping integrates all channels of the retailer creating the same strategy for every channel. Consumers can choose between many different online and offline channels, like websites, social media, brick-and-mortar shops, webshops, smartphone applications, catalogs, and many more. The main purpose for omnichannel shopping is that consumers tend to switch between channels throughout the purchasing process and they will have the possibility to experience all channels together as one channel. Moving across multiple channels is becoming the norm for consumers who expect a consistent and seamless shopping journey across these channels and experience the omnichannel journey (Melero et al., 2016; Xu & Jackson, 2019; Blázquez, 2014). An overview of the retail models can be seen in figure 1.

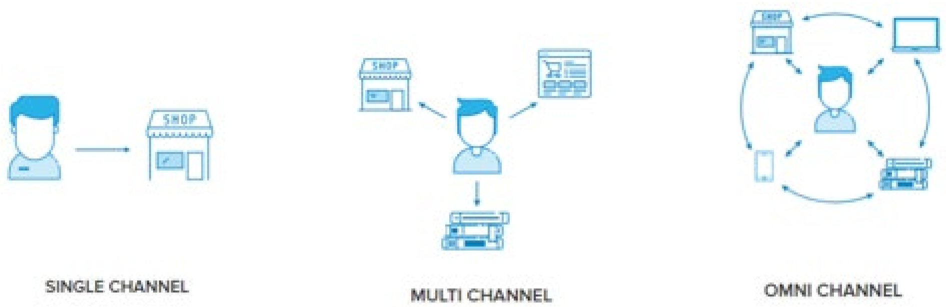

Figure 1 – Retail models (adapted from Magestore, 2020)

The literature around omnichannel shopping emphasizes several key features that have an impact on retailing activities. Omnichannel retailing focuses on integrating channels linking stock flow in different channels and creating properly communicating information flows and databases with the coordination of various channel processes and business models (Chen et al., 2016). There is a wide range of channels available and consumers can choose and use their preferred channel throughout the customer journey. Integrated omnichannel retail systems will have the benefits of more efficient product distribution and delivery, increased performance, and improved satisfaction of consumer needs (Mirzabeiki & Saghiri, 2020; Dunne & Carver, 2013). Omnichannel retail systems have many applications and will provide a lot of benefits to both the consumer and the retailer. Furthermore, data management and integration are crucial in omnichannel integration and execution (Li et al., 2015). Access to data in all available channels within a retail environment can facilitate the evaluation of processes and support decisions on the choice of channels.

There is an explosive growth of big data in the retail industry due to the increasing digital presence of retailers. Big data is a collection of data sets that become too large and become too difficult to process when using traditional database management tools (Li et al.,2015). Research using big data systems of retailers is limited but it contains valuable information for both the retailers and researchers. A purchase done by a consumer provides the retailer with a lot of information, including data about the transaction (price paid and quantity purchased), consumer data (gender, age, and location), and purchase data (online purchases with a delivery method and offline purchases). Retailers that draw effective insights from big data can make better predictions about consumer behavior and better target their consumers (Grewal et al., 2017).

Understanding consumers is important to implement targeted strategies. Surveys are common methods for understanding the behavior patterns of consumers. However, surveys are performed at a single point in time and do not follow changing trends. This makes it difficult to measure changing behavior of consumers. When consumers purchase a product, this purchase is recorded in a system where every transaction is stored (Green et al., 2020). In this way, transaction data provides a better picture of consumers, follows their purchases in real-time, and follows changing trends over time. Customer cards are widely used by retailers and consumers can use customer cards to save for discounts or to keep track of all their purchases. The retailer's purpose of the customer card is to collect widespread transaction- and socio-demographic data from consumers (Pauler & Dick, 2006). These data can be used to investigate consumer segments, understand their behavior, and offer targeted incentives to individual consumers. The literature review shows that few researchers have access to transaction data and most of the transaction data researches that are performed use data mining techniques (Böttcher et al., 2009; Hosseini et al., 2010). There are very few studies exploring the applications of transactional data using customer cards (Green et al., 2020; Aiello et al., 2020) and these are mostly performed in grocery retailing. However, these studies indicate that customer card data gives a lot of valuable insights into the knowledge and understanding of consumer behavior.

2.2 THE CUSTOMER JOURNEY

The decision-making process of consumers has changed significantly since the emergence of online shopping. Consumers experience different purchase decision-making processes, shaped by a variety of contextual, environmental, societal, and individual factors (Grewal & Roggeveen, 2020). To understand these different decision-making processes, retailers need to understand the characteristics of the journeys that consumers take. A lot of literature is published about the purchase process of consumers. Edelman and Singer (2015) describe the purchase process as a five-step model called the customer decision journey. The five-step model consists of considering, evaluating, purchasing, the loyalty loop, and a new journey. The loyalty loop is a concept in which the consumer goes through three stages: enjoy, advocate, bond. The five-step model also implies that a journey has an impact on future journeys in which past experiences impact current and future experiences. According to Lemon and Verhoef (2016), the consumer decision process is divided into three stages. The first stage is the pre-purchase stage, followed by the purchase stage and the post-purchase stage. Liang and Lai

(2002) developed a five-stage consumer decision-making process that includes need recognition, searching for information, evaluation of alternatives, purchasing, and post-purchasing. The different consumer decision-making processes seem to have quite some similarities.

The five-step model defined by Edelman and Singer, (2015) shows similarities with the five-stage consumer decision-making process that is defined by Liang and Lai., (2002). They both incorporate the evaluation of alternatives and the purchase stage. The five-stage model by Liang and Lai, (2002) is more elaborate on the pre-purchase stages, splitting the considering stage of Edelman and Singer, (2015) into a need recognition stage and searching for information stage. The five-step model by Edelman and Singer, (2015) is more elaborate on the post-purchase stage, considering a loyalty loop as after-sales service and creating a new journey stage that combines past experiences with current and future experiences. The three-stage model defined by Lemon and Verhoef, (2016) tries to capture all the stages defined in the other models in three general stages, namely pre-purchase, purchase, and post-purchase. In summary, this research will follow the three-stage model designed by Lemon and Verhoef (2016) which can be seen in figure 2. The focus of this research will be on the purchase stage. The three stages will be defined below, and a range of touchpoints will be explained.

Figure 2 – The customer journey (Lemon and Verhoef, 2016)

Stage 1: Pre-purchase

During the first stage of the customer journey, the consumer is stimulated by a product or service and the consumer will search for information. Stimulation is something that is not necessarily visible, and it can happen subconsciously. Consumers can be stimulated through a search engine, word-of-mouth, website, testimonial, or even a social or blog post (Stirista, 2020). Stimulation and need recognition are two aspects that are linked with each other. Consumers can be subconsciously stimulated by an advertisement and will eventually consciously recognize a need for that product. Similarly, consumers that have recognized a need for a type of product can be stimulated by references to find a product that fits their needs. After the consumer is stimulated and has recognized a need for a type of product, they will search for information about the product. The consumer is aware of the product or brand and wants to evaluate the product or brand. They will actively research the product and will compare it against possible competitors while asking for opinions and consulting reviews. The duration of this stage is very product-dependent and can take months or seconds to complete. At the end of the first stage, the consumer will express a preference towards a specific product.

Stage 2: Purchase

The second stage of the customer journey consists of the purchase and delivery of the product. The consumer has collected all the information about the product and will make a purchase decision. The purchase stage involves purchase actions and related activities for fulfilling a transaction (Huang and Benyoucef, 2017). There are several options to choose from when purchasing a product. The product can be purchased online via different types of online channels or it can be purchased in a physical store.

Stage 3: Post-purchase

The third stage of the customer journey consists of the after-sales service. The after-sales service is important to create consumer loyalty and increase consumer satisfaction about the product they have bought. After-sales service includes the concepts of retention and advocacy (Stirista, 2020). The concept of retention includes the service provided by the retailer that will make the consumer come back again. Examples of retention can be rewarding like sending emails or ads offering coupons, discounts, or free shipping to the consumer or it can be seeking information by sending emails to ask feedback, ask questions or upsell a service to the consumer. The concept of advocacy includes the consumer who publicly supports or recommends the product or service that they have purchased. Examples of advocacy are sharing positive feedback about the product on the website, to their friends via word of mouth, social media, or review pages.

Customer experience management research is increasingly concerned with the long-term evolution of customer experience journeys across multiple service cycles (Siebert et al., 2020). Verhoef et al. (2016) introduce a process model that adds past experiences and future experiences to the customer journey. The process model is based on consumer experiences of consumers and that consumers will adapt their future behavior on past and current experiences with channels and retailers. The past experiences include previous purchases and experiences with a retailer and a shopping channel. The model states that previous experiences influence consumer behavior. If a consumer has a good experience with the online channel, they will be more likely to purchase a product again in the online channel. On the other hand, if a consumer has a bad experience with the online channel, they will be more likely to purchase a product via another channel. Past experiences have an impact on current behavior and current behavior will have an impact on future decisions. Consumers that have repeating positive experiences with a type of channel will have increased loyalty to that channel (Gensler et al., 2012). This means that if a consumer has positive experiences with purchasing a product via a mobile application, the consumer will have a strong intention to use this channel again if this is possible.

2.3 CONSUMER BEHAVIOR

Consumer behavior is the study of consumers and the processes they use to make certain decisions. Hoyer et al. (2010) define consumer behavior as the totality of consumers' decisions concerning the acquisition, consumption, and disposition of goods, services, time, and ideas. This research focuses on the acquisition behavior of consumers. Consumers have many channels to choose from throughout their decision process which makes consumer behavior more versatile but also more complex. Retailers need to have a better understanding of consumer decision processes to increase consumer satisfaction and improve retail performance (Puccinelli et al., 2009). The consumer decision process depends on several aspects that can influence the behavior of the consumer.

There is quite some research on the aspects that influence consumer shopping behavior focusing on sociodemographic and spatial factors. Sociodemographic factors include factors like age, gender, education, and income. First, several studies indicate that the age of the consumer harms online shopping (Farag et al., 2007; Irawan, 2015; Maat and Konings, 2018). These studies argue a linear relationship between the age of the consumer and online shopping, and it shows that the higher the age of the consumer, the lower the probability of online shopping. On the other hand, Farag et al. (2006) and Zhen et al. (2018) indicate that there is a nonlinear relationship between age and online shopping. The research argues that the probability of online shopping increases until the age of 33 and decreases after the age of 33. Secondly, several studies argue that the gender of the consumer has an impact on online shopping but there are contradicting results. On the one hand, Maat and Konings (2018) argue that women are more likely to adopt online shopping than men. However, Farag et al. (2006, 2007) argue that men are more inclined to purchase articles online. On the other hand, Lee et al. (2015) indicate that there is no significant impact of gender on online shopping. Third, several studies argue that the education level of consumers positively influences the engagement of online shopping (Farag et al., 2006, Farag et al., 2007; De Blasio, 2008; Zhen et al., 2018). The higher the education level of the consumer, the more they are inclined to purchase articles online. Fourth, some studies indicate that the income of consumers has a positive effect on the engagement of online shopping (Farag et al., 2007; De Blasio, 2008; Zhen et al., 2018). Consumers with a higher income tend to engage more in purchasing articles online.

Next to the sociodemographic factors, consumer behavior can also be influenced by spatial factors. Rural markets tend to be different from urban markets in several ways. Two hypotheses try to capture the influence of spatial factors on consumer behavior which are called the innovation diffusion hypothesis and the efficiency hypothesis. The efficiency hypothesis states that rural consumers tend to purchase more online as they have longer travel times to physical stores (Farag et al., 2006). Demographically, rural areas have lower population densities and rural markets are served by fewer and smaller physical stores that have lower product variety (Clarke et al., 2015; Kirby-Hawkins et al., 2018). This creates lower product availability within the direct environment of the consumer which leads to longer travel times. In the case of grocery shopping, consumers are also more likely to purchase groceries online when they live further away from the supermarket (Kirby-Hawkins et al., 2018). Secondly, the innovation diffusion hypothesis states that urban residents are more likely to adopt online shopping, as they tend to have a higher education level and they use the internet more often for other purposes (Farag et al., 2006). Sinai and Waldfogel (2004) indicate that there are high correlations between online retailing for clothes and books and access to the nearest physical store. De Blasio (2008) also argues that urban consumers use online banking more frequently than rural consumers. Additionally, Patel et al. (2015) indicate different results when looking at decisions to purchase products in physical stores in urban and rural areas. Their findings suggest that urban consumers shop more frequently in physical stores when purchasing clothing and jewelry, while rural consumers shop more frequently in physical stores for sports goods and interior products. These researches show that there are contradicting results in different retail sectors when focusing on spatial differences.

2.4 OPENING ONLINE AND OFFLINE CHANNELS

The retail sector is changing rapidly with the increasing digital presence of stores. The digitization process has an impact on physical stores with implications extending far beyond e-commerce (Pauwels et al., 2011). Mobile applications are used in physical stores to compare prices, evaluate products, compare health, environmental and ethical considerations, in addition to facilitating payment or creating shopping lists (Ström et al., 2014; Groß, 2015; Pantano & Priporas, 2016). In the rapidly evolving retail landscape, the question is raised whether physical stores will still be needed in the future. The digitization process will have an impact on the retail landscape, but it is less clear in what way it will have an impact. The effects of digitization can be seen by the “cannibalization” of the transfer of sales from physical stores to online (Doherty & Ellis-Chadwick, 2010; Hernant & Rosengren, 2017). According to Weltevreden and Rotem-Mindali (2007), the cannibalization of physical sales due to online channels is a process that develops over time, as consumers will increase their online shopping. Furthermore, Nierop et al. (2011) and Melis et al. (2015) argue that adding an informational website decreased consumers’ offline purchase frequency as some consumers reduced the number of shopping trips. These studies focus only on the addition of an online channel in the retail environment of a retailer. The results might be different when looking at the addition of a physical store. Firstly, online channels have different value propositions than physical stores (Reinartz et al., 2019; Verhoef et al., 2007). Online channels excel on the availability of information about products creating the possibility to search for information and compare competing products. Secondly, there are differences in the consumer decision-making process when shopping at various channels. Online channels are more used for utilitarian shopping motives, whereas retail stores are more used for hedonic shopping motives (Childers et al., 2001; Haas & Kenning, 2014).

Avery et al. (2012) make some inroads into analyzing the impact of adding a physical retail store to an online store. Their findings conclude that the impact will depend on two key aspects, namely the complementarity and transparency of the new offline channel. The complementarity of the new offline channel implies to which extent the offline channel is a substitute for the current channels. If the offline channel is a substitute for the current channels, this will result in a negative impact on the current channels. The transparency of the offline channel implies to which extent the new channel is known to consumers. If the new channel is experimental, the impact will be visible in the long term, as consumers must get used to this channel. The research suggests that adding a new channel decreases sales in current channels in the short run but increases sales in the long run.

2.5 SHOPPING TRIP BEHAVIOR

Consumer shopping destination choice is a complex behavior and it is a choice made to maximize preferences within constraints (Kitamura et al., 1998; Hägerstrand, 1970). The constraints include spatial, temporal, and personal factors. Personal constraints include preferences of individuals, attitudes, household characteristics, and destination alternatives. Spatial constraints include the separation of origin and shopping destination and include travel time and travel costs. Temporal constraints include the store opening hours and the time of day. Furthermore, decisions about shopping trip behavior can be driven by a range of factors like the attractiveness of the shopping location (size, composition, price level, and level of service), accessibility (travel costs, parking, public transport, walking, and cycling access), traveler characteristics and the nature of the trip itself (Zhu et al., 2006; Scott & He, 2012; Shobeirinejad et al., 2013). The combination of constraints and factors influences the behavior of consumers and affects shopping trips.

Consumer store choice modeling is widely used by researchers and one of the aspects that is important for store choice modeling is assortment. The most widely used theory, known as the "law of retail gravitation", states that consumers prefer larger assortments. The law of retail gravitation suggests that the probability of choosing a retail store is positively related to its size but inversely related to its distance from the shopper's home (Reilly, 1931). The model implies that consumers are willing to travel further distances to larger retail centers, as they are more attractive to consumers. Furthermore, the size of the store is a product of the number of categories offered and the number of items within each category (Levy & Weitz, 2009). Larger stores offer more product categories and have a wider variety of products within a product category. Briesch et al. (2009) analyzed five different categories of assortment and indicate that they all positively impact store choice. Households will choose stores that offer more product categories and that offer more variety of products within a product category. However, Briesch et al. (2009) also argue that travel distance harms consumer store choice which shows that the law of retail gravitation is still operative in current conditions.

With the increasing popularity of online shopping, it is also important to identify the impact on offline shopping trip behavior. Four types of impacts are investigated by Mokhtarian (2008): substitution (shopping trips are replaced by online shopping); complementary (due to online shopping, the frequency of shopping trips increases); modification (due to online shopping,

shopping trips to physical stores are not replaced but are altered); and neutrality (online shopping has no impact on shopping travel). There is quite a lot of research on the impact of online shopping on shopping trips and there are contradicting results. Some studies argue that online shopping reduces shopping trips indicating a substitutional effect (Weltevreden & Rotem-Mindali, 2009; Joewono et al., 2020; Irawan, 2015), while other studies suggest that online shopping increases shopping trips indicating a complementary effect (Lee et al., 2017; Xi et al., 2018). Furthermore, Farag et al. (2007) argue that consumers spend less time on shopping trips to physical stores when they purchase more frequently online which indicates a modification effect. Additionally, Calderwood and Freathy (2014) discuss that the frequency of shopping trips is not affected by online shopping which shows a neutrality effect.

2.6 CONCLUSION

This chapter has given insights into different concepts of consumer shopping behavior. The use of the internet changes the shopping behavior of consumers as it increases the number of channels that consumers can use. Moving between different channels is becoming the norm for consumers and they expect a seamless experience through all channels (Melero et. al, 2016; Xu & Jackson, 2019; Blázquez, 2014). Understanding these consumers in an omnichannel environment is important. Omnichannel shopping behavior will be investigated within the customer journey. The customer journey consists of three stages, defined by Lemon and Verhoef (2016), namely the pre-purchase, purchase, and post-purchase stages. This research will try to elaborate on the current literature by exploring the shopping behavior of consumers in an omnichannel environment during the purchase stage of the customer journey.

There is an explosive growth of big data in the retail industry due to the increasing digital presence of retailers. Transactional data of customer cards can be used to analyze changing consumer behavior over time. However, there are only a few studies that focus on transactional data of consumers (Green et al., 2020; Aiello et al., 2020) and these are mainly in grocery retailing. This gives opportunities to incorporate transactional data in other retail sectors to get a better picture of consumers. Consumers differ in characteristics and their behavior. The literature review shows that there are contradicting results when looking at socio-demographic and spatial characteristics of consumers and their behavior. Kirby-Hawkins et al., (2018) discuss that higher travel times to physical stores lead to more online shopping and Farag et al. (2006) discuss that urban residents shop more online due to higher education and more use of the internet. Retailers need to have a better understanding of the behavior of different consumers to create strategies for different retail markets. It is important to analyze these characteristics further to gain more insights into consumer behavior.

Furthermore, research into the effects of opening additional channels in an existing omnichannel environment is limited. The digitization process will have an impact on the retail landscape, however, it is less clear what the impact will be. The main findings by other studies conclude that there is a cannibalization effect of sales from physical stores to online with the addition of a shopping channel (Hernant & Rosengren, 2017; Irawan, 2015) or that there is a complementary effect which means that the shopping frequency will increase with the addition of a shopping channel (Lee et al., 2017; Avery et al., 2012). It is important to understand what the effect of opening a shopping channel is on the overall performance of the retailer. This research will try to gain insight into the effect of opening additional physical channels in an existing omnichannel environment.

3. METHODOLOGY

This chapter will introduce the methodology that will be used in this research describing the theoretical foundations and explaining the method of data collection. The methodology will be used to predict the choice of shopping channels when an additional physical store is opened. Section 3.1 discusses the data collection method and section 3.2 introduces the data collection locations. Sections 3.3, 3.4, and 3.5 outlines the tools that will be used to analyze the data within this research, and this chapter ends with a conclusion in section 3.6.

3.1 DATA COLLECTION

The type of data that will be used in this research is transaction data collected from Decathlon. Decathlon is one of the biggest sports retailers in the world and they are an omnipresent retailer that provides both online and offline sales possibilities. Their stores are generally located within city centers, retail parks, or peripheral locations. This research uses revealed data to investigate the impact of the opening of physical stores in a region where, in the last decade, a limited number of physical stores have been opened. There are four catchment areas where two or three physical stores have been opened in the Netherlands, namely Eindhoven, Amsterdam, Utrecht, and Rotterdam. The catchment areas of Utrecht and Amsterdam will be investigated from January 1, 2015, until January 1, 2021. The catchment areas of Eindhoven and Rotterdam will be investigated from January 1, 2014, until January 1, 2021, to investigate the change in consumer shopping behavior two years before the first store opening in these catchment areas. There are also regions, Kerkrade and Apeldoorn, with only one physical Decathlon store. These regions will be used to investigate the natural changes in consumer shopping behavior. These two physical stores have been there for a long time and there are no store openings in the region that could have affected the performance. Data is collected from a large database of transactions to gain insight into the shopping channel behavior of consumers. The data concerns the purchase history of consumers. Analyzing the consumer behavior within different channels is divided into two sections:

- Customer cards (section 3.1.1)
- Channel shopping behavior (section 3.1.2)

3.1.1 CUSTOMER CARDS

Every consumer who has purchased a product in an online channel is obliged to create a customer account to complete their purchase. This customer account is known as the customer card which can be used in every channel to purchase products and to maintain an overview of what is purchased. Few studies are exploring the applications of transactional data using customer cards (Green et al., 2020; Aiello et al., 2020), and these were primarily performed in grocery retailing and not in sports retailing. With the help of Decathlon customer card data, the transactional data of every customer card can be acquired and followed over time. The customer card includes information on the age, gender, and home address of the consumer. However, it is not obligatory to use a customer card in physical stores which means that consumers might not always use their customer card. This is reflected in the percentage of customer cards used in physical stores, with only 20% of the consumers scanning their customer cards.

3.1.2 CHANNEL SHOPPING BEHAVIOR

Consumers move between online and offline channels when purchasing products. Consumers who shop at Decathlon have a lot of purchase and delivery possibilities and the different possibilities can be seen in figure 3. Purchasing a product can be done via two main shopping channels, namely online and physical channels. Online channels have three methods of delivery which are Click and Collect, pickup points, and home delivery. Deliveries using Click and Collect will be delivered to a physical channel and deliveries using pickup points will be delivered to supermarkets or businesses that work in cooperation with the delivery company. Consumers can purchase products in physical channels and they have the option to order products online if the consumer cannot take the products home or if the products are not available in the physical store. The online order follows the same procedure as the online channel and also has the same delivery possibilities. The physical channel has a maximum of four different shopping alternatives when purchasing a product in a physical store. A catchment area has a maximum of three physical stores and consumers can also choose to shop at a physical store outside the catchment area.

The behavior of consumers can change when an additional offline channel opens in a catchment area. Mokhtarian (2008) has identified four types of impacts, namely substitution, complementary, modification, and neutrality. Online shopping is increasing in popularity and it is important to identify the impact of online shopping on shopping channel behavior. Additionally, different types of characteristics of consumers might impact consumer behavior. The literature review shows that there are contradicting results regarding the effects of socio-demographic characteristics (Farag et al., 2006; Farag et al., 2007; Zhen et al., 2018) and spatial characteristics (Farag et al., 2007; Kirby-Hawkins et al., 2018). Transaction data is available on every customer card of consumers who shop at Decathlon. Customer cards have information about the purchase and delivery of every product that is purchased using the customer card. The purchase information consists of the number of products that are purchased, the number of transactions that are made, and the amount that is spent per transaction. Using customer card characteristics (age, gender, and the home address of the consumer) and transaction data (purchase location and type of delivery), the behavior of consumers can be identified when the retail environment changes.

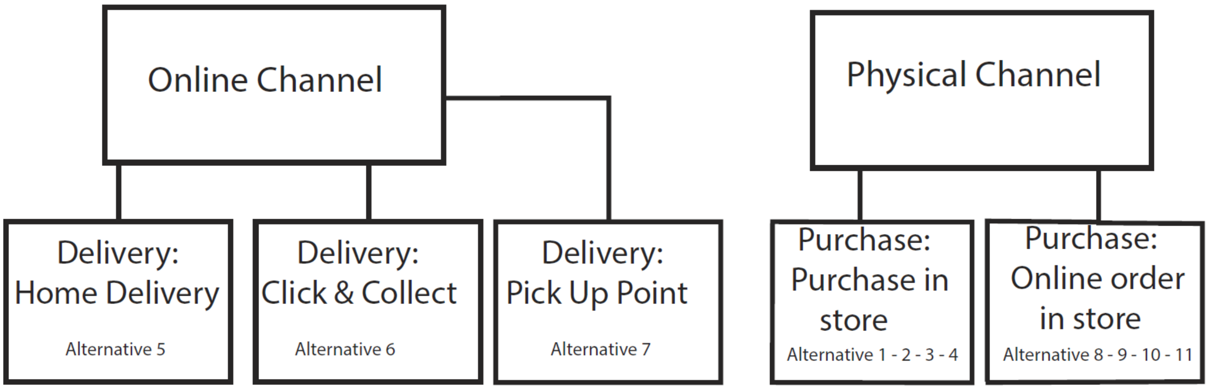

Figure 3 – Shopping channels

3.1.3 VARIABLES

The list of variables that are included in the estimation of the shopping channel choice models is based on the literature study. As a start, the sociodemographic variables age and gender are included in the dataset. The gender of the consumer is known in most of the customer cards, however, the age of the consumer is missing in 10% of the customer card information. The customer ID number, transaction ID number, and the total number of purchases are attributes that are included in the dataset to identify the transaction characteristics of every consumer. Store size, travel distance, and parking facilities are attributes that are included based on the location where the product has been purchased. Table 1 shows the collected attributes and Table 2 shows an example of what the dataset looks like.

Table 1 – Collected attributes

ID	Label	Value Name
Cust_ID	Customer ID number	Numeric
ID_Pur	Transaction ID number	Numeric
Index	Alternative	Numeric (shown in figure 3)
Pur_Tot	Total purchases by customer	Numeric
Age	Age	Numeric
Gender	Gender	Female
		Unknown
		Male
Parking	Parking facilities	No parking
		Paid parking
		Free parking
Size	Size of store (m2)	Numeric
Distance	Distance to store (km)	Numeric

Table 2 – Dataset Example

Cust_ID	ID_Pur	Index	Pur_tot	Age	Gender	Parking	Size	Distance
1	1923	3	3	42	Male	Free parking	2470	0
1	1923	4	3	42	Male	No parking	0	0
1	1923	5	3	42	Male	No parking	0	0
1	1923	6	3	42	Male	No parking	0	0
1	1923	9	3	42	Male	Paid parking	4361	0
23	82843	1	12	67	Female	Paid parking	1821	10,632
23	82843	2	12	67	Female	Free parking	4152	5,321
23	82843	3	12	67	Female	Paid parking	4170	0
23	82843	4	12	67	Female	No parking	0	0
23	82843	5	12	67	Female	No parking	0	0
23	82843	6	12	67	Female	No parking	0	0
23	82843	7	12	67	Female	Paid parking	1821	10,632
23	82843	8	12	67	Female	Free parking	4152	5,321
23	82843	9	12	67	Female	Free parking	3520	0
23	13	1	12	67	Female	Paid parking	1821	10,632
23	13	3	12	67	Female	Paid parking	2919	0
23	13	4	12	67	Female	No parking	0	0
23	13	5	12	67	Female	No parking	0	0
23	13	6	12	67	Female	No parking	0	0
23	13	7	12	67	Female	Paid parking	1821	10,632
23	13	9	12	67	Female	Free parking	2470	0

3.2 SHOPPING LOCATIONS

Data is collected from consumers that live in the catchment area of eleven Decathlon stores. The catchment area is the area from which the Decathlon attracts their consumers that use the physical stores in the catchment area and the catchment areas are defined by Decathlon. Figure 4 shows all the Decathlon stores that are situated in the Netherlands with the eleven Decathlon stores investigated in this research encircled in red:

- Decathlon Eindhoven (section 3.2.1)
- Decathlon Best (section 3.2.1)
- Decathlon Amsterdam Arena (section 3.2.2)
- Decathlon Amsterdam Kinkerstraat (section 3.2.2)
- Decathlon Amsterdam Noord (section 3.2.2)
- Decathlon Utrecht Vredenburg (section 3.2.3)
- Decathlon Utrecht The Wall (section 3.2.3)
- Decathlon Rotterdam Coolsingel (section 3.2.4)
- Decathlon Rotterdam Alexandrium (section 3.2.4)
- Decathlon Kerkrade (section 3.2.5)
- Decathlon Apeldoorn (section 3.2.6)

Figure 4 – Decathlon stores in the Netherlands

3.2.1 CATCHMENT EINDHOVEN

There are two Decathlon stores located within the catchment area of Eindhoven. The first store that opened was Decathlon Best which is in a peripheral location in the North of Eindhoven. The store consists of a 5,876 m^2 Gross Leasable Area (GLA) and it opened on the 21st of March 2012. The store is located next to a highway junction and it has free parking for visitors. The second store that opened was Decathlon Eindhoven which is located within the city center of Eindhoven. The store consists of 1,145 m^2 GLA and it opened on the 19th of October 2016. The store is close to multiple paid parking garages and it is located within walking distance of the central train station in Eindhoven. The catchment area of the Eindhoven catchment area, which can be seen in figure 5, contains approximately 800,000 inhabitants.

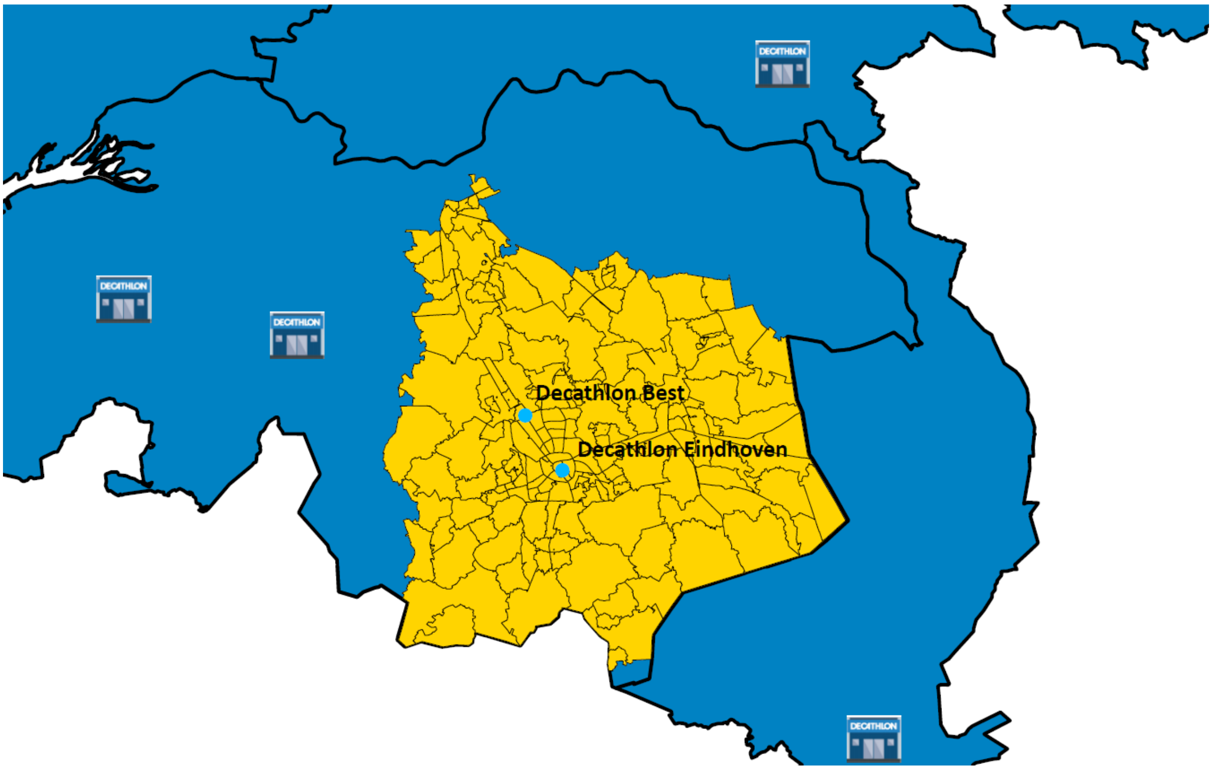

Figure 5 – Catchment area Eindhoven

3.2.2 CATCHMENT AMSTERDAM

The catchment area of Amsterdam consists of three Decathlon stores. The first store is Decathlon Amsterdam Arena which is in a shopping center on the outskirts of Amsterdam. The store consists of 3,904 m^2 GLA and it opened on the 1st of January 2000. The store is located close to a highway junction and a train station and there are paid parking garages close to the store. The second store is Decathlon Amsterdam Noord which is in a shopping center on the outskirts of Amsterdam. The store consists of 3,352 m^2 GLA and it opened on the 17th of October 2019. The store is located close to a highway junction and there are paid parking garages close to the store. The third store is Decathlon Amsterdam Kinkerstraat which is in the city center of Amsterdam. The store consists of 250 m^2 GLA and it opened on the 15th of April 2020. The store is located close to a tram station and there are paid parking garages close to the store. The catchment area of the Amsterdam catchment area, which can be seen in figure 6, contains approximately 1,200,000 inhabitants.

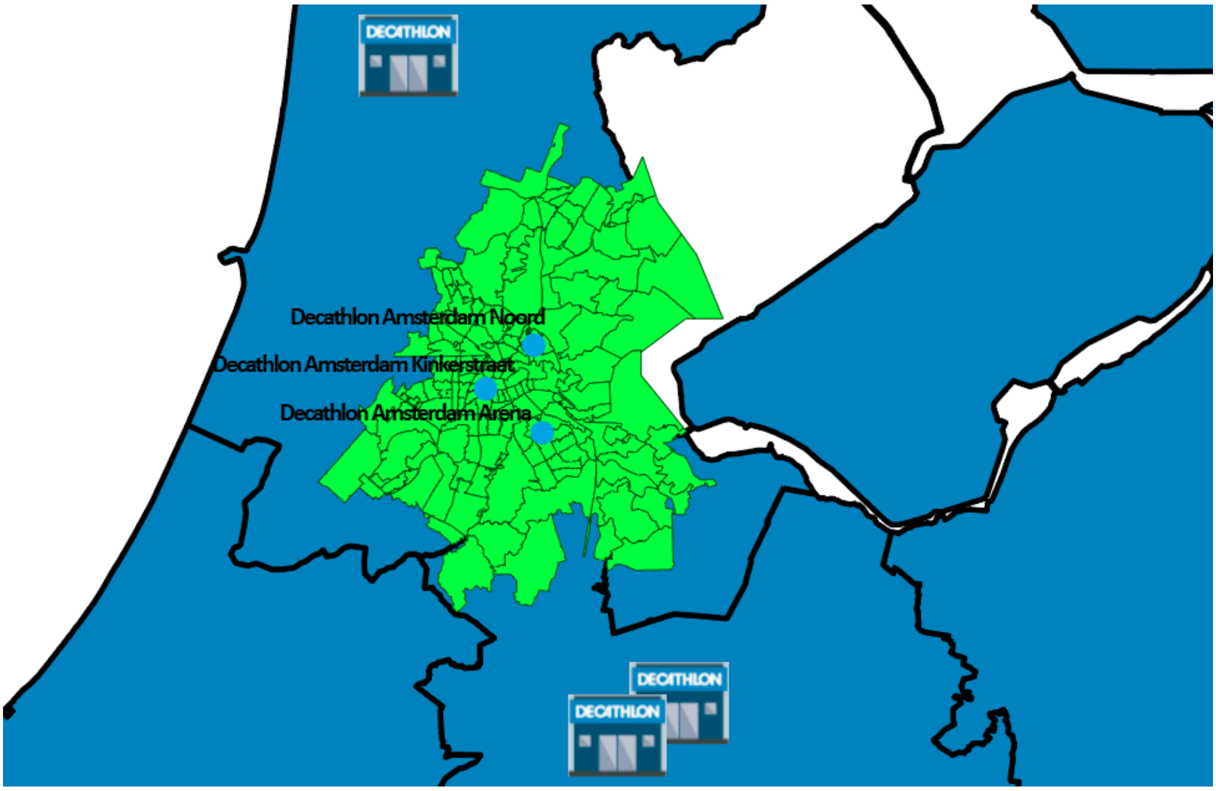

Figure 6 – Catchment area Amsterdam

3.2.3 CATCHMENT UTRECHT

There are two Decathlon stores located within the catchment area of Utrecht. The first store is Decathlon Utrecht The Wall which is in a peripheral location in the East of Utrecht. The store consists of 4,629 m^2 GLA and it opened on the 28th of November 2018. The store is located next to a highway junction and it has paid parking for visitors. The second store is Decathlon Utrecht Vredenburg which is in the city center of Utrecht. The store consists of 3,832 m^2 GLA and it opened on the 16th of October 2019. The store is located close to the central train station of Utrecht and the store is close to multiple paid parking garages. The catchment area of the Utrecht catchment area, which can be seen in figure 7, contains approximately 800,000 inhabitants.

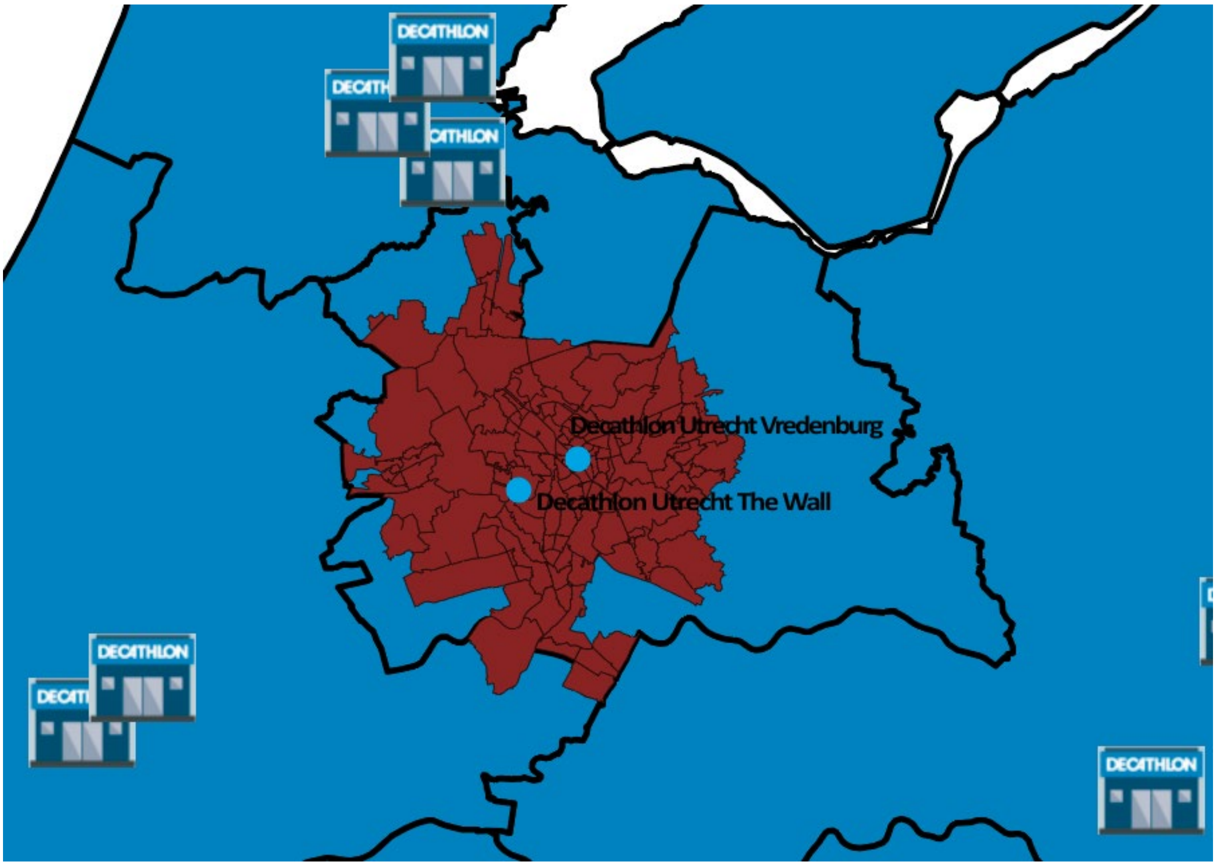

Figure 7 – Catchment area Utrecht

3.2.4 CATCHMENT ROTTERDAM

The catchment area of Rotterdam consists of two Decathlon stores. The first store is Decathlon Rotterdam Coolsingel which is in the city center of Rotterdam. The store consists of 6,294 m^2 GLA and it opened on the 6th of April 2016. The store is located close to the central train station of Rotterdam and there are multiple paid parking garages close to the store. The second store in Rotterdam Alexandrium is in a shopping center on the outskirts of Rotterdam. The store consists of 2,023 m^2 GLA and it opened on the 17th of October 2018. The store is located close to a train station and the shopping center has a paid parking garage. The catchment area of the Rotterdam catchment area, which can be seen in figure 8, contains approximately 1,100,000 inhabitants.

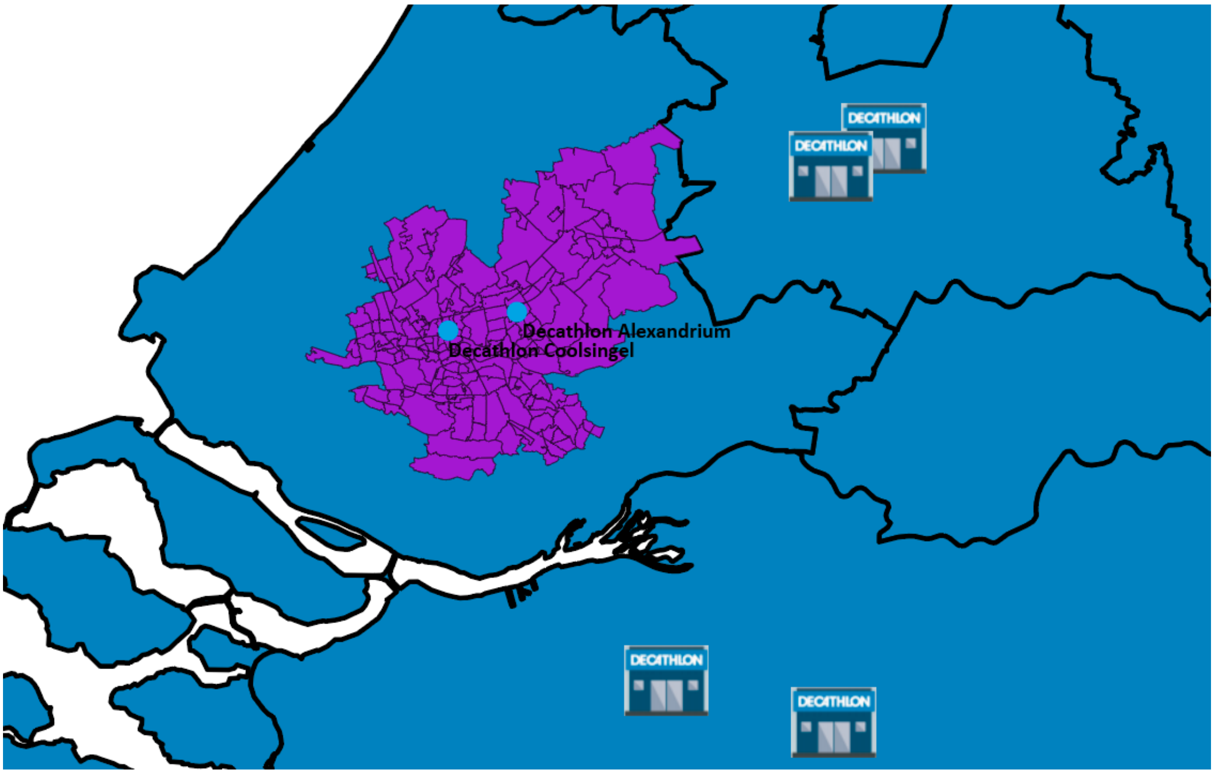

Figure 8 – Catchment area Rotterdam

3.2.5 CATCHMENT KERKRADE

The catchment area of Kerkrade consists of one Decathlon store. Decathlon Kerkrade will be used as a control group to control for changing variables. The store is in a shopping center on the outskirts of Kerkrade. The store consists of 4,832 m^2 GLA and it opened on the 13th of November 2003. The store is located close to a highway junction and there is free parking available at the store. The catchment area of the Kerkrade catchment area, which can be seen in figure 9, contains approximately 500,000 inhabitants.

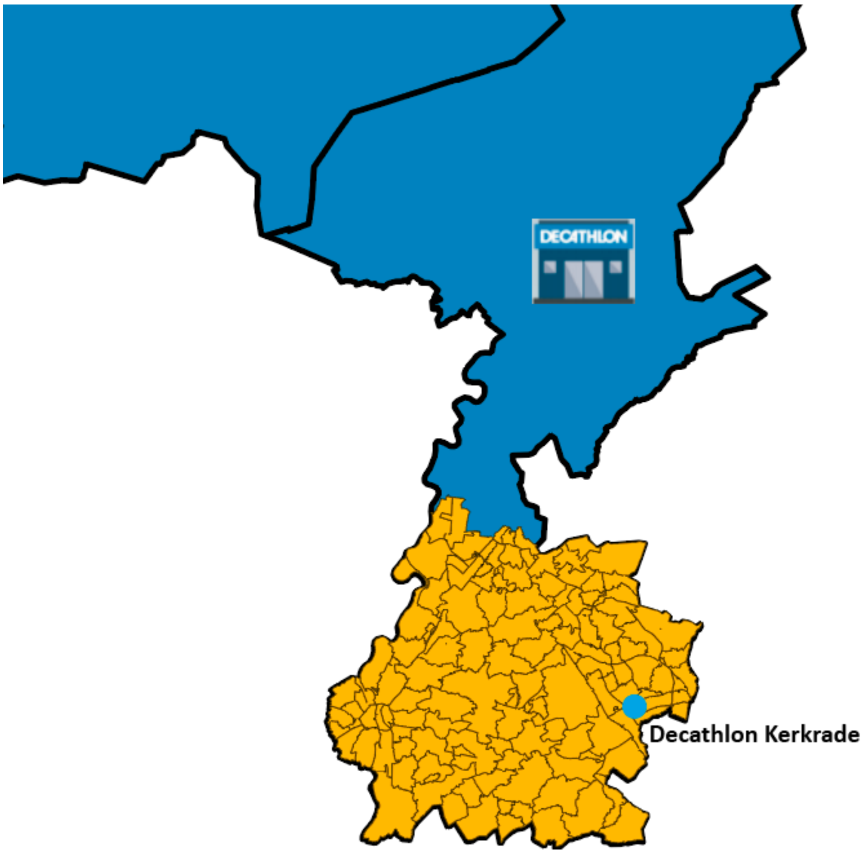

Figure 9 – Catchment area Kerkrade

3.2.6 CATCHMENT APELDOORN

The catchment area of Apeldoorn consists of one Decathlon store. Decathlon Apeldoorn will be used as a control group to control for changing variables. The store is in a shopping center on the outskirts of Apeldoorn. The store consists of 4,116 m^2 GLA and it opened on the 5th of November 2014. The store is located close to a highway junction and there is free parking available at the store. The catchment area of the Apeldoorn catchment area, which can be seen in figure 10, contains approximately 400,000 inhabitants.

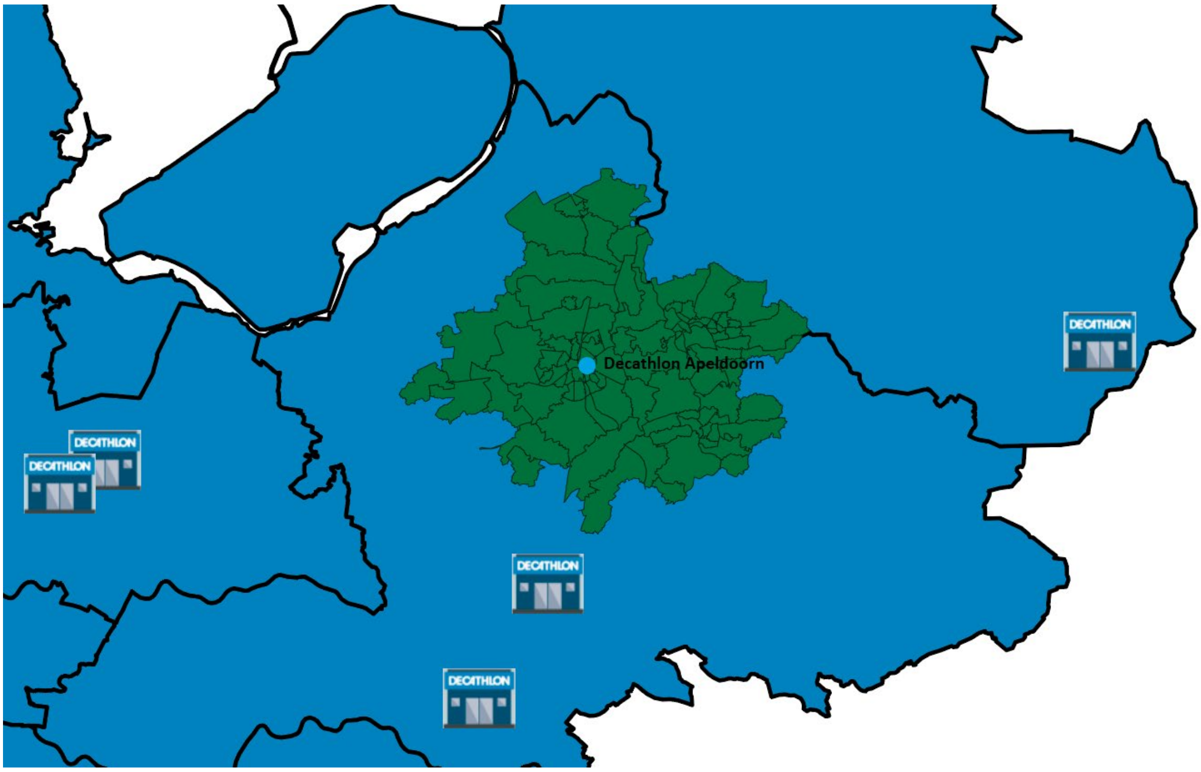

Figure 10 – Catchment area Apeldoorn

3.3 RANDOM UTILITY THEORY

The basic theory behind decision-making models is based on an important assumption. Each decision-maker is assumed to act rationally which means that the decision-maker makes choices that maximize their utility. When multiple alternatives are presented, the decision-maker makes a trade-off between the attributes of the alternatives, and the alternative with the highest utility will be chosen. Within this research, the alternatives are the different shopping channels consisting of the online channel and the physical stores and the different delivery methods which are home delivery, delivery to a pickup point, and Click and Collect. The utility (U_{ni}) for an alternative *i* for individual *n* can be divided into an observed or deterministic component (V_{ni}) and an unobserved or error component (ε_{ni}). The observed component consists of the attributes of alternatives and the relative importance of the attributes can be determined using observed choice behavior. It is almost impossible to observe all attributes for every decision-maker, so the error component captures the unmeasured factors that affect the utility of alternatives. Hensher, Rose, and Greene (2015) describe the utility function as:

$$U_{ni} = V_{ni} + \varepsilon_{ni} = \sum_{k=1}^{k} \beta_k X_{nik} + \varepsilon_{ni} \quad (1)$$

U_{ni} = Overall utility of alternative *i* for individual *n*
V_{ni} = Observed component of the utility of alternative *i* for individual *n*
ε_{ni} = Unobserved component of the utility of alternative *i* for individual *n*
β_k = Utility weight for attribute *k*
X_{nik} =score of attribute *k* of alternative *i* for individual *n*

When the utility function has been defined, the alternatives can be compared with each other. Utility theory assumes that decision-makers will maximize their utility which means that the probability that they will choose alternative *i* over any other alternative *j* can be determined. Hensher, Rose, and Greene (2015) describe the probability function as:

$$P_{ni} = prob(U_{ni} > U_{nj}\ ;\ \forall\, j \neq i, j = 1,2,\dots,J) \quad (2)$$

P_{ni} = Probability of individual *n* choosing alternative *i* out of a set of *J* alternatives

3.4 DISCRETE CHOICE MODELS

By observing many consumer choices, the probability of choosing an alternative based on the characteristics of the alternative can be derived using discrete choice models. Discrete choice models are based on random utility theory. The most used discrete choice models are multinomial logit models (MNL model) and mixed logit models (ML model) and these will be explained in this section.

3.4.1 MULTINOMIAL LOGIT MODEL (MNL MODEL)

The MNL model is widely used to analyze choice data. The MNL model assumes that the variance of the unobserved component of all alternatives is constant and that the error components are independent of each other. The MNL model estimates the mean preferences of decision-makers. The probability of an individual *n* choosing alternative *i* from the set of *J* alternatives is described as follows (Hensher, Rose & Greene 2015):

$$P_{ni} = \frac{e^{V_{ni}}}{\sum_{j=1}^{J} e^{V_{nj}}} \tag{3}$$

P_{ni}= The probability that individual *n* chooses alternative *i* from the set of *J* alternatives
$e^{V_{nj}}$ =Exponential of the observed utility of alternative *j* for individual *n*

3.4.2 MIXED LOGIT MODEL (ML MODEL)

The ML model is an extension of the MNL model. MNL models assume that there is homogeneity among shopping alternatives. In contrast to MNL models, ML models try to capture random taste variation and allow for correlation in unobserved factors over time. This is useful in this research because there might be uneven competition between shopping alternatives in a catchment area. Introducing a new physical store in a catchment area may affect the other existing physical stores relatively more than the online alternatives. ML models are also based on random utility theory, but the structural utility equation is defined differently. ML models can be specified to account for this uneven competition between the same type of alternatives (physical stores). By adding a common random component to the physical stores, the probability of choosing one of these physical stores decreases. This random component can be a random variable representing a normal distribution with a certain standard deviation. The decrease in the probability of the physical stores increases with an increasing standard deviation of the distribution.

The ML model calculates choice probabilities by repeatedly applying the MNL model. For each individual, a random value is drawn from the normal distribution and added to the utility of each physical store, and the individual choice probabilities are then calculated. This is repeated R times for each individual and the probabilities are averaged across R repetitions. By adding the same random value to all physical stores where the individual can choose from, the utilities of these alternatives become correlated and therefore their choice probabilities will decrease. For a good representation of the Normal distribution, so-called Halton draws should be used, as every draw will tend to fill in the spaces that were left empty by the previous observations. With Halton draws, a lower number of draws (R) may suffice, although many repetitions (500-1000) are still preferred.

3.5 GOODNESS OF FIT

There are several methods to measure the predictive power of a model. The goodness of fit method indicates whether a model is performing well at predicting the observed values within a dataset. The most used method is McFadden's Rho-squared, which is based on the log-likelihood. Both measures will be explained in this section.

3.5.1 LOG-LIKELIHOOD

In general, the parameters in discrete choice models are estimated by optimizing the likelihood function, which is defined by Hensher, Rose, and Greene (2015) as:

$$L = \prod_{n=1}^{N} \prod_{i=1}^{J} {P_{ni}}^{Y_{ni}} \tag{4}$$

P_{ni} =Probability of respondent *n* choosing alternative *i*
y_{ni} =1 if alternative *i* was chosen by respondent *n*, otherwise 0

To ease computations, the log-likelihood function is used:

$$LL = ln\ (L) = \ln\left(\prod_{n=1}^{N} \prod_{i=1}^{J} {P_{ni}}^{Y_{ni}}\right) = \sum_{n=1}^{N} \sum_{i=1}^{J} y_{ni}\ ln\ P_{ni} \tag{5}$$

$LL(\beta)$ =Log-likelihood function
N =Sample size

The parameters, the β_k's, are estimated by varying the parameters to maximize *LL*.

3.5.2 MC FADDEN'S RHO-SQUARED

McFadden's Rho-squared is used to measure how well the model fits the data. It compares the model with the estimated parameters and is formulated by Hensher, Rose, and Greene (2015) as follows:

$$\rho^2 = 1 - \frac{LL(\beta)}{LL(0)} \tag{6}$$

$LL(0)$ =Log-likelihood if all parameters are 0

The result of this equation gives a value between 0 and 1. The closer the value is to 1, the better the model predicts observed values. Values higher than 0.2 imply that the model is good at predicting observed values.

3.6 CONCLUSION

This chapter has introduced the methodology that will be used in this research describing the theoretical foundations and explaining the method of data collection. The research will be performed for the sports retailer Decathlon in which six catchment areas will be analyzed. Data will be collected from a large database of transactions to gain insight into the shopping channel behavior of consumers in the different catchment areas. Discrete choice models will be estimated on transaction data collected with customer cards. The theory behind the random utility theory, discrete choice models, and measuring the goodness of fit has been explained in this chapter and will be used for the model estimation in the next chapter.

4. DATA ANALYSIS AND RESULTS

The previous chapter has outlined the data collection methods and the discrete choice models that will be used to analyze the data in this chapter. This chapter will present the results of the discrete choice model estimations. The original dataset including customer card transactions is very large and includes more than four million entries. The dataset is filtered using the following conditions:

1. The consumer must make at least one transaction every two years (according to his/her customer card)
2. The consumer must make at least one transaction in a physical store from 2014 until 2020
3. 10% of the customer cards with the most transactions will be excluded

First, consumers must make at least one transaction every two years to investigate changing shopping behavior within several years. Secondly, consumers must make at least one transaction in a physical store, so consumers that only have online transactions are removed. It is not obligatory to use a customer card in a physical store, so there is too much uncertainty that these consumers do not use their customer cards when shopping in a physical store. Third, so-called outliers are removed from the data. 10% of the consumers with the highest number of transactions are removed. This will be done to exclude consumers that have an exceptionally large number of transactions. In total, 166,212 customer cards will be analyzed including 1,256,696 transactions. Table 3 gives an overview of the number of analyzed transactions and customer cards per catchment area. The catchment areas of Amsterdam and Eindhoven have the most analyzed customer cards and transactions, while the catchment area of Utrecht has the least number of analyzed transactions and customer cards.

Table 3 – Number of analyzed transactions and customer cards per catchment area

Catchment Area	Analyzed customer cards	Analyzed transactions
Amsterdam	35,046	221,413
Apeldoorn	20,908	184,200
Eindhoven	41,886	353,551
Kerkrade	23,551	205,285
Rotterdam	28,754	191,168
Utrecht	16,067	101,079
Total	166,212	1,256,696

This chapter will be divided into four sections. Section 4.1 introduces the demographic analysis of the consumers that are included in the final dataset and the descriptive analysis concerning the number of transactions that have been made per month in the different catchment areas. Additionally, section 4.2 explains the data preparation for modeling consumer behavior, presents the results of the different ML models for each catchment area, and discusses the results of the catchment areas. Section 4.3 presents the results of the model that can be used to predict the effects of new store openings within a catchment area and this chapter ends with a conclusion in section 4.4.

4.1 DESCRIPTIVE ANALYSIS

Online and offline transactions are analyzed from 2014 until 2021. Figure 11 shows the total number of transactions per month from 2014 until 2021. The figure combines all the online and offline transactions of the six catchment areas that will be analyzed. Most transactions are made in July and August and the least transactions are made in February and March. In total, the number of transactions continues to grow each year when looking at the six catchment areas. The largest growth can be seen in the years 2015 and 2016. One reason to explain the increase in transactions might be that the number of stores in the catchment areas increased over time.

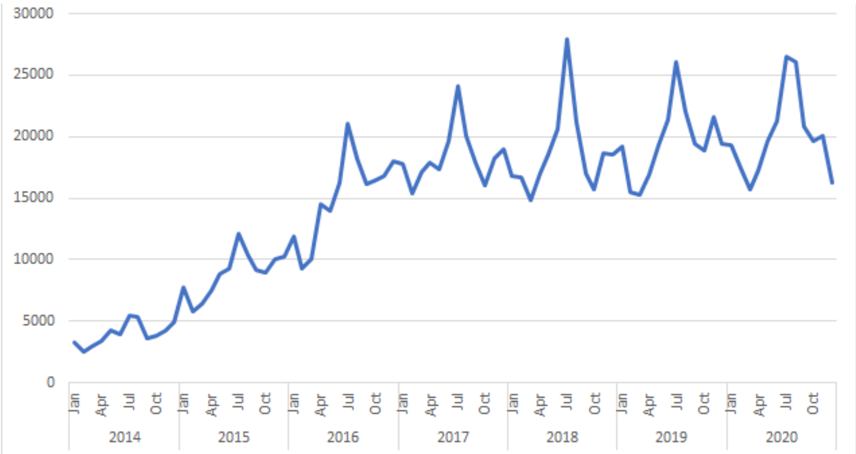

Figure 11 – Total number of transactions per month (2014 -> 2021)

Figure 12 shows the total number of transactions per channel per month for the six catchment areas. This gives a more detailed insight into the number of transactions in the last seven years. Most of the transactions have been made in physical stores, but this is declining slowly since 2018. On the other hand, transactions in online channels have been increasing since 2014. Since the start of the Covid-19 pandemic in 2020, online transactions overtook the number of physical transactions in several months during the year. This shows that the online channel is becoming an important shopping channel within the omnichannel environment of Decathlon. Online orders in physical stores have the lowest number of transactions and the number of transactions stays quite stable in the last seven years.

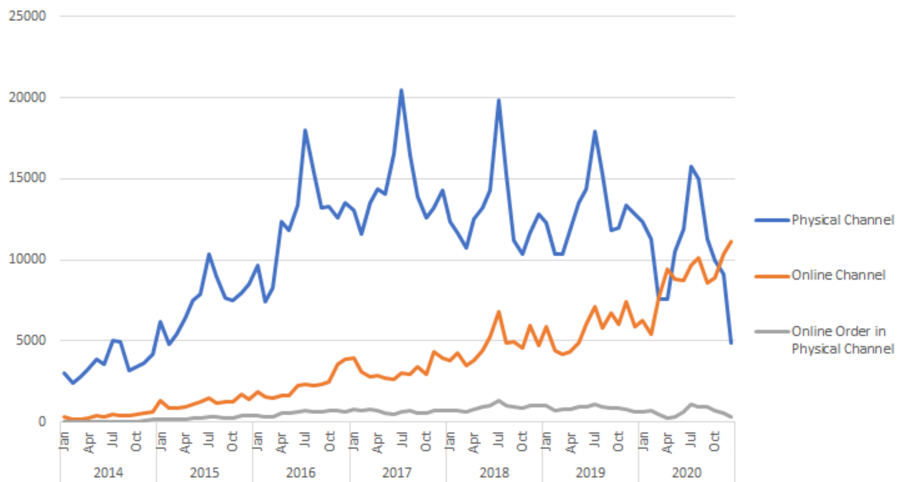

Figure 12 – Total number of transactions per channel per month (2014 -> 2021)

The demographic characteristics of age and gender have been included in the dataset. To determine whether the sample population is representative of the total population in the catchment areas, the distribution of the sample is compared to the total distribution of the population in the catchment areas. The age group <18 years old are not included in the demographic characteristics since they will not own a customer card. Table 4 shows the difference in the distribution of age. There is quite a difference between the sample population and the total population. In total, 112,350 analyzed consumers have indicated their age, and this is about 2.6% of the total population. For the different age groups, there are considerably more consumers in the age group 35-44 years old and 45-54 years old in the sample population compared to the total population. On the other hand, there are fewer consumers in the age group 18-35 years old and the age group >55 years old in the sample population compared to the total population.

Table 4 – Age sample-population

Age	Sample population	Percentage	Population catchment areas (CBS, 2020b)	Percentage
18 - 35 years	30,029	27%	1,286,029	30%
35 - 44 years	31,496	28%	667,895	15%
45 - 54 years	31,525	28%	737,290	17%
>55 years	19,300	17%	1,661,165	38%
Total	112,350	100%	4,352,379	100%
Unknown	53,862			

Table 5 shows the difference in the distribution of gender of the population that is 18 years and older. There is a difference between the sample population and the total population in the different catchment areas regarding gender. First, 3,743 consumers have not indicated their gender in their customer card which is considerably lower than the number of consumers that have not indicated their age. In total 162,469 analyzed consumers have indicated their gender in the customer card. This is about 3.7% of the total population in the catchment areas. 53% of the sample population has indicated their gender as female and 47% indicated their gender as male. This is somewhat different than the overall population in the different catchment areas which consists of 51% females and 49% males.

Table 5 – Gender sample-population (Age: >18 years old)

Gender	Sample population	Percentage	Population catchment areas (CBS, 2020b)	Percentage
Male	76,986	47%	2,139,434	49%
Female	85,483	53%	2,212,945	51%
Total	162,469	100%	4,352,379	100%
Unknown	3,743			

To determine whether the sample population is representative compared to the total population in the catchment area, the chi-square test can be used. For both the age and the gender, the p-value of the chi-square test is lower than 0.05 which means that the sample population is not representative of the population in the catchment area. The descriptive analysis can also be determined per catchment area and this will be described in the following sections.

4.1.1 CATCHMENT EINDHOVEN

The catchment area of Eindhoven consists of two physical stores and the analysis will be performed from 2014 until 2021. The data consists of 41,886 customer cards that have made 353,551 transactions. The first physical store that was opened in the catchment area of Eindhoven is located on the edge of the city of Eindhoven and has been open since 2012. The second physical store was opened in 2016 in the city center of Eindhoven. Figure 13 shows the total number of transactions per month in the catchment area of Eindhoven in the last seven years. The total number of transactions has a growing trend until 2018 and decreases in 2019 and 2020.

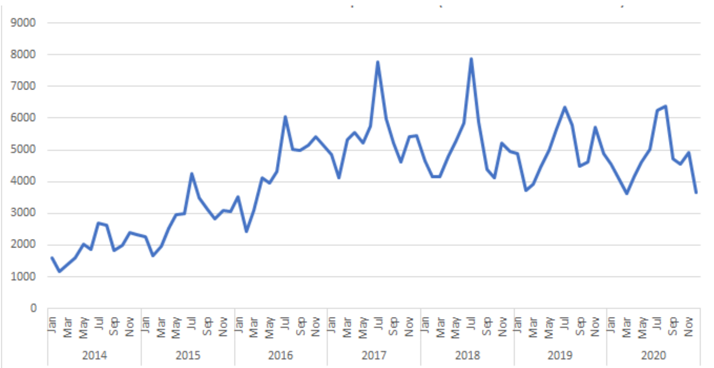

Figure 13 – Total number of transactions per month (catchment area Eindhoven)

Figure 14 gives a more detailed overview than figure 13 when adding the different purchase channels to the total number of transactions per month. The blue line indicates the transactions in physical stores, the orange line indicates the transactions in online channels by consumers living in the catchment area and the grey line indicates the online orders in physical stores. The red dotted lines indicate the opening date of the two stores that are in the catchment area. The transactions in physical stores have been increasing from 2014 until 2018. Furthermore, online channel transactions have an increasing trend since 2014. During the Covid-19 lockdown period in 2020, online channel transactions have overtaken physical channel transactions.

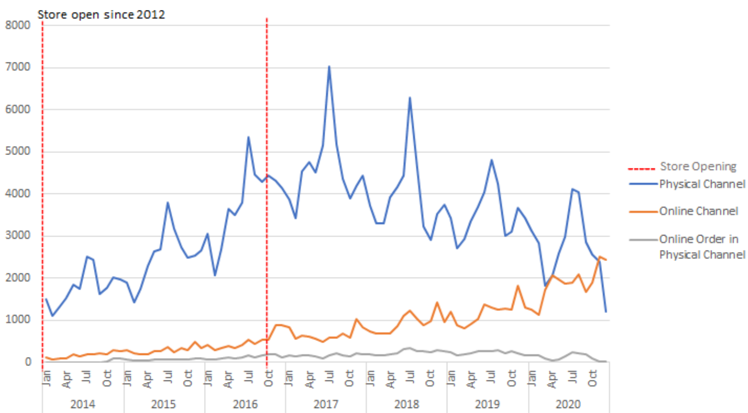

Figure 14 – Total number of transactions per channel per month (catchment area Eindhoven)

To determine whether the sample population is representative of the total population in the catchment area of Eindhoven regarding age and gender, the distribution of the sample is compared to the total distribution of the population in the catchment area. Table 6 shows the difference in the distribution of age in the catchment area of Eindhoven. There is quite a difference between the sample population and the total population. In total, 29,975 analyzed consumers have indicated their age, and this is 3.8% of the total population. For the different age groups, there are considerably more consumers in the age group 35-44 years old and 45-

54 years old in the sample population compared to the total population. On the other hand, there are fewer consumers in the age group >55 years old in the sample population compared to the total population.

Table 6 – Age sample-population catchment area Eindhoven

Age	Sample population	Percentage	Population catchment area (CBS, 2020b)	Percentage
18 - 35 years	8,113	27%	212,328	27%
35 - 44 years	8,315	28%	114,680	14%
45 - 54 years	8,637	29%	136,335	17%
>55 years	4,910	16%	328,005	41%
Total	29,975	100%	791,348	100%
Unknown	11,911			

Table 7 shows the difference in the distribution of gender of the population that is 18 years and older. There is a slight difference between the sample population and the total population in the different catchment areas regarding gender. In total 40,757 analyzed consumers have indicated their gender in the customer card. This is 5.2% of the total population in the catchment area. 53% of the sample population has indicated their gender as female and 47% indicated their gender as male. This is somewhat different than the overall population in the catchment area of Eindhoven which consists of 50% females and 50% males.

Table 7 – Gender sample-population catchment area Eindhoven (Age: >18 years old)

Gender	Sample population	Percentage	Population catchment area (CBS, 2020b)	Percentage
Male	18,969	47%	394,793	50%
Female	21,788	53%	396,555	50%
Total	40,757	100%	791,348	100%
Unknown	1,129			

To determine whether the sample population is representative compared to the total population in the catchment area, the chi-square test can be used. For both the age and the gender, the p-value of the chi-square test is lower than 0.05 which means that the sample population is not representative of the population in the catchment area.

4.1.2 CATCHMENT AMSTERDAM

The catchment area of Amsterdam consists of three physical stores and the analysis will be performed from 2015 until 2021. The data consists of 35,046 customer cards that have made 221,413 transactions. The first store was opened in Amsterdam in 2000 and is located on the edge of Amsterdam. The second store was opened in the North of Amsterdam in 2019 and the third store opened in 2020 in the city center of Amsterdam. Figure 15 shows the total number of transactions in the catchment area of Amsterdam.

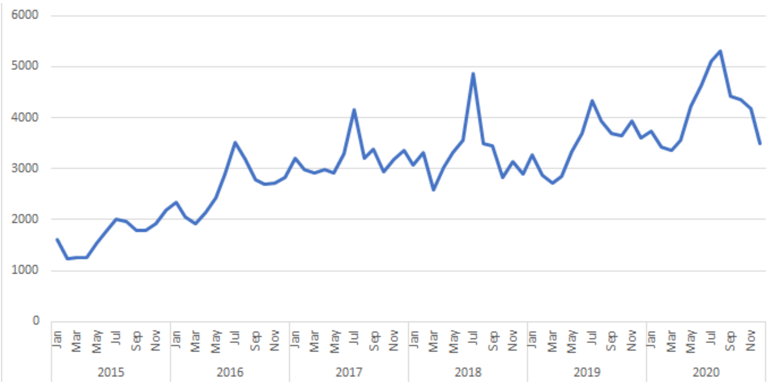

Figure 15 – Total number of transactions per month (catchment area Amsterdam)

Figure 16 shows that the physical channel is the most used shopping channel until the outbreak of Covid-19 in March 2020. The online channel has an increasing trend in the catchment area, while the offline channel is decreasing since 2017 and shows an increase at the end of 2019, possibly because of the opening of a new store.

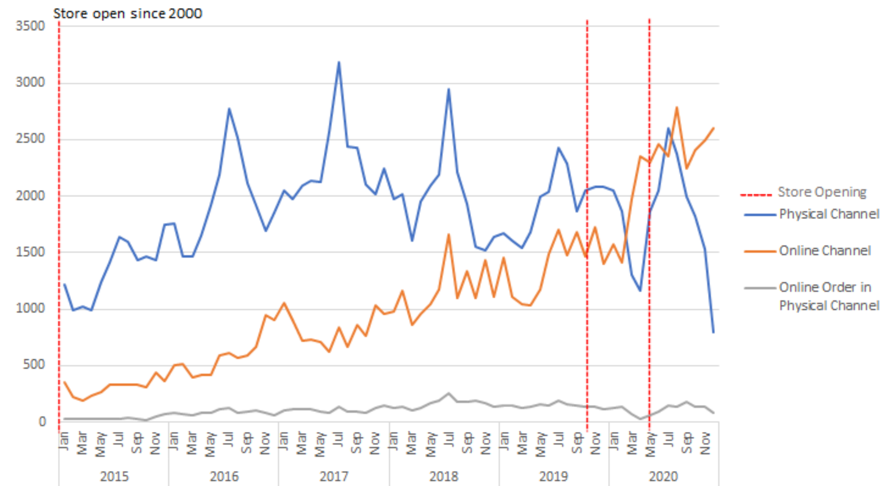

Figure 16 – Total number of transactions per channel per month (catchment area Amsterdam)

Table 8 shows the difference in the distribution of age in the catchment area of Amsterdam. In total, 21,854 analyzed consumers have indicated their age, and this is 2.2% of the total population. For the different age groups, there are considerably more consumers in the age group 35-44 years old and 45-54 years old in the sample population compared to the total population. On the other hand, there are fewer consumers in the age group 18-35 years old and >55 years old in the sample population compared to the total population.

Table 8 – Age sample-population catchment area Amsterdam

Age	Sample population	Percentage	Population catchment area (CBS, 2020b)	Percentage
18 - 35 years	6,486	30%	342,516	34%
35 - 44 years	6,513	30%	168,995	17%
45 - 54 years	6,057	28%	163,170	16%
>55 years	2,798	13%	329,965	33%
Total	21,854	100%	1,004,646	100%
Unknown	13,192			

Table 9 shows the difference in the distribution of gender of the population that is 18 years and older. In total 34,319 analyzed consumers have indicated their gender in the customer card. This is 3.4% of the total population in the catchment area. 51% of the sample population has indicated their gender as female and 49% indicated their gender as male which is the same as the overall population in the catchment area of Amsterdam.

Table 9 – Gender sample-population catchment area Amsterdam (Age: >18 years old)

Gender	Sample population	Percentage	Population catchment area (CBS, 2020b)	Percentage
Male	16,776	49%	492,264	49%
Female	17,543	51%	512,382	51%
Total	34,319	100%	1,004,646	100%
Unknown	727			

To determine whether the sample population is representative compared to the total population in the catchment area, the chi-square test can be used. For the age of the consumer, the p-value of the chi-square test is lower than 0.05 which means that the sample population is not representative of the population in the catchment area. For the gender of the consumer, the p-value of the chi-square test has a value of 0.67 which means that the sample population is representative of the population in the catchment area.

4.1.3 CATCHMENT UTRECHT

The catchment area of Utrecht consists of two physical stores and the analysis will be performed from 2015 until 2021. The data consists of 16,067 customer cards that have made 101,079 transactions. The total number of transactions per month can be seen in figure 17. The catchment area of Utrecht has two new stores that have been opened in 2018 on the edge of the city of Utrecht and in 2019 in the city center of Utrecht which explains the lower number of analyzed transactions.

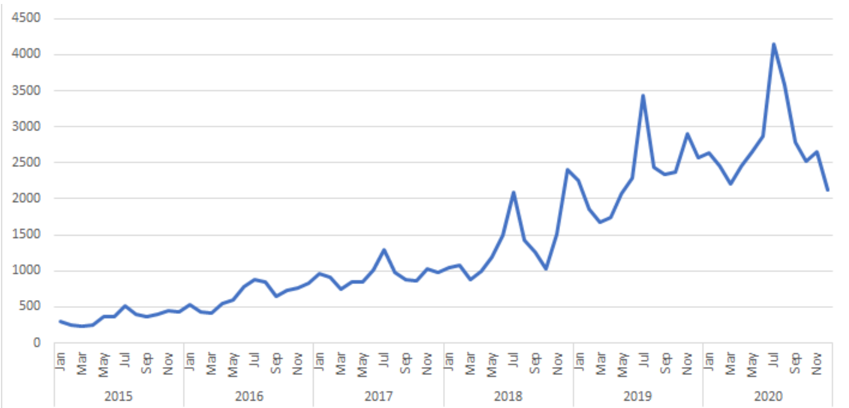

Figure 17 – Total number of transactions per month (catchment area Utrecht)

Figure 18 shows that the online channel is the most used shopping channel by the consumers living in the catchment area until the first store opened in 2018. This is also the reason why the number of transactions is rather low until 2018. With the opening of the first and the second store in the catchment area of Utrecht, the number of transactions in physical channels increases in combination with the increasing trend of the online channel.

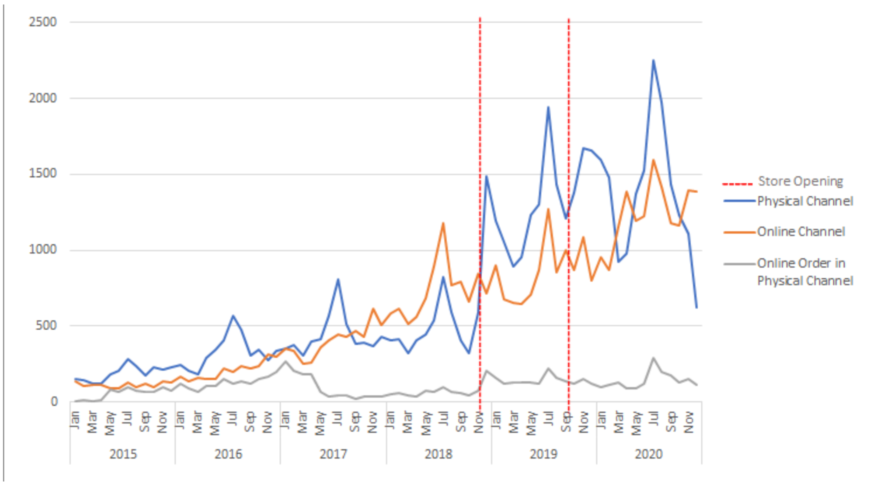

Figure 18 – Total number of transactions per channel per month (catchment area Utrecht)

Table 10 shows the difference in the distribution of age in the catchment area of Utrecht. In total, 8,505 analyzed consumers have indicated their age, and this is 1.3% of the total population. For the different age groups, there are considerably more consumers in the age group 35-44 years old, 18-35 years old, and 45-54 years old in the sample population compared to the total population. On the other hand, there are fewer consumers in the age group >55 years old in the sample population compared to the total population.

Table 10 – Age sample-population catchment area Utrecht

Age	Sample population	Percentage	Population catchment area (CBS, 2020b)	Percentage
18 - 35 years	2,913	34%	205,064	32%
35 - 44 years	2,723	32%	103,940	16%
45 - 54 years	2,179	26%	108,675	17%
>55 years	690	8%	219,620	34%
Total	8,505	100%	637,299	100%
Unknown	7,562			

Table 11 shows the difference in the distribution of gender of the population that is 18 years and older. In total 15,709 analyzed consumers have indicated their gender in the customer card. This is 2.5% of the total population in the catchment area. 53% of the sample population has indicated their gender as female and 47% indicated their gender as male. This is somewhat different than the overall population in the catchment area of Utrecht which consists of 51% females and 49% males.

Table 11 – Gender sample-population catchment area Utrecht (Age: >18 years old)

Gender	Sample population	Percentage	Population catchment area (CBS, 2020b)	Percentage
Male	7,383	47%	311,662	49%
Female	8,326	53%	325,637	51%
Total	15,709	100%	637,299	100%
Unknown	358			

To determine whether the sample population is representative compared to the total population in the catchment area, the chi-square test can be used. For the age and gender, the p-value of the chi-square test is lower than 0.05 which means that the sample population is not representative of the population in the catchment area.

4.1.4 CATCHMENT ROTTERDAM

The catchment area of Rotterdam consists of two physical stores and the analysis will be performed from 2014 until 2021. The first store was opened in the year 2016 in the city center of Rotterdam and the second store was opened in the year 2018 on the edge of the city of Rotterdam. The data consists of 28,754 customer cards that have made 191,168 transactions. Figure 19 shows the total number of transactions per month in the catchment area of Rotterdam. There is an explosive growth of transactions in the year 2016 and that there is an increasing trend of transactions until 2021.

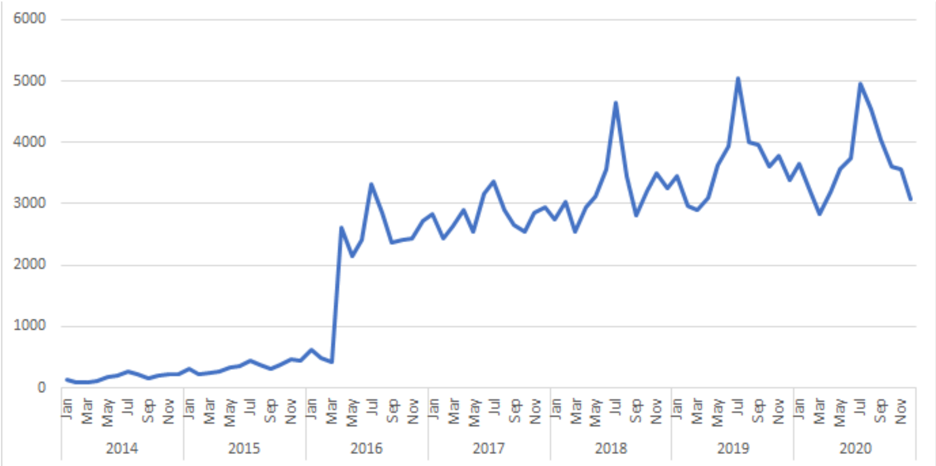

Figure 19 – Total number of transactions per month (catchment area Rotterdam)

Figure 20 shows that there is a huge increase in physical channel transactions when the first physical store is opened, but the transactions stay quite stable until the year 2020. The number of online channel transactions has also been increasing since 2016 and overtakes the physical channel transactions in 2020.

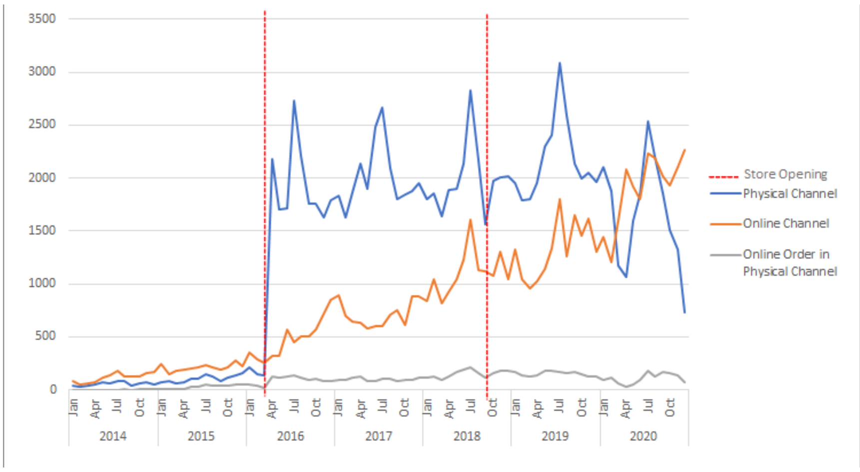

Figure 20 – Total number of transactions per channel per month (catchment area Rotterdam)

Table 12 shows the difference in the distribution of age in the catchment area of Rotterdam. In total, 18,238 analyzed consumers have indicated their age, and this is 1.6% of the total population. For the different age groups, there are considerably more consumers in the age group 35-44 years old, 18-35 years old, and 45-54 years old in the sample population compared to the total population. On the other hand, there are fewer consumers in the age group >55 years old in the sample population compared to the total population.

Table 12 – Age sample-population catchment area Rotterdam

Age	Sample population	Percentage	Population catchment area (CBS, 2020b)	Percentage
18 - 35 years	5,820	32%	341,051	30%
35 - 44 years	5,709	31%	181,055	16%
45 - 54 years	4,595	25%	195,925	17%
>55 years	2,114	12%	431,710	38%
Total	18,238	100%	1,149,741	100%
Unknown	10,516			

Table 13 shows the difference in the distribution of gender of the population that is 18 years and older. In total 28,301 analyzed consumers have indicated their gender in the customer card. This is 2.5% of the total population in the catchment area. 53% of the sample population has indicated their gender as female and 47% indicated their gender as male. This is somewhat different than the overall population in the catchment area of Rotterdam which consists of 51% females and 49% males.

Table 13 – Gender sample-population catchment area Rotterdam (Age: >18 years old)

Gender	Sample population	Percentage	Population catchment area (CBS, 2020b)	Percentage
Male	13,389	47%	562,397	49%
Female	14,912	53%	587,344	51%
Total	28,301	100%	1,149,741	100%
Unknown	453			

To determine whether the sample population is representative compared to the total population in the catchment area, the chi-square test can be used. For the age and gender, the p-value of the chi-square test is lower than 0.05 which means that the sample population is not representative of the population in the catchment area.

4.1.5 CATCHMENT KERKRADE

The catchment area of Kerkrade consists of one physical store and the analysis will be performed from 2014 until 2021. The data consists of 23,551 analyzed customer cards that have made 205,285 transactions. The store in Kerkrade is located on the edge of Kerkrade and has been open since 2003. Figure 21 shows the total number of transactions per month in the catchment area of Kerkrade. The total number of transactions has a growing trend until 2018 and decreases in the years after.

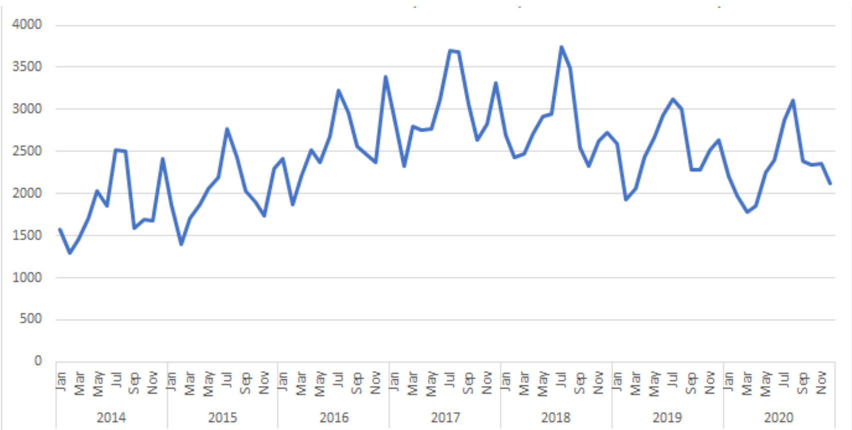

Figure 21 – Total number of transactions per channel per month (catchment area Kerkrade)

Figure 22 gives a more detailed overview than figure 21 when adding the different purchase channels to the total number of transactions per month. The physical channel is the most used shopping channel. The online channel has an increasing trend in the catchment area, while the offline channel is decreasing since 2017. This means that the decreasing trend in figure 21 is mostly caused by the decrease in physical channel transactions.

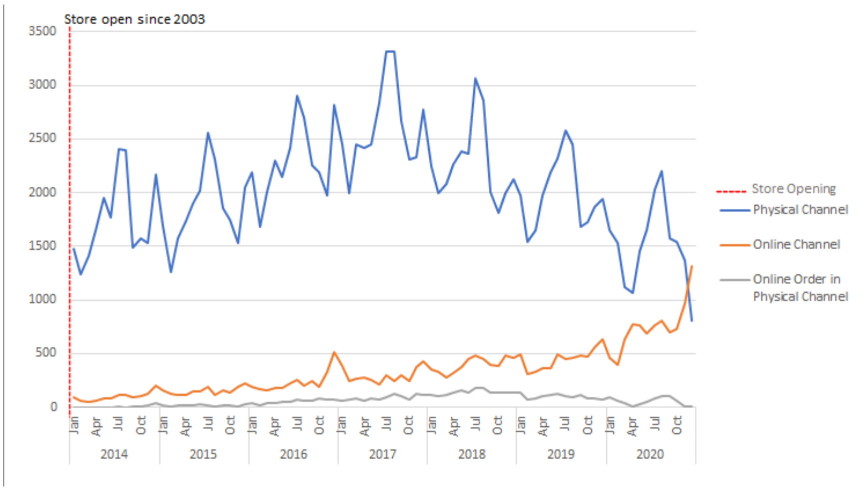

Figure 22 – Total number of transactions per channel per month (catchment area Kerkrade)

Table 14 shows the difference in the distribution of age in the catchment area of Kerkrade. In total, 18,138 analyzed consumers have indicated their age, and this is 4.2% of the total population. For the different age groups, there are considerably more consumers in the age group 35-44 years old and 45-54 years old in the sample population compared to the total population. On the other hand, there are fewer consumers in the age group 18-35 years old and >55 years old in the sample population compared to the total population.

Table 14 – Age sample-population catchment area Kerkrade

Age	Sample population	Percentage	Population catchment area (CBS, 2020b)	Percentage
18 - 35 years	2,424	13%	102,621	24%
35 - 44 years	3,970	22%	51,815	12%
45 - 54 years	5,622	31%	71,955	17%
>55 years	6,122	34%	201,750	47%
Total	18,138	100%	428,141	100%
Unknown	5,413			

Table 15 shows the difference in the distribution of gender of the population that is 18 years and older. In total 22,699 analyzed consumers have indicated their gender in the customer card. This is 5.3% of the total population in the catchment area. 49% of the sample population has indicated their gender as female and 51% indicated their gender as male. This is somewhat different than the overall population in the catchment area of Kerkrade which consists of 51% females and 49% males.

Table 15 – Gender sample-population catchment area Kerkrade (Age: >18 years old)

Gender	Sample population	Percentage	Population catchment area (CBS, 2020b)	Percentage
Male	11,508	51%	210,514	49%
Female	11,191	49%	217,627	51%
Total	22,699	100%	428,141	100%
Unknown	852			

To determine whether the sample population is representative compared to the total population in the catchment area, the chi-square test can be used. For the age and gender, the p-value of the chi-square test is lower than 0.05 which means that the sample population is not representative of the population in the catchment area.

4.1.6 CATCHMENT APELDOORN

The catchment area of Apeldoorn consists of one physical store and the analysis will be performed from 2015 until 2021. The first store was opened in the year 2014 and is located on the edge of Apeldoorn. The data consists of 20,908 customer cards that have made 184,200 transactions in the last six years. Figure 23 shows the total number of transactions per month in the catchment area of Apeldoorn. The total number of transactions follows a stable trend with peaks in the summer months.

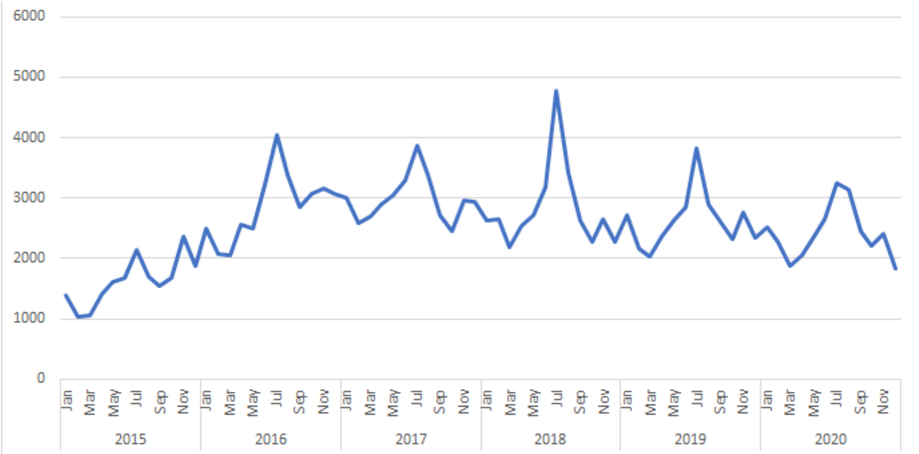

Figure 23 – Total number of transactions per channel per month (catchment area Apeldoorn)

Figure 24 shows that the transactions in physical channels have been increasing until 2018 and show a decrease in the years after that. Online channel transactions have a slowly increasing trend since the year 2015.

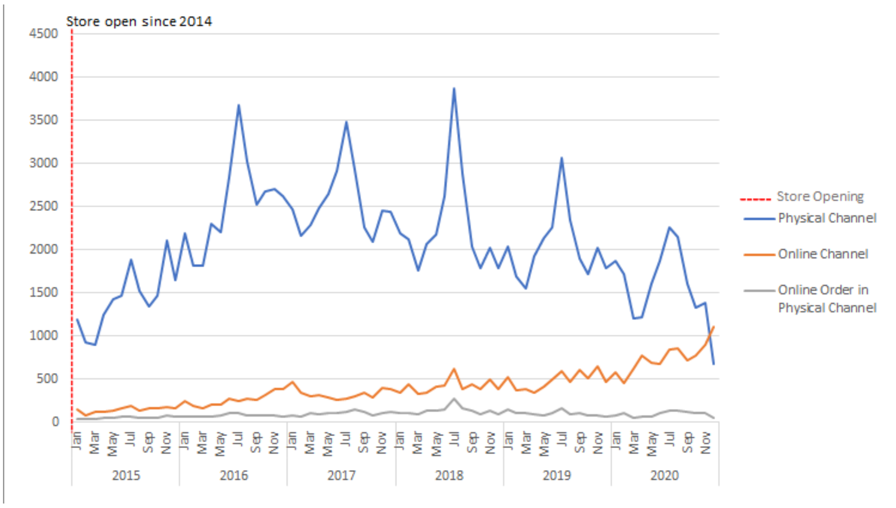

Figure 24 – Total number of transactions per channel per month (catchment area Apeldoorn)

Table 16 shows the difference in the distribution of age in the catchment area of Apeldoorn. In total, 15,640 analyzed consumers have indicated their age, and this is 4.6% of the total population. For the different age groups, there are considerably more consumers in the age group 35-44 years old, 18-35 years old, and 45-54 years old in the sample population compared to the total population. On the other hand, there are fewer consumers in the age group >55 years old in the sample population compared to the total population.

Table 16 - Age sample-population catchment area Apeldoorn

Age	Sample population	Percentage	Population catchment area (CBS, 2020b)	Percentage
18 - 35 years	4,273	27%	82,449	24%
35 - 44 years	4,266	27%	47,410	14%
45 - 54 years	4,435	28%	61,230	18%
>55 years	2,666	17%	150,115	44%
Total	15,640	100%	341,204	100%
Unknown	5,268			

Table 17 shows the difference in the distribution of gender of the population that is 18 years and older. In total 20,684 analyzed consumers have indicated their gender in the customer card. This is 6.1% of the total population in the catchment area. 53% of the sample population has indicated their gender as female and 47% indicated their gender as male. This is somewhat different than the overall population in the catchment area of Apeldoorn which consists of 51% females and 49% males.

Table 17 – Gender sample-population catchment area Apeldoorn (Age: >18 years old)

Gender	Sample population	Percentage	Population catchment area (CBS, 2020b)	Percentage
Male	8,961	43%	167,804	49%
Female	11,723	57%	173,400	51%
Total	20,684	100%	341,204	100%
Unknown	224			

To determine whether the sample population is representative compared to the total population in the catchment area, the chi-square test can be used. For the age and gender, the p-value of the chi-square test is lower than 0.05 which means that the sample population is not representative of the population in the catchment area.

4.2 MODEL ESTIMATION

The model estimation is performed using the software package called R. The goal of the model estimation is to gain insight into individual decision-making when an additional physical store is opened and the effects on omnichannel shopping behavior in a catchment area. The data will be estimated using the ML model described in chapter 3.4. The MNL model is unable to correctly account for competition between shopping alternatives of the same kind (e.g. physical stores) which is not ideal for this research. Therefore, the choice has been made to analyze the data using the ML model. The ML model will be used to estimate the datasets in the six different catchment areas and to estimate all the catchment areas in one model. The parking variable is estimated in the model with all the catchment areas because there is not enough parking variation when assessing every catchment area separately. This section starts with the data preparation and the remaining sections discuss the results of the ML models for every catchment area separately. The last section includes a discussion of the results.

4.2.1 DATA PREPARATION

In this section, the data preparation in R will be discussed explaining the variables that have been imported into R and the alteration and addition of variables. The full list of variables can be found in Appendix A. The first step is to prepare the data set that was introduced in section 3.1.3 and this can be seen in Table 18. Most of the imported variables have numeric values and the consumers' age and gender are coded into different groups. First, the 'number of alternatives' (Nalt) variable is added to show the number of shopping alternatives the consumer can choose from when deciding where to shop. Secondly, the 'choice' (Chosen) variable is a dummy coded variable that is added to indicate which shopping alternative has been chosen by the consumer. The 'total purchases by consumer' (pur_tot) variable is added to the dataset to indicate the total number of transactions that the consumer has made. Additionally, the variables age, gender, and parking have a numeric value for each category.

Consumers move between online and offline channels when purchasing products. Consumers who shop at Decathlon have a lot of purchase and delivery possibilities and the number of available possibilities differs depending on the number of shopping alternatives that are present in a catchment area. Purchasing a product can be done via two main shopping channels, namely online and physical channels. Online channels have three shopping alternatives that differ in the methods of delivery which are Click and Collect, pickup points, and home delivery. Consumers can also purchase products in physical channels and they, additionally, have the option to order products online if the consumer cannot take the products home or if the products are not available in the physical store. The online order follows the same procedure as the online channel and also has three same delivery possibilities. The physical channel has a maximum of four different shopping alternatives when purchasing a product in a physical store. A catchment area has a maximum of three physical stores and consumers can also choose to shop at a physical store outside the catchment area. Table 18 shows that there are 5 situations considering the number of alternatives that a consumer can choose from when deciding where to shop. When there are no physical stores in the catchment area, the only available shopping alternatives are online shopping (3 alternatives) and shopping at a physical store outside the catchment area (2 alternatives). When a physical store opens in a catchment area, the number of shopping alternatives increases by two, as consumers have the option to purchase a product in the new physical store and they can order a product online in the new physical store. During the Covid-19 lockdown, it was not possible to purchase a product in a physical store or to collect Click and Collect packages at physical stores. The only purchase and delivery possibilities that were available during this period were the online purchase with home delivery and online purchase with pickup point delivery.

Table 18 – Imported variables

ID	Label	Value Name	Value Numeric
Cust_ID	Customer ID number	Numeric	
ID_Pur	Transaction ID number	Numeric	
Nalt	Number of alternatives	No physical stores in catchment area	5
		1 physical store in catchment area	7
		2 physical stores in catchment area	9
		3 physical stores in catchment area	11
		Covid-19 Lockdown	2
Index	Alternative	Purchase in physical store 1	1
		Purchase in physical store 2	2
		Purchase in physical store 3	3
		Purchase in store outside the catchment area	4
		Online purchase + Home delivery	5
		Online purchase + Pickup point delivery	6
		Online purchase + Click and Collect delivery	7
		Online order in physical store 1	8
		Online order in physical store 2	9
		Online order in physical store 3	10
		Online order in physical store outside the catchment area	11
Chosen	Choice	Not chosen	0
		Chosen	1
Pur_Tot	Total purchases by customer	Numeric	
Age	Age	Unknown	0
		<35 years	1
		35-44 years	2
		45-54 years	3
		>55 years	4
Gender	Gender	Female	2
		Unknown	0
		Male	1
Parking	Parking facilities	No parking	0
		Paid parking	1
		Free parking	2

The data preparation also includes coding the imported categorical variables and adding additional categorical variables. Several methods can be used when coding variables. The simplest and most common coding method is dummy coding. Dummy coding is a way to transform categorical variables into dichotomous variables that have a value of either zero or one. A dummy variable has a value of one for one of the levels of the categorical variable and zeroes for all other levels. In a ML model, a dummy variable with a value of zero will cause its coefficient to disappear from the equation and the value of one causes the coefficient to function as a supplemental intercept, because of the multiplication of the observed value by one. Dummy variables are very useful to identify specific characteristics of subsets of alternatives. Eight dummy variables have been added to the dataset and these can be seen in Table 19.

The first five dummy variables in Table 19 focus on the different purchase and delivery possibilities for consumers who shop at Decathlon. In this way, the effects of the different purchase and delivery possibilities can be predicted. The ‘online purchase with home delivery’ (OnlH) variable is a dummy variable that has a value of one when the alternative is an online purchase with home delivery (Index = 5 in table 18). The ‘online purchase with pickup point delivery’ (OnlP) variable is a dummy variable that has a value of one when the alternative is an online purchase with pickup point delivery (Index = 6 in Table 18). The ‘online purchase with Click and Collect delivery' (OnlCC) variable is a dummy variable that has a value of one when the purchase is an online purchase with Click and Collect delivery to a store (index = 7 in Table 18). The dummy variable 'purchase in a physical store' (ST) has a value of one when the shopping alternative consists of a purchase in a physical store (Index = 1,2,3,4,8,9,10,11 in Table 18). The 'online order in a physical store with a delivery option' (Del) variable is a dummy variable that is used as to measures the effect of the delivery of products with an online order from a physical store. The dummy variable has a value of one when the index includes an online order in a physical store (Index = 8,9,10,11 in Table 18).

Furthermore, a dummy variable is used to include the distance to stores outside the catchment area, as this distance is not known. In this way, the average effect of traveling to a store outside the catchment area can be measured. The variable 'purchase in a physical store outside the catchment area' (Dist3) is a dummy variable that has a value of one when the purchase is done in a physical store outside the catchment area (Index = 4,11 in Table 18). Additionally, a dummy variable is used to see if there are differences between shopping in a store in a city center or locations on the outskirts of a city. Using this dummy variable, more detailed insight into preferred store locations can be measured. The 'physical store located in a city center' (CST) variable has a value of one when the purchase is made in a physical store located in a city center (this may be applicable if Index = 1,2,3,4,8,9,10,11 in Table 18). Besides the city center variable, a dummy variable is used to gain insight into differences between shopping behavior and parking facilities. The 'parking facilities' (ParkST12) variable includes the options of paid parking, free parking, and no parking (in case of online purchase). The parking variable has a value of zero when there are no parking facilities. The parking variable also has a value of zero when there are paid parking facilities nearby the physical store and a value of one when there are free parking facilities nearby the physical store.

Table 19 – Dummy coded variables

ID	Label	Value Name	Value Numeric
OnlH	Is the online purchase with home delivery?	Yes	1
		No	0
OnlP	Is the online purchase with pickup point delivery?	Yes	1
		No	0
OnlCC	Is the online purchase with Click & Collect delivery?	Yes	1
		No	0
ST	Is the purchase in a store?	Yes	1
		No	0
Del	Is the purchase an online order in a physical store?	Yes	1
		No	0
Dist3	Is the purchase in a store outside the catchment region?	Yes	1
		No	0
CST	Is the purchase in a store located in a city center?	Yes	1
		No	0
ParkST12	Is there free parking facilitated nearby the physical store?	Yes	1
		No	0

For this research, the imported variables age and gender need to be recoded. For the age variable, the age categories 35-44 years old and 45-54 years old are combined with the 'unknown' age category and this category will have the numeric value of zero. The age category 18-34 years old will have the numeric value of minus one and the age category >55 years old will have the numeric value of one. For the gender variable, the female category is changed to minus one and the other categories remain the same. The recoded variables and the changes in numeric values can be seen in Table 20.

Table 20 – Recoded variables

ID	Value Name	Value Numeric (Old)	Value Numeric (new)
Age	Unknown	0	0
	18-34 years	1	-1
	35-44 years	2	0
	45-54 years	3	0
	>55 years	4	1
Gender	Unknown	0	0
	Male	1	1
	Female	2	-1

Next to recoding the imported age and gender variables, some additional variables are added to the dataset. These are the variables AST and GST which are a combination of the recoded variables Age and Gender and the dummy variable ST. The variable AST is made up of the recoded variable Age multiplied by the dummy coded variable ST. The variable AST can have the numeric value minus one, zero, and one. The variable AST is equal to one if an older consumer visits a store, minus one if a young person visits a store, and zero otherwise. The variable GST is made up of the recoded variable Gender multiplied by the dummy variable ST. The variable GST can have the numeric value minus one (if a female visits a store), zero, and one (if a male visits a store). These AST and GST variables are so-called context variables and can be seen in Table 21.

Table 21 – Context variables

ID	Label	Value Numeric
AST	The Recoded variable Age multiplied by the dummy variable ST	-1
		0
		1
GST	The Recoded variable Gender multiplied by the dummy variable ST	-1
		0
		1

4.2.2 CATCHMENT EINDHOVEN

The first model estimation is performed in the catchment area of Eindhoven where there has been a store opening before the year 2014 and one in 2016. First, the MNL model for the catchment area of Eindhoven will be compared with the ML model for the catchment area of Eindhoven. The results of the MNL model are shown in Table 22 and the results of the ML model are shown in Table 23. The estimates in Table 22 and Table 23 show the effect of the attribute with a higher β-estimate indicating a stronger preference for the attribute. Every attribute that is included in the MNL and ML model has a significant parameter which indicates that all attributes contribute to the utility of the shopping alternatives to a statistically significant extent. The intercept parameter of online purchase with Click and Collect delivery is set at zero for identification. The MNL model has a pseudo-rho-squared (ρ^2) value of 0.48 indicating that the model is performing very well and is good at predicting observed values.

Table 22 – MNL model Eindhoven

ID	Label	β
ST	Purchase in a physical store	1.426 ***
OnlH	Purchase online + Home delivery	0.891 ***
OnlP	Purchase online + Pickup point delivery	-0.267 ***
Del	Online order in a physical store	-3.190 ***
SizeST	Size of the store (m2)	4.600 e−4 ***
DistST	Distance to store in catchment area (km)	-0.119 ***
Dist3	Distance to store outside catchment area (dummy)	-3.579 ***
AST	Age * purchase in a physical store	0.341 ***
GST	Gender * purchase in a physical store	0.167 ***
Significance codes: 0 "***" 0.001 "**" 0.01 "*" 0.05 "." 0.1 " " 1		
Log Likelihood model		-389700
Log Likelihood null model		-750175
pseudo rho-squared		0.48

The ML model has a pseudo-rho-squared (ρ^2) value of 0.52 indicating that the model is performing very well and is better at predicting observed choices than the MNL model. The main difference between the two models is the utility for the 'purchase in a physical store' (ST) variable. Purchasing products in a physical store has a utility of 1.426 in the MNL model and a utility of 2.356 in the ML model with a standard deviation of 1.860. The standard deviation (σ) of the physical store (ST) variable represents the common component discussed in section 3.4.2 and measures whether the physical stores compete more with each other than with the other (online) alternatives. The high standard deviation indicates that the physical stores compete more with each other. The MNL model is not able to correctly account for competition between the physical shopping alternatives and is, therefore, less accurate in the estimation of the physical store (ST) choices.

Table 23 – ML model Eindhoven

ID	Label	β	σ
ST	Purchase in a physical store	2.356 ***	1.860 ***
OnlH	Purchase online + Home delivery	0.891 ***	
OnlP	Purchase online + Pickup point delivery	-0.267 ***	
Del	Online order in a physical store	-3.190 ***	
SizeST	Size of the store (m2)	4.780 e−4 ***	
DistST	Distance to store in catchment area (km)	-0.119 ***	
Dist3	Distance to store outside catchment area (dummy)	-4.210 ***	
AST	Age * purchase in a physical store	0.341 ***	
GST	Gender * purchase in a physical store	0.167 ***	
Significance codes: 0 "***" 0.001 "**" 0.01 "*" 0.05 "." 0.1 " " 1			
Log Likelihood model			-362910
Log Likelihood null model			-750175
pseudo rho-squared			0.52

The results for the different purchase and delivery possibilities can be compared with each other. It can be said that the ‘purchase in a physical store’ (ST) variable has a positive utility equal to 2.356. The ‘online order in a physical store’ (Del) variable has a negative utility, meaning that the standard utility of that store decreases by 3.190 and becomes negative. This means that ordering in a physical store is considered less attractive than taking home the products. The ‘online purchase with home delivery’ (OnlH) variable has a positive utility equal to 0.891. The variable ‘online ordering and pickup point delivery’ (OnlP) has a negative utility which means that the consumers in the Eindhoven catchment area consider this an unattractive alternative. It also means that the ‘online purchase with Click and Collect delivery’ (OnlCC) variable has a higher utility as this parameter is kept at zero. Thus, it can be said that purchasing products online with home delivery has the highest utility compared to the other delivery methods when purchasing products online.

Furthermore, the ‘size of the physical store’ (SizeST) variable has a positive utility which means that the larger the store, the higher the overall utility of that store. The effect of store size is very small, however, the average size of a Decathlon store is almost 3000 m^2. The standard utility of a store (equal to 2.356) increases by 0.478 for every 1000 m^2, so this indicates that the size of the physical store has a considerable impact on the utility of a physical store. The ‘distance to the store’ (DistST) variable has a negative utility which means that the further the consumer must travel to the store, the lower the overall utility of that store. For every 10km traveling, the utility of a store decreases by 1.19. As the distance to stores outside the catchment area is unknown, this distance effect is measured by the ‘distance to a physical store outside the catchment area’ (Dist3) variable. Given the estimates for both distance variables (DistST and Dist3), it can be concluded that the average effect of traveling to a physical store outside the catchment area is perceived as traveling 35.3 kilometers (-4.210/-0.119) to a physical store.

Additionally, two context variables are added to the ML model in the catchment area of Eindhoven to measure the effects of age and gender on visiting a physical store. The recoded age variable (AST) has a positive utility which indicates that the age group >55 years old has a higher overall utility for shopping in a physical store compared to the age group 35-54 years old. It also indicates that the age group 18-35 years old has a lower overall utility for shopping in a physical store. The recoded gender variable (GST) also has a positive utility indicating that men tend to have a higher overall utility for shopping in a physical store compared to women.

The MNL model and the ML model can be used to predict the probability of choosing each shopping alternative based on the observed shopping data and the estimated parameters. The probability can be predicted for every transaction and the average predicted probability of the shopping alternatives can be calculated depending on the number of available shopping alternatives and the residential location of the consumers. There are three shopping alternative situations in the catchment area of Eindhoven, namely one store in the catchment area, two stores in the catchment area, and the Covid-19 lockdown period with no stores being available. Additionally, the observed probabilities of the shopping alternatives can be calculated and these can be compared to the predicted probabilities to see how well the model can predict observed probabilities. The observed probabilities have been calculated for the shopping alternative situation when two stores have been opened in the catchment area. The average predicted probabilities of the shopping alternatives using the MNL model for the three situations and the observed probabilities when two stores have been opened in the catchment area are shown in Table 24.

Purchasing products in store 1 has the highest probability when there is one store located in the catchment area. Store 1 is the physical store located in Best and has a predicted probability of 0.76. This indicates that, on average, consumers have a 76% chance of choosing the physical store in Best for their purchases when there is one store located in the catchment area. The online channel transactions (online purchase + home delivery, online purchase + pickup point delivery, and online purchase + Click and Collect delivery) have a combined probability of 21%. Within this channel, online purchase with home delivery has the highest probability of 12%. Additionally, purchasing products outside the catchment area has a probability of 3% and online orders in a physical store have a combined probability of 0.4%.

The probabilities change when the second store is opened in the catchment area. The second store is a smaller store located in the city center of Eindhoven. The probability that a consumer chooses to purchase in the store in Eindhoven is equal to 8%. The opening of the store in Eindhoven impacts the first store in Best the most, as the probability that a consumer chooses to purchase a product in Best decreases by 7%-points. However, the physical store in Best still has the highest probability with a probability of 69%. The combined probability of the online channel decreases by 1%-point with the opening of the store in Eindhoven. The probability of purchasing products outside the catchment area and the probability of ordering products online in a physical store decreases slightly.

Furthermore, the probabilities have also been calculated for the Covid-19 lockdown period in December 2020. During the Covid-19 lockdown, the only available shopping alternatives were to purchase the product online with home delivery or delivery to a pickup point. The online purchase with home delivery has the highest probability of 76% and the online purchase with pickup point delivery has a probability of 24%. Next to the predicted probabilities, Table 24 also shows the observed probabilities when there are two stores in the catchment area. The observed probabilities and the predicted probabilities have similar values, but the main difference is the observed probabilities of the physical stores. The predicted probabilities of the MNL model for the physical store in Best are 8%-points higher than the observed probabilities and the predicted probabilities for the physical store in Eindhoven are 3%-points lower than the observed probabilities. Furthermore, the predicted probability for the online orders in the physical store in Best is 90% lower than the observed probabilities.

Table 24 – Predicted probabilities MNL model Eindhoven and observed probabilities

Catchment area Eindhoven	1 store in catchment area	2 stores in catchment area	COVID-19 lockdown	Observed probabilities with 2 stores in catchment area
Purchase in store 1	0.76	0.69	-	0.61
Purchase in store 2	-	0.08	-	0.11
Purchase in store outside catchment area	0.03	0.02	-	0.03
Online purchase + home delivery	0.12	0.11	0.76	0.13
Online purchase + pickup point delivery	0.04	0.04	0.24	0.05
Online purchase + click and collect delivery	0.05	0.05	-	0.05
Online order in store 1	0.003	0.003	-	0.03
Online order in store 2	-	0.003	-	0.01
Online order in store outside catchment area	0.001	0.001	-	0.001

The average predicted probabilities of the shopping alternatives using the ML model for the three situations and the observed probabilities when two stores have been opened in the catchment area are shown in Table 25. Purchasing products in store 1 has the highest probability when there is one store located in the catchment area. Store 1 is the physical store located in Best and consumers have a 71% chance of choosing the physical store in Best for their purchases when there is one store located in the catchment area. The online channel transactions (online purchase + home delivery, online purchase + pickup point delivery, and online purchase + Click and Collect delivery) have a combined probability of 24%. Within this channel, online purchase with home delivery has the highest probability of 14%. Additionally, purchasing products outside the catchment area has a probability of 2% and online orders in a physical store have a combined probability of 3.1%.

The probabilities change when the second store is opened in the catchment area. The second store is a smaller store located in the city center of Eindhoven. The probability that a consumer chooses to purchase in the store in Eindhoven is equal to 11%. The opening of the store in Eindhoven impacts the first store in Best the most, as the probability that a consumer chooses to purchase a product in Best decreases by 8%-points. However, the physical store in Best still has the highest probability with a probability of 63%. The combined probability of the online channel decreases by 3%-points with the opening of the store in Eindhoven. The probability of purchasing products outside the catchment area and the probability of ordering products online in a physical store decreases slightly.

Furthermore, the probabilities have also been calculated for the Covid-19 lockdown period in December 2020. During the Covid-19 lockdown, the only available shopping alternatives were to purchase the product online with home delivery or delivery to a pickup point. The online purchase with home delivery has the highest probability of 76% and the online purchase with pickup point delivery has a probability of 24%. Next to the predicted probabilities, Table 25 also shows the observed probabilities when there are two stores in the catchment area. The observed probabilities and the predicted probabilities have similar values indicating that the model is performing well at predicting observed values.

Table 25 – Predicted probabilities ML model Eindhoven and observed probabilities dataset Eindhoven

Catchment area Eindhoven	1 store in catchment area	2 stores in catchment area	COVID-19 lockdown	Observed probabilities with 2 stores in catchment area
Purchase in store 1	0.709	0.631	-	0.61
Purchase in store 2	-	0.106	-	0.11
Purchase in store outside catchment area	0.023	0.020	-	0.03
Online purchase + home delivery	0.139	0.124	0.76	0.13
Online purchase + pickup point delivery	0.044	0.039	0.24	0.05
Online purchase + click and collect delivery	0.06	0.05	-	0.05
Online order in store 1	0.029	0.026	-	0.03
Online order in store 2	-	0.004	-	0.01
Online order in store outside catchment area	0.0010	0.0009	-	0.001

The main difference between the predicted probabilities of the MNL model and the ML model is the probabilities for the purchases made in a physical store. Purchasing products in the physical store in Best (Purchase in store 1 in Table 24 and Table 25) has a probability of 69% in the MNL model and a probability of 63% in the ML model when two stores have been opened in the catchment area of Eindhoven. When looking at the predicted probabilities of every shopping alternative, the probabilities of the ML model are closer to the observed probabilities than the probabilities of the MNL model. This shows that the ML model is more accurate at predicting the observed probabilities. This seems logical as the ML contains a common random component for the physical stores (represented by the standard deviation of the 'ST'-parameter). Therefore, the ML model will be used for the model estimation of every catchment area.

Additionally, the turning point can be calculated when the estimated utility of shopping in a physical store overtakes the utility of purchasing products online. The online shopping alternative with the highest utility is online purchase with home delivery. This variable will be used for the calculation as this is the point where the consumer will have the same preference for shopping online (with home delivery) and going to a physical store. The estimated sum of utility of purchasing in a physical store includes three variables, namely the variables purchase in a physical store (ST), size of the store (SizeST), and distance to the store (DistST). Furthermore, the age (AST), and gender (GST) of every consumer influence the utility of the physical store.

The turning point for the store located in Best has a value between 29 and 37 kilometers (depending on age and gender). This means that when a consumer must travel more than 29-37 km to the store in Best, the consumer would on average prefer ordering online (with home delivery) rather than going to the store. For the store in Eindhoven, this would be the case if the consumer must travel more than 12 to 20 kilometers. The reason why consumers want to travel shorter distances to the Eindhoven store is that this store is considerably smaller than the one in Best. The calculations for the physical stores in Best and Eindhoven can be seen in Table 26 and Table 27 and the calculations for all the catchment areas can be found in Appendix C.

Table 26 – Turning point store Best

Turning point store Best	Age 18 - 34 Gender female	Age 35 - 54 Gender female	Age > 55 Gender female	Age 18 - 34 Gender male	Age 35 - 54 Gender male	Age > 55 Gender male
ST	2.356	2.356	2.356	2.356	2.356	2.356
SizeST (5193 m2)	2.482	2.482	2.482	2.482	2.482	2.482
AST	-0.341	0	0.341	-0.341	0	0.341
GST	-0.167	-0.167	-0.167	0.167	0.167	0.167
Total	4.331	4.672	5.012	4.664	5.005	5.345
OnlH	0.891	0.891	0.891	0.891	0.891	0.891
Total - OnlH	3.440	3.781	4.121	3.774	4.114	4.455
DistST	-0.119	-0.119	-0.119	-0.119	-0.119	-0.119
Turning point (km)	28.9	31.7	34.6	31.6	34.5	37.4

Table 27 – Turning point store Eindhoven

Turning point store Eindhoven	Age 18 - 34 Gender female	Age 35 - 54 Gender female	Age > 55 Gender female	Age 18 - 34 Gender male	Age 35 - 54 Gender male	Age > 55 Gender male
ST	2.356	2.356	2.356	2.356	2.356	2.356
SizeST (965 m2)	0.461	0.461	0.461	0.461	0.461	0.461
AST	-0.341	0	0.341	-0.341	0	0.341
GST	-0.167	-0.167	-0.167	0.167	0.167	0.167
Total	2.310	2.651	2.991	2.643	2.984	3.324
OnlH	0.891	0.891	0.891	0.891	0.891	0.891
Total - OnlH	1.419	1.760	2.100	1.753	2.093	2.434
DistST	-0.119	-0.119	-0.119	-0.119	-0.119	-0.119
Turning point (km)	11.9	14.8	17.6	14.7	17.6	20.4

4.2.3 CATCHMENT AMSTERDAM

The next model estimation, using the ML model, is performed in the catchment area of Amsterdam where the first store opened in 2000, the second store in 2019, and the third store in 2020. The results of the ML model are shown in Table 28. Every attribute that is included in the model has significant attribute levels which indicate that all attributes contribute to the utility of the shopping alternatives to a statistically significant extent. The ML model has a pseudo-rho-squared (ρ^2) value of 0.37 indicating that the model is performing well and is good at predicting observed values.

Table 28 – ML model Amsterdam

ID	Label	β	σ
ST	Purchase in a physical store	2.342 ***	1.749 ***
OnlH	Purchase online + Home delivery	1.381 ***	
OnlP	Purchase online + Pickup point delivery	0.0366 **	
Del	Online order in a physical store	-2.860 ***	
SizeST	Size of the store (m2)	3.806 e−4 ***	
DistST	Distance to store in catchment area (km)	-0.115 ***	
Dist3	Distance to store outside catchment area (dummy)	-3.754 ***	
AST	Age * purchase in a physical store	0.354 ***	
GST	Gender * purchase in a physical store	0.205 ***	
Significance codes: 0 "***" 0.001 "**" 0.01 "*" 0.05 "." 0.1 " " 1			
Log Likelihood model			-284030
Log Likelihood null model			-450745
pseudo rho-squared			0.37

The ML model can be used to predict the probability of choosing each shopping alternative based on the observed shopping data and the estimated parameters. The probability can be predicted for every transaction and the average predicted probability of the shopping alternatives can be calculated depending on the number of available shopping alternatives and the residential location of the consumers. There are four shopping alternative situations in the catchment area of Amsterdam, namely one store in the catchment area, two stores in the catchment area, three stores in the catchment area, and the Covid-19 lockdown period. The observed probabilities have been calculated for the shopping alternative situation when three stores have been opened in the catchment area. The average predicted probabilities of the shopping alternatives using the ML model for the three situations and the observed probabilities when three stores have been opened in the catchment area are shown in Table 29. The observed probabilities are somewhat different than the predicted probabilities when there are three stores in the catchment area. Additionally, the turning point can be calculated when the estimated utility of shopping in a physical store overtakes the utility of purchasing products online. The calculations for the three stores located in Amsterdam can be seen in Appendix C.

Table 29 – Predicted probabilities ML model Amsterdam and observed probabilities dataset Amsterdam

Catchment area Amsterdam	1 store in catchment area	2 stores in catchment area	3 stores in catchment area	COVID-19 lockdown	Observed probabilities with 3 stores in catchment area
Purchase in store 1	0.51	0.38	0.34	-	0.25
Purchase in store 2	-	0.24	0.22	-	0.10
Purchase in store 3	-	-	0.09	-	0.03
Purchase in store outside catchment area	0.05	0.04	0.03	-	0.06
Online purchase + home delivery	0.27	0.20	0.18	0.79	0.35
Online purchase + pickup point delivery	0.07	0.05	0.05	0.21	0.11
Online purchase + click and collect delivery	0.07	0.05	0.04	-	0.08
Online order in store 1	0.03	0.02	0.02	-	0.01
Online order in store 2	-	0.02	0.01	-	0.01
Online order in store 3	-	-	0.007	-	0.005
Online order in store outside catchment area	0.003	0.002	0.002	-	0.003

4.2.4 CATCHMENT UTRECHT

The next model estimation, using the ML model, is performed in the catchment area of Utrecht where the first store opened in 2018 and the second store in 2019. The results of the ML model are shown in Table 30. Every attribute that is included in the model has significant attribute levels which indicate that all attributes contribute to the utility of the shopping alternatives to a statistically significant extent. The ML model has a pseudo-rho-squared (ρ^2) value of 0.2 indicating that the model is performing well and is good at predicting observed values.

Table 30 – ML model Utrecht

ID	Label	β	σ
ST	Purchase in a physical store	2.392 ***	1.541 ***
OnlH	Purchase online + Home delivery	1.364 ***	
OnlP	Purchase online + Pickup point delivery	0.263 ***	
Del	Online order in a physical store	-1.910 ***	
SizeST	Size of the store (m2)	1.598 e−4 ***	
DistST	Distance to store in catchment area (km)	-0.118 ***	
Dist3	Distance to store outside catchment area (dummy)	-1.856 ***	
AST	Age * purchase in a physical store	0.113 ***	
GST	Gender * purchase in a physical store	0.200 ***	
Significance codes: 0 "***" 0.001 "**" 0.01 "*" 0.05 "." 0.1 " " 1			
Log Likelihood model			-153430
Log Likelihood null model			-192993
pseudo rho-squared			0.2

The ML model can be used to predict the probability of choosing each shopping alternative based on the data that has been imported and analyzed. The probability can be predicted for every transaction and the average predicted probability of the shopping alternatives can be calculated depending on the number of available shopping alternatives and the residential location of the consumers. There are four shopping alternative situations in the catchment area of Utrecht, namely no store in the catchment area, one store in the catchment area, two stores in the catchment area, and the Covid-19 lockdown period. The average predicted probabilities of the four shopping alternative situations are shown in Table 31. Next to the predicted probabilities, Table 31 also shows the observed probabilities when there are two stores in the catchment area. The observed probabilities and the predicted probabilities have rather similar values indicating that the model is performing well at predicting observed values. Additionally, the turning point can be calculated when the estimated utility of shopping in a physical store overtakes the utility of purchasing products online. The turning point for the store Utrecht The Wall has a value between 12 and 17 kilometers. For the store in Utrecht Vredenburg, this would be the case if the consumer must travel more than 10 to 16 kilometers. The calculations for the two stores located in Utrecht can be seen in Appendix C.

Table 31 – Predicted probabilities ML model Utrecht and observed probabilities dataset Utrecht

Catchment area Utrecht	No store in catchment area	1 store in catchment area	2 stores in catchment area	COVID-19 lockdown	Observed probabilities with 2 stores in catchment area
Purchase in store 1	-	0.31	0.24	-	0.28
Purchase in store 2	-	-	0.20	-	0.20
Purchase in store outside catchment area	0.29	0.19	0.14	-	0.05
Online purchase + home delivery	0.42	0.27	0.21	0.75	0.24
Online purchase + pickup point delivery	0.14	0.09	0.07	0.25	0.08
Online purchase + click and collect delivery	0.11	0.07	0.05	-	0.10
Online order in store 1	-	0.05	0.04	-	0.02
Online order in store 2	-	-	0.03	-	0.02
Online order in store outside catchment area	0.04	0.03	0.02	-	0.01

4.2.5 CATCHMENT ROTTERDAM

The next model estimation, using the ML model, is performed in the catchment area of Rotterdam where the first store opened in 2016 and the second store in 2018. The results of the ML model are shown in Table 32. Every attribute that is included in the model has significant attribute levels which indicate that all attributes contribute to the utility of the shopping alternatives to a statistically significant extent. The ML model has a pseudo-rho-squared (ρ^2) value of 0.32 indicating that the model is performing well and is good at predicting observed values.

Table 32 – ML model Rotterdam

ID	Label	β	σ
ST	Purchase in a physical store	2.491 ***	1.559 ***
OnlH	Purchase online + Home delivery	1.218 ***	
OnlP	Purchase online + Pickup point delivery	-0.0682 ***	
Del	Online order in a physical store	-2.714 ***	
SizeST	Size of the store (m2)	2.418 e−4 ***	
DistST	Distance to store in catchment area (km)	-0.174 ***	
Dist3	Distance to store outside catchment area (dummy)	-3.021 ***	
AST	Age * purchase in a physical store	0.255 ***	
GST	Gender * purchase in a physical store	0.184 ***	
Significance codes: 0 "***" 0.001 "**" 0.01 "*" 0.05 "." 0.1 " " 1			
Log Likelihood model			-265810
Log Likelihood null model			-391707
pseudo rho-squared			0.32

The ML model can be used to predict the probability of choosing each shopping alternative based on the data that has been imported and analyzed. The probability can be predicted for every transaction and the average predicted probability of the shopping alternatives can be calculated depending on the number of available shopping alternatives and the residential location of the consumers. There are four shopping alternative situations in the catchment area of Rotterdam, namely no store in the catchment area, one store in the catchment area, two stores in the catchment area, and the Covid-19 lockdown period. The average predicted probabilities of the four shopping alternative situations are shown in Table 33. Next to the predicted probabilities, Table 33 also shows the observed probabilities when there are two stores in the catchment area. The observed probabilities and the predicted probabilities have similar values indicating that the model is performing well at predicting observed values. Additionally, the turning point can be calculated when the estimated utility of shopping in a physical store overtakes the utility of purchasing products online. The turning point for the store Rotterdam Coolsingel has a value between 11 and 16 kilometers. For the store in Rotterdam Alexandrium, this would be the case if the consumer must travel more than 7 to 12 kilometers. The calculations for the two stores located in Rotterdam can be seen in Appendix C.

Table 33 – Predicted probabilities ML model Rotterdam and observed probabilities dataset Rotterdam

Catchment area Rotterdam	No store in catchment area	1 store in catchment area	2 stores in catchment area	COVID-19 lockdown	Observed probabilities with 2 stores in catchment area
Purchase in store 1	-	0.47	0.39	-	0.32
Purchase in store 2	-	-	0.15	-	0.15
Purchase in store outside catchment area	0.17	0.09	0.07	-	0.07
Online purchase + home delivery	0.52	0.26	0.22	0.78	0.25
Online purchase + pickup point delivery	0.14	0.07	0.06	0.22	0.09
Online purchase + click and collect delivery	0.15	0.08	0.06	-	0.08
Online order in store 1	-	0.03	0.03	-	0.02
Online order in store 2	-	-	0.01	-	0.01
Online order in store outside catchment area	0.01	0.01	0.005	-	0.004

4.2.6 CATCHMENT KERKRADE

The next model estimation, using the ML model, is performed in the catchment area of Kerkrade where the store opened in 2003. The results of the ML model are shown in Table 34. Every attribute that is included in the model has significant attribute levels which indicate that all attributes contribute to the utility of the shopping alternatives to a statistically significant extent. The variable for store size (SizeST) could not be estimated using the model of Kerkrade, so this variable is excluded in the model estimation. The ML model has a pseudo-rho-squared (ρ^2) value of 0.69 indicating that the model is performing well and is good at predicting observed values.

Table 34 – ML model Kerkrade

ID	Label	β	σ
ST	Purchase in a physical store	9.382 ***	1.991 ***
OnlH	Purchase online + Home delivery	5.335 ***	
OnlP	Purchase online + Pickup point delivery	3.846 ***	
Del	Online order in a physical store	-3.444 ***	
SizeST	Size of the store (m2)	n/a	
DistST	Distance to store in catchment area (km)	-0.113 ***	
Dist3	Distance to store outside catchment area (dummy)	-5.166 ***	
AST	Age * purchase in a physical store	0.924	
GST	Gender * purchase in a physical store	0.162	
Significance codes: 0 "***" 0.001 "**" 0.01 "*" 0.05 "." 0.1 " " 1			
Log Likelihood model			-122020
Log Likelihood null model			-398559
pseudo rho-squared			0.69

The ML model can be used to predict the probability of choosing each shopping alternative based on the data that has been imported and analyzed. The probability can be predicted for every transaction and the average predicted probability of the shopping alternatives can be calculated depending on the number of available shopping alternatives and the residential location of the consumers. There are two shopping alternative situations in the catchment area of Kerkrade, namely one store in the catchment area and the Covid-19 lockdown period. The average predicted probabilities of the two shopping alternative situations are shown in Table 35. Next to the predicted probabilities, Table 35 also shows the observed probabilities when there is one store in the catchment area. The observed probabilities and the predicted probabilities have similar values indicating that the model is performing well at predicting observed values. Additionally, the turning point can be calculated when the estimated utility of shopping in a physical store overtakes the utility of purchasing products online. The turning point for the store located in Kerkrade has a value between 26 and 45 kilometers. The calculations for the store located in Kerkrade can be seen in Appendix C.

Table 35 – Predicted probabilities ML model Kerkrade and observed probabilities dataset Kerkrade

Catchment area Kerkrade	1 store in catchment area	COVID-19 lockdown	Observed probabilities with 1 store in catchment area
Purchase in store 1	0.79	-	0.82
Purchase in store outside catchment area	0.02	-	0.02
Online purchase + home delivery	0.13	0.82	0.11
Online purchase + pickup point delivery	0.03	0.18	0.02
Online purchase + click and collect delivery	0.001	-	0.001
Online order in store 1	0.03	-	0.03
Online order in store outside catchment area	0.001	-	0.001

4.2.7 CATCHMENT APELDOORN

The next model estimation, using the ML model, is performed in the catchment area of Apeldoorn where the store opened in 2014. The results of the ML model are shown in Table 36. Every attribute that is included in the model has significant attribute levels which indicate that all attributes contribute to the utility of the shopping alternatives to a statistically significant extent. The ML model has a pseudo-rho-squared (ρ^2) value of 0.61 indicating that the model is performing very well and is good at predicting observed values.

Table 36 – ML model Apeldoorn

ID	Label	β	σ
ST	Purchase in a physical store	2.854 ***	1.799 ***
OnlH	Purchase online + Home delivery	0.994 ***	
OnlP	Purchase online + Pickup point delivery	-0.731 ***	
Del	Online order in a physical store	-3.093 ***	
SizeST	Size of the store (m2)	5.268e−4 ***	
DistST	Distance to store in catchment area (km)	-0.101 ***	
Dist3	Distance to store outside catchment area (dummy)	-4.296 ***	
AST	Age * purchase in a physical store	0.242 ***	
GST	Gender * purchase in a physical store	0.118 ***	
Significance codes: 0 "***" 0.001 "**" 0.01 "*" 0.05 "." 0.1 " " 1			
Log Likelihood model			-140400
Log Likelihood null model			-357581
pseudo rho-squared			0.61

The ML model can be used to predict the probability of choosing each shopping alternative based on the data that has been imported and analyzed. The probability can be predicted for every transaction and the average predicted probability of the shopping alternatives can be calculated depending on the number of available shopping alternatives and the residential location of the consumers. There are two shopping alternative situations in the catchment area of Apeldoorn, namely one store in the catchment area and the Covid-19 lockdown period. The average predicted probabilities of the two shopping alternative situations are shown in Table 37. Next to the predicted probabilities, Table 37 also shows the observed probabilities when there is one store in the catchment area. The observed probabilities and the predicted probabilities have similar values indicating that the model is performing well at predicting observed values. Additionally, the turning point can be calculated when the estimated utility of shopping in a physical store overtakes the utility of purchasing products online. The turning point for the store located in Apeldoorn has a value between 33 and 40 kilometers. The calculations for the store located in Apeldoorn can be seen in Appendix C.

Table 37 – Predicted probabilities ML model Apeldoorn and observed probabilities dataset Apeldoorn

Catchment area Apeldoorn	1 store in catchment area	COVID-19 lockdown	Observed probabilities with 1 store in catchment area
Purchase in store 1	0.76	-	0.78
Purchase in store outside catchment area	0.03	-	0.03
Online purchase + home delivery	0.12	0.85	0.10
Online purchase + pickup point delivery	0.02	0.15	0.02
Online purchase + click and collect delivery	0.04	-	0.04
Online order in store 1	0.03	-	0.04
Online order in store outside catchment area	0.001	-	0.001

4.2.8 ALL CATCHMENT AREAS

The data has been analyzed for each of the six different catchment areas. It is also possible to analyze the data using samples from every catchment in one large dataset. This dataset will be analyzed using the ML model and the results will be presented in this section. First, the decision was made to exclude the sample population from Kerkrade, as the model for Kerkrade in section 4.2.6 could not be estimated with all variables and the estimated parameters have rather extreme values. Secondly, two variables are added to the dataset to give more insight into the shopping behavior of consumers. The first variable is parking which includes the options of paid parking, free parking, and no parking (in case of online purchase). The parking variable has a value of zero when there are paid parking facilities around the physical store and the parking variable has a value of one when there are free parking facilities around the physical store, so the expected parameter of parking will be positive. The second variable is added to see if there are differences in utility between city center stores and stores located in retail parks or on the outskirts of a city. The second variable has a value of one when a purchase has been made in a physical store in a city center. These variables could not be analyzed in the individual catchment area analysis since there was too little variation among the shopping alternatives. Within the model using all catchment areas, the correlation between the parking and city center variable will decrease, as there is more variation between parking facilities and the location of the physical store.

The dataset of the sample consists of 5,000 customer cards from each catchment area. In total, 25,000 customer cards will be analyzed that have made 182,251 transactions. The number of analyzed customer cards and transactions from the sample can be seen in Table 38. Table 38 also shows the total analyzed customer cards and transactions that were collected in the five catchment areas. The sample consists of 17.5% of the total number of collected customer cards and 17.3% of the total number of collected transactions.

Table 38 – Analyzed customer cards and transactions of sample

Catchment Area	Analyzed customer cards	Analyzed transactions
Amsterdam	5,000	31,310
Apeldoorn	5,000	44,396
Eindhoven	5,000	41,926
Rotterdam	5,000	33,227
Utrecht	5,000	31,392
Total sample	25,000	182,251
Total all 5 catchment areas	142,661	1,051,411
Sample (%)	17.5%	17.3%

The results of the ML model are shown in Table 39. Every attribute that is included in the model has significant attribute levels which indicate that all attributes contribute to the utility of the shopping alternatives to a statistically significant extent. The intercept parameter online of purchase with Click and Collect delivery is set at zero for identification. The ML model representing all catchment areas has a pseudo-rho-squared (ρ^2) value of 0.41 indicating that the model is performing very well and is good at predicting observed values. The standard deviation (σ) of the physical store (ST) variable measures whether the physical stores compete more with each other than with the other (online) alternatives. The higher the standard deviation, the more the physical stores compete with each other.

Table 39 – ML model including all catchment areas

ID	Label	β	σ
ST	Purchase in a physical store	3.158 ***	1.795 ***
OnlH	Purchase online + Home delivery	1.199 ***	
OnlP	Purchase online + Pickup point delivery	-0.0987 ***	
Del	Online order in a physical store	-2.773 ***	
ParkST12	Parking facilities	0.0949 ***	
SizeST	Size of the store (m2)	2.229 e−4 ***	
DistST	Distance to store in catchment area (km)	-0.104***	
Dist3	Distance to store outside catchment area (dummy)	-3.283 ***	
AST	Age * purchase in a physical store	0.308 ***	
GST	Gender * purchase in a physical store	0.147 ***	
CST	City center location * purchase in a physical store	-0.810 ***	
Significance codes: 0 "***" 0.001 "**" 0.01 "*" 0.05 "." 0.1 " " 1			
Log Likelihood model			-216540
Log Likelihood null model			-366897
pseudo rho-squared			0.41

The results for the different purchase and delivery possibilities can be compared with each other. It can be said that the 'purchase in a physical store' (ST) variable has a positive utility equal to 3.158. The 'online order in a physical store' (Del) variable has a negative utility, meaning that the standard utility of that store decreases by 2.773. This means that ordering in a physical store is considered less attractive than taking home the products. The 'online purchase with home delivery' (OnlH) variable has a positive utility equal to 1.199. The variable 'online ordering and pickup point delivery' (OnlP) has a negative utility which means that the consumers in all catchment areas consider this an unattractive alternative. It also means that the 'online purchase with Click and Collect delivery' (OnlCC) variable has a higher utility as this parameter is kept at zero. Thus, it can be said that purchasing products online with home delivery has the highest utility compared to the other delivery methods when purchasing products online.

Furthermore, the 'size of the physical store' (SizeST) variable has a positive utility which means that the larger the store, the higher the overall utility of that store. The effect of store size is very small, however, the average size of a Decathlon store is almost 3000 m^2. The standard utility of a physical store (equal to 3.158) increases by 0.223 for every 1000 m^2, so this indicates that the size of the physical store has a considerable impact on the overall utility of a store. The 'distance to the store' (DistST) variable has a negative utility which means that the further the consumer must travel to the store, the lower the overall utility of that store. For every 10km traveling, the utility of a store decreases by 1.04. As the distance to stores outside the catchment area is unknown, this distance effect is measured by the 'distance to a physical store outside the catchment area' (Dist3) variable. Given the estimates for both distance variables (DistST and Dist3), it can be concluded that the average effect of traveling to a store outside the catchment area is perceived as traveling 31.6 kilometers (-3.283/-0.104) to a physical store.

Besides the size of the physical store and the distance to the physical store affecting the overall utility of purchasing in a store, the 'parking facilities' (ParkST12) variable and 'city center location' (CST) variable were added to this model. The parking facilities have a positive utility which means that there is a positive effect of parking on overall store utility. The utility of the physical store is zero if there are paid parking facilities close to the store and the utility of the physical store increases by 0.0949 if there are free parking facilities close to the store. The utility is equal to zero if there is an online purchase as no parking is required to purchase products online. The city center location variable has a negative utility which indicates that if a store is in a city center, the overall utility of that store will decrease by 0.81. This variable shows that consumers have a lower utility for physical stores that are located in the city center.

Additionally, two context variables are added to the ML model to measure the effects of age and gender on visiting a physical store. The context age variable (AST) has a positive utility which indicates that the age group >55 years old has a higher overall utility for shopping in a physical store compared to the age group 35-54 years old. It also indicates that the age group 18-35 years old has a lower overall utility for shopping in a physical store. The context gender variable (GST) also has a positive utility indicating that men tend to have a higher overall utility for shopping in a physical store compared to women.

The ML model can be used to predict the probability of choosing each shopping alternative based on data used to estimate the model (see Table 38). The probability can be predicted for every transaction and the average predicted probability of the shopping alternatives can be calculated depending on the number of available shopping alternatives and the residential location of the consumers. There are five shopping alternative situations in the model with all catchments, namely no store in the catchment area, one store in the catchment area, two stores in the catchment area, three stores in the catchment area, and the Covid-19 lockdown period. The average predicted probabilities of the five shopping alternative situations are shown in Table 40. Next to the predicted probabilities, Table 40 also shows the observed probabilities when there are two stores in the catchment area. The observed probabilities and the predicted probabilities have similar values indicating that the model is performing well at predicting observed values.

Table 40 – Predicted probabilities ML model all catchment areas and observed probabilities dataset all catchment

Model using all catchment areas	No stores in catchment area	1 store in catchment area	2 stores in catchment area	3 stores in catchment area	COVID-19 lockdown	Observed probabilities with 2 stores in catchment area
Purchase in store 1	-	0.58	0.49	0.44	-	0.45
Purchase in store 2	-	-	0.14	0.12	-	0.14
Purchase in store 3	-	-	-	0.09	-	-
Purchase in store outside catchment area	0.19	0.07	0.06	0.06	-	0.04
Online purchase + home delivery	0.51	0.20	0.17	0.15	0.79	0.19
Online purchase + pickup point delivery	0.14	0.05	0.05	0.04	0.21	0.06
Online purchase + click and collect delivery	0.15	0.06	0.05	0.05	-	0.07
Online order in store 1	-	0.04	0.03	0.03	-	0.02
Online order in store 2	-	-	0.01	0.01	-	0.01
Online order in store 3	-	-	-	0.01	-	-
Online order in store outside catchment area	0.01	0.005	0.004	0.004	-	0.004

Purchasing products online with home delivery has the highest probability when there is no store located in the catchment area with a probability of 0.51. This indicates that, on average, consumers have a 51% chance of choosing online shopping with home delivery for their purchases when there is no store located in the catchment area. The online shopping channels have a combined probability of 80%. Purchasing products in a store outside the catchment area has a probability of 19%.

Purchasing products in store 1 has the highest probability (58%) when there is one store located in the catchment area. The online channel transactions have a combined probability of 31%. Within this channel, online purchase with home delivery has the highest probability of 20%. Additionally, purchasing products outside the catchment area has a probability of 7% and online orders in a physical store have a combined probability of 4.5%.

The probabilities change when the second store is opened in the catchment area. The probability that a consumer chooses to purchase in store 2 is equal to 14%. The opening of store 2 impacts store 1 the most, as the probability that a consumer chooses to purchase a product in store 1 decreases by 9%-points. However, physical store 1 still has the highest probability with a probability of 49%. The combined probability of the online channel decreases by 4%-points with the opening of store 2. The probability of purchasing products outside the catchment area and the probability of ordering products online in a physical store decreases slightly. The probabilities also change when the third store is opened in the catchment area. The probability that a consumer chooses to purchase in store 3 is equal to 9%. Physical store 1 has the highest probability with 44%. The combined probability of the online channel decreases by 3%-points with the opening of store 3. The probability of purchasing products outside the catchment area and the probability of ordering products online in a physical store decreases slightly.

Furthermore, the probabilities have also been calculated for the Covid-19 lockdown period in December 2020. During the Covid-19 lockdown, the only available shopping alternatives were to purchase the product online with home delivery or delivery to a pickup point. The online purchase with home delivery has the highest probability of 79% and the online purchase with pickup point delivery has a probability of 21%.

4.2.9 DISCUSSION AND CONCLUSION

The ML models estimated in the previous sections have given insights into the shopping alternative preferences of consumers who shop at Decathlon. However, the estimated models of the different catchment areas show similarities but also some different results which will be discussed in this section. First, comparing the estimated models for the different catchment areas shows that all the utilities are statistically significant. The results of the ML model of the different catchment areas are shown in Table 41. The lowest model fit is in the catchment area of Utrecht which has a ρ^2 value of 0.2, so this means that all the models perform well at predicting the observed values. The rho-squared value is different among the different catchment areas which can be explained by the differences in opening years of physical stores in the catchment areas. The first physical store that was opened in the catchment area of Rotterdam and Utrecht was in 2016 and 2018, respectively. This could indicate that online shopping and shopping in a physical store outside the catchment area were not perfectly taken into account within the ML model and this leads to a lower rho-squared value.

Table 41 – ML model results of the different catchment areas

ID	Label	Eindhoven	Amsterdam	Utrecht	Rotterdam	Kerkrade	Apeldoorn	All catchment areas (excluding Kerkrade)
ST	Purchase in a physical store	2.356	2.342	2.392	2.491	9.382	2.854	3.158
ST (σ)	Standard deviation purchase in a physical store	1.86	1.749	1.541	1.559	1.991	1.799	1.795
OnlH	Purchase online + Home delivery	0.891	1.381	1.364	1.218	5.335	0.994	1.199
OnlP	Purchase online + Pickup point delivery	-0.267	0.0366	0.263	-0.0682	3.846	-0.731	-0.0987
Del	Online order in a physical store	-3.19	-2.86	-1.91	-2.714	-3.444	-3.093	-2.773
ParkST12	Parking facilities	-	-	-	-	-	-	0.0949
SizeST	Size of the store (m2)	4.780 e−4	3.806 e−4	1.598 e−4	2.418 e−4	-	5.268 e−4	2.229 e−4
DistST	Distance to store in catchment area (km)	-0.119	-0.115	-0.118	-0.174	-0.113	-0.101	-0.104
Dist3	Distance to store outside catchment area (dummy)	-4.21	-3.754	-1.856	-3.021	-5.166	-4.296	-3.283
AST	Age * purchase in a physical store	0.341	0.354	0.113	0.255	0.924	0.242	0.308
GST	Gender * purchase in a physical store	0.167	0.205	0.2	0.184	0.162	0.118	0.147
CST	City center location * purchase in a physical store	-	-	-	-	-	-	-0.81
Log-likelihood model		-362910	-284030	-153430	-265810	-122020	-140400	-216540
Log-likelihood null model		-750175	-450745	-192993	-391707	-398559	-357581	-366897
Rho-squared		0.52	0.37	0.2	0.32	0.69	0.61	0.41

The utilities for the different purchase and delivery possibilities can be seen in Table 41. The utilities for the variables are quite similar across the catchment areas except for the catchment area in Kerkrade. The model for Kerkrade could not be estimated with all the variables and the estimated parameters have extreme values. Figure 25 shows the utilities for the purchase and delivery possibilities excluding the catchment area of Kerkrade. The 'purchase in a physical store' (ST) variable has the highest value within every estimated model. This seems to be consistent with the findings in the descriptive analysis shown in section 4.1 where purchasing products in a store has the highest number of transactions. The 'purchase in a physical store' (ST) variable has a higher utility value in the model using all catchment areas than in the other models. This can be explained by the 'city center location' (CST) variable which has a negative utility indicating that if a store is in a city center, the overall utility of that store will decrease by 0.81. The 'purchase in a physical store' (ST) variable will have a utility of 2.35 if the physical store is located in a city center and the variable will have a utility of 3.16 if the physical store is not located in a city center. The utility values for the 'purchase in a physical store' (ST) variable of the other catchment areas are in between the two utility values of the model using all catchment areas. The 'online order in a physical store' (Del) variable has been added to evaluate the utility for consumers to order products online from a physical store. This variable has a strong negative utility within the estimated model of every catchment area which indicates that ordering in a physical store is considered less attractive than taking home the products. This is also consistent with the findings in section 4.1 where this alternative has the lowest share in the total number of transactions.

When looking at the online channel variables, the 'online purchase with home delivery' (OnlH) has the highest utility within every estimated model. This is consistent across the different catchment areas indicating that home delivery is the most preferred delivery option with the highest utility. The 'online purchase with Click and Collect delivery' (OnlCC) variable has a utility of zero. The variable 'online ordering and pickup point delivery' (OnlP) has both positive and negative utility values indicating that delivery with pickup point delivery can have a lower or higher utility value than Click and Collect delivery depending on the catchment area. For the catchment areas of Eindhoven, Rotterdam, Apeldoorn, and the model using all catchment areas, Click and Collect delivery has a higher overall utility than pickup point delivery. On the other hand, for the catchment areas of Amsterdam, Kerkrade, and Utrecht, Click and Collect delivery has a lower overall utility than pickup point delivery. This means that there is no clear preference between the Click and Collect delivery and pickup point delivery method. The two delivery methods are also quite similar, as they both consist of the delivery of a package to a specific location where it can be collected. Using Click and Collect delivery, the location where the package can be delivered is a physical Decathlon retail store and the pickup point location can be a range of retail stores. This indicates that for the catchment areas of Eindhoven, Rotterdam, and Apeldoorn, the retail stores might be located at more attractive locations which are better accessible as they have a higher overall utility than the pickup point locations.

Figure 25 – Utility values for purchase and delivery possibilities (excluding catchment area Kerkrade)

Next to the variables for the purchase and delivery possibilities, a variable was added to measure the effect of the size of the store. The variable for the size of the physical store (SizeST) influences the utility of purchasing a product in a physical store. Additionally, a parking variable is added to the model with all catchment areas to measure the effect of parking on the overall utility of purchasing a product in a store. The size and parking utilities can also be seen in Table 41. Looking at the size variable, the utilities have a very low positive value, but this is quite consistent among the different catchment areas. The average size of a Decathlon store is around 3000 m^2, which means that the utility of the size of the store can have a considerable effect on the utility of a store. The effect of the size of the physical store on the utility of a physical store can be seen in figure 26 where five different store sizes have been used to calculate the utility for the store size. For example, the utility for the store size for the store in the catchment area of Rotterdam will become 0.73 if the store has a size of 3000 m^2. Thus, the size of a store has a considerable impact on the overall utility of a store. This is also in line with the findings of Briesch et al. (2009) that consumers will choose larger stores that offer more product categories and more variety of products within a product category. Furthermore, the parking facilities variable (ParkST12) has a positive utility which means that there is a positive effect of parking on overall store utility. The utility of the physical store is zero if there are paid parking facilities close to the store and the utility of the physical store increases by 0.0949 if there are free parking facilities close to the store. By definition, the utility of the online purchase alternatives is not affected by parking facilities. Free parking facilities have a higher impact on the overall utility of a store than paid parking facilities which seems obvious.

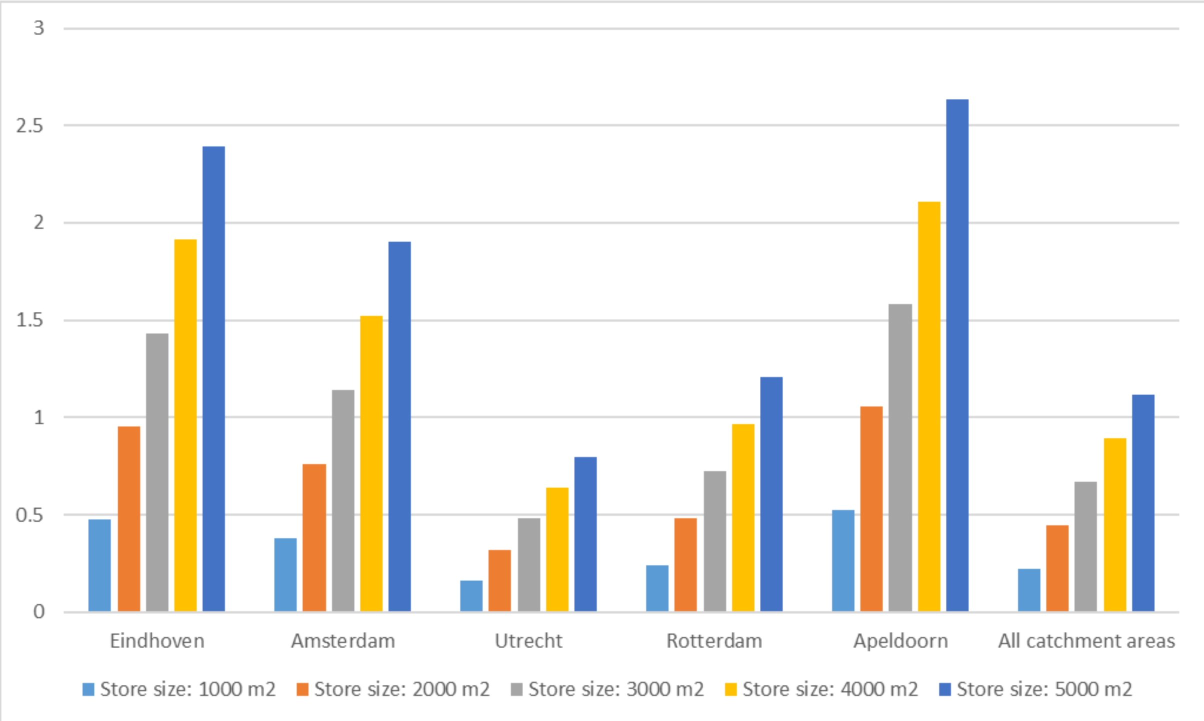

Figure 26 – Utility values for the size of the physical store (SizeST) considering five different store sizes

Additionally, the distance to a physical store is analyzed within the different model estimations to evaluate store preferences based on distance. The utilities for the 'distance to a physical store within the catchment area' (DistST) variable can also be seen in Table 41. The effect of the distance to the physical store on the utility of a physical store can be seen in figure 27 where six different distances have been used to calculate the utility for the distance to the physical store within the catchment area. The different catchment areas have similar parameters for the distance to a physical store within the catchment area. The distance parameters range between -0.101 for the catchment area of Amsterdam and -0.174 for the catchment area of Rotterdam. The catchment area of Rotterdam has a stronger negative utility value than the other catchment areas which could indicate that the physical stores in Rotterdam are less attractive for consumers living in the catchment area of Rotterdam. Consumers living in this catchment area appear to be more distance sensitive. Overall, the higher the travel distance to a physical store, the lower the preference for shopping at a physical store. This is in line with the efficiency hypothesis presented by Farag et al. (2006). The efficiency hypothesis states that consumers who live further away tend to purchase more online as they have longer travel times. Consumers tend to be efficient in the way they shop and longer travel distances lead to less efficiency meaning that consumers will have more preferences for shopping online. This also seems the case for this research, as the parameter for distance creates lower preferences to shop in a physical store when the distance to a store increases.

Next to the variable for measuring the effect of distance to a physical store within the catchment area, another variable has been added that measures the effect of distance to a physical store outside the catchment area (Dist3) which is a dummy variable that has a value of one when the purchase is made in a store outside of the catchment area. The distance to a physical store parameter of the different models can also be seen in Table 41. Given the estimates for both distance variables (DistST and Dist3), the perceived distance to the physical store outside the catchment area can be identified. There are quite a few differences between the different catchment areas. The lowest distances are found in the catchment areas of Utrecht and Rotterdam, which can be explained by the fact that the first store was opened in 2016 and 2018 respectively. This means that the consumers had to go to stores outside of the catchment area to make a purchase which creates a lower threshold to shop outside the catchment area. The highest distances are found in the catchment areas of Apeldoorn and Kerkrade, which can be explained by the fact that these stores are quite isolated from other stores outside the catchment area which makes the threshold higher to shop in a physical store outside the catchment area. The catchment area of Kerkrade and Apeldoorn also consists of one physical store meaning that the average distance between a physical store and the consumers living in the catchment area is higher.

Figure 27 – Utility values for the distance to the physical store (DistST) considering six different distances

Furthermore, sociodemographic variables were added to the different models to evaluate store preference per age group and gender. The utility values related to the sociodemographic variables of the different models can be seen in both Table 41 and figure 28. The different catchment areas have similar parameters for both the age (AST) variable and the gender (GST) variable. Looking at the age (AST) variable, the parameters range between 0.113 in the catchment area of Utrecht and 0.354 in the catchment area of Amsterdam. This indicates that the age group >55 years tend to have a higher preference for purchasing products in a physical store and that the age group 18 - 35 years old tend to have a lower preference for purchasing products in a physical store than the in-between age group. This also means that the age group of 18 - 35 years old will have higher preferences for online shopping and the age group >55 years old have lower preferences for online shopping. This is in line with the findings of Farag et al. (2007) and Irawan (2015), as they state that the higher the age of the consumer, the lower the probability of online shopping. Looking at the gender variable (GST), the parameters range between 0.118 in the catchment area of Apeldoorn and 0.205 in the catchment area of Amsterdam. This indicates that men tend to have higher preferences for shopping in a physical store and women have lower preferences for shopping in a physical store. This also means that women are more inclined to shop online. This is in line with the findings of Maat and Konings (2018) who stated that women are more likely to shop online.

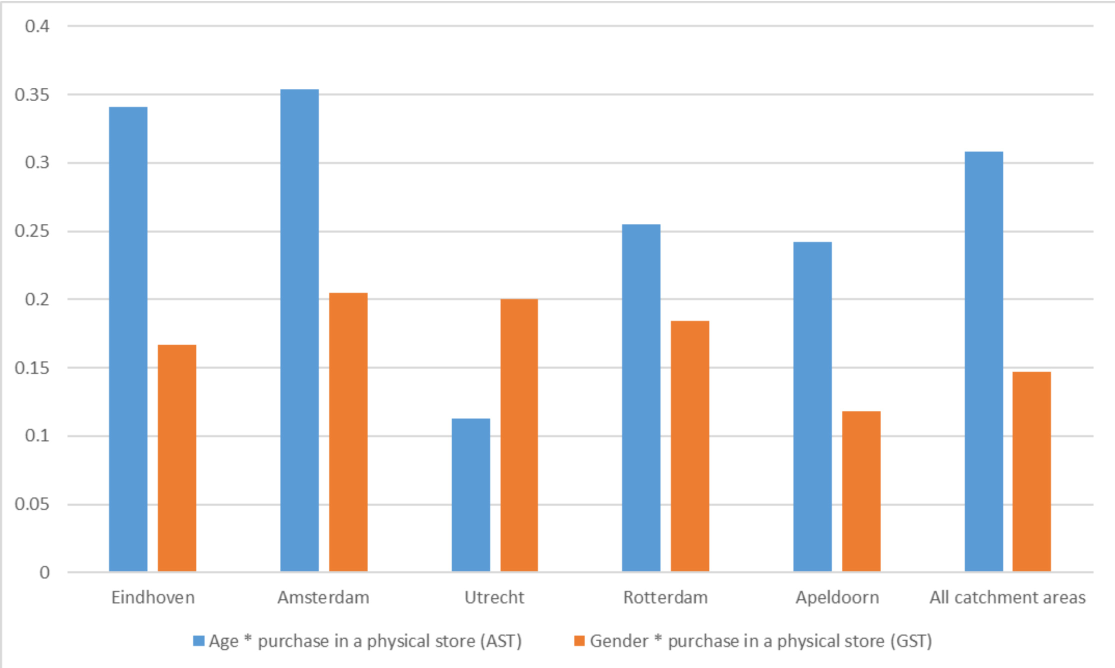

Figure 28 – Utility values for the age and gender variable (excluding catchment area Kerkrade)

The estimated ML models were used to predict the average probability of the shopping alternatives for the different catchment areas depending on the number of available shopping alternatives and the residential location of the consumers. Table 42 shows the average probabilities when there is no store located in the catchment areas. The value 'n/a' indicates that this option is not applicable for the specific catchment area. The catchment area of Utrecht and Rotterdam have similar probabilities. Purchasing products online with home delivery has the highest probability in every catchment area indicating that online shopping is preferred when there is no physical store available. Purchasing products in a store outside the catchment area has a higher probability in the catchment area of Utrecht, probably because the stores outside the catchment area of Utrecht are pretty well accessible.

Table 42 – Average predicted probabilities with no store located in the catchment area

No store in catchment area	Catchment Eindhoven	Catchment Amsterdam	Catchment Utrecht	Catchment Rotterdam	Catchment Kerkrade	Catchment Apeldoorn	All catchments
Purchase in store outside catchment area	n/a	n/a	0.29	0.17	n/a	n/a	0.19
Online purchase + home delivery	n/a	n/a	0.42	0.52	n/a	n/a	0.51
Online purchase + pickup point delivery	n/a	n/a	0.14	0.14	n/a	n/a	0.14
Online purchase + click and collect delivery	n/a	n/a	0.11	0.15	n/a	n/a	0.15
Online order in store outside catchment area	n/a	n/a	0.04	0.01	n/a	n/a	0.01

The average predicted probabilities change when there is one store located in the catchment area and this can be seen in Table 43. This situation applies to every catchment area. There are quite some similarities among the different catchment areas. The shopping alternative with the highest probability is named physical store 1. This indicates that the physical store plays an important role in the catchment areas and the omnichannel shopping environment. The probability of shopping in a physical store outside a catchment area decreases due to the availability of a store with a lower travel distance for the consumers living in the catchment area. Online shopping also experiences a decrease in probability, as the physical store in the catchment area is more attractive. This also shows that the opening of a physical store in a catchment area has a substitution effect on the other shopping alternatives which is in line with the findings of Joewono et al. (2020).

Table 43 – Average predicted probabilities with one store located in the catchment area

One store in catchment area	Catchment Eindhoven	Catchment Amsterdam	Catchment Utrecht	Catchment Rotterdam	Catchment Kerkrade	Catchment Apeldoorn	All catchments
Purchase in store 1	0.71	0.51	0.31	0.47	0.79	0.76	0.58
Purchase in store outside catchment area	0.02	0.05	0.19	0.09	0.02	0.03	0.07
Online purchase + home delivery	0.14	0.27	0.27	0.26	0.13	0.12	0.20
Online purchase + pickup point delivery	0.04	0.07	0.09	0.07	0.03	0.02	0.05
Online purchase + click and collect delivery	0.06	0.07	0.07	0.08	0.001	0.04	0.06
Online order in store 1	0.03	0.03	0.05	0.03	0.03	0.03	0.04
Online order in store outside catchment area	0.001	0.003	0.03	0.01	0.001	0.001	0.005

Adding a second store to the catchment area also means that the probability of choosing each alternative changes and this can be seen in Table 44. Adding a second store to the catchment area has the most impact on the probability of shopping in the first store. Consumers have an additional option to purchase products in a physical store in the catchment area which means that they must decide between two stores instead of one. This also indicates that the opening of a second physical store in a catchment area has a substitution effect on the purchases in the other physical stores. Additionally, the probability of online shopping and shopping in a physical store outside the catchment area decreases slightly making the choice of these alternatives less attractive.

Table 44 – Average predicted probabilities with two stores located in the catchment area

Two stores in catchment area	Catchment Eindhoven	Catchment Amsterdam	Catchment Utrecht	Catchment Rotterdam	Catchment Kerkrade	Catchment Apeldoorn	All catchments
Purchase in store 1	0.63	0.38	0.24	0.39	n/a	n/a	0.49
Purchase in store 2	0.11	0.24	0.20	0.15	n/a	n/a	0.14
Purchase in store outside catchment area	0.02	0.04	0.14	0.07	n/a	n/a	0.06
Online purchase + home delivery	0.12	0.20	0.21	0.22	n/a	n/a	0.17
Online purchase + pickup point delivery	0.04	0.05	0.07	0.06	n/a	n/a	0.05
Online purchase + click and collect delivery	0.05	0.05	0.05	0.06	n/a	n/a	0.05
Online order in store 1	0.03	0.02	0.04	0.03	n/a	n/a	0.03
Online order in store 2	0.004	0.02	0.03	0.01	n/a	n/a	0.01
Online order in store outside catchment area	0.001	0.002	0.02	0.005	n/a	n/a	0.004

The average predicted probabilities change when there are three stores located in the catchment area and this can be seen in Table 45. This alternative is only applicable to the catchment area of Amsterdam. The combined probability of shopping in a physical store increases, while the individual probabilities of shopping in physical store 1 and store 2 decreases. This indicates that cannibalization of transactions from one physical store to the new physical store is possible and it means that the opening of a third physical store in a catchment area has a substitution effect on the purchases in the other physical stores. The probabilities of shopping in a physical store outside a catchment and online shopping show a slight decrease, as the physical stores in the catchment area are more attractive.

Table 45 – Average predicted probabilities with three stores located in the catchment area

Three stores in catchment area	Catchment Eindhoven	Catchment Amsterdam	Catchment Utrecht	Catchment Rotterdam	Catchment Kerkrade	Catchment Apeldoorn	All catchments
Purchase in store 1	n/a	0.34	n/a	n/a	n/a	n/a	0.44
Purchase in store 2	n/a	0.22	n/a	n/a	n/a	n/a	0.12
Purchase in store 3	n/a	0.09	n/a	n/a	n/a	n/a	0.09
Purchase in store outside catchment area	n/a	0.03	n/a	n/a	n/a	n/a	0.06
Online purchase + home delivery	n/a	0.18	n/a	n/a	n/a	n/a	0.15
Online purchase + pickup point delivery	n/a	0.05	n/a	n/a	n/a	n/a	0.04
Online purchase + click and collect delivery	n/a	0.04	n/a	n/a	n/a	n/a	0.05
Online order in store 1	n/a	0.02	n/a	n/a	n/a	n/a	0.03
Online order in store 2	n/a	0.01	n/a	n/a	n/a	n/a	0.01
Online order in store 3	n/a	0.01	n/a	n/a	n/a	n/a	0.01
Online order in store outside catchment area	n/a	0.002	n/a	n/a	n/a	n/a	0.004

Additionally, the average predicted probabilities have been calculated for the relatively short Covid-19 lockdown period in December 2020. During this lockdown, it was not possible to shop in a physical store which means that most of the shopping alternatives could not be used. Only two shopping alternatives were available, namely online shopping with home delivery and online shopping with pickup point delivery. Table 46 shows the probabilities in the different catchment areas during the lockdown. Online shopping with the home delivery alternative has the highest probability.

Table 46 – Average predicted probabilities during Covid-19 lockdown period

Covid-19 lockdown period	Catchment Eindhoven	Catchment Amsterdam	Catchment Utrecht	Catchment Rotterdam	Catchment Kerkrade	Catchment Apeldoorn	All catchments
Online purchase + home delivery	0.76	0.79	0.75	0.78	0.82	0.85	0.79
Online purchase + pickup point delivery	0.24	0.21	0.25	0.22	0.18	0.15	0.21

Furthermore, the turning point has been calculated when the estimated utility of shopping in a physical store overtakes the utility of purchasing products online. The online purchase alternative with the highest utility is the online purchase with home delivery in every catchment area. This utility has been used to calculate the turning point where the consumer will have the same utility for shopping online (with home delivery) and going to a physical store. The estimated sum of utility of purchasing in a physical store includes three variables, namely the variables purchase in a physical store (ST), size of the store (SizeST), and distance to the store (DistST). Furthermore, the age (AST), and gender (GST) of every consumer influence the utility of the physical store. The turning points for every physical store in the different catchment areas can be seen in Table 47. The turning point can be seen as the range of the physical store where consumers living inside the turning point range of a physical store will have a higher probability of choosing that physical store. The physical stores in Best, Kerkrade, and Apeldoorn have the highest turning points indicating that they reach a lot of consumers in their catchment area. Most of the stores have a range between 10 and 20 kilometers. Having multiple physical stores in the catchment area affects the range of the physical stores, as there are more physical stores to choose from. Consumers are more likely to choose a physical store that has the shortest distance from their home address, while also keeping the size of the physical store in mind.

Table 47 – Turning points of every store in the different catchment areas

Turning point stores	Age 18 - 34 Gender female	Age 35 - 54 Gender female	Age > 55 Gender female	Age 18 - 34 Gender male	Age 35 - 54 Gender male	Age > 55 Gender male
Catchment area Eindhoven						
Store Best	28.9	31.7	34.6	31.6	34.5	37.4
Store Eindhoven	11.9	14.8	17.6	14.7	17.6	20.4
Catchment area Amsterdam						
Store Amsterdam Arena	13.2	16.2	19.3	16.7	19.8	22.9
Store Amsterdam Noord	13.7	16.8	19.9	17.3	20.4	23.4
Store Amsterdam Kinkerstraat	9.1	12.0	15.0	12.0	14.9	17.9
Catchment area Utrecht						
Store Utrecht The Wall	11.7	12.7	13.7	15.1	16.1	17.0
Store Utrecht Vredenburg	10.4	11.4	12.3	13.8	14.7	15.7
Catchment area Rotterdam						
Store Rotterdam Coolsingel	10.9	12.3	13.8	13.0	14.5	15.9
Store Rotterdam Alexandrium	7.3	8.8	10.3	9.5	10.9	12.4
Catchment area Kerkrade						
Store Kerkrade	26.2	34.4	42.6	29.1	37.2	45.4
Catchment area Apeldoorn						
Store Apeldoorn	32.9	35.3	37.7	35.2	37.6	40.0

The results of the model estimations show that the mixed logit model can be used to gain insight into individual decision-making when an additional physical store is opened. The next section describes a case study of an additional store opening in the catchment area of Eindhoven.

4.3 CASE STUDY: PREDICTING CONSUMER BEHAVIOR

The model estimation that has been conducted in the previous section resulted in estimated utility and probability values that can be used to predict the changing omnichannel shopping behavior of consumers in future store openings. For this research, a case study is conducted in the catchment area of Eindhoven to predict the changing shopping behavior when another physical store is added. The physical store that is added to the catchment area is a store located in the city center of Helmond to expand the number of consumers. The location of the physical store in Helmond can be seen in figure 29.

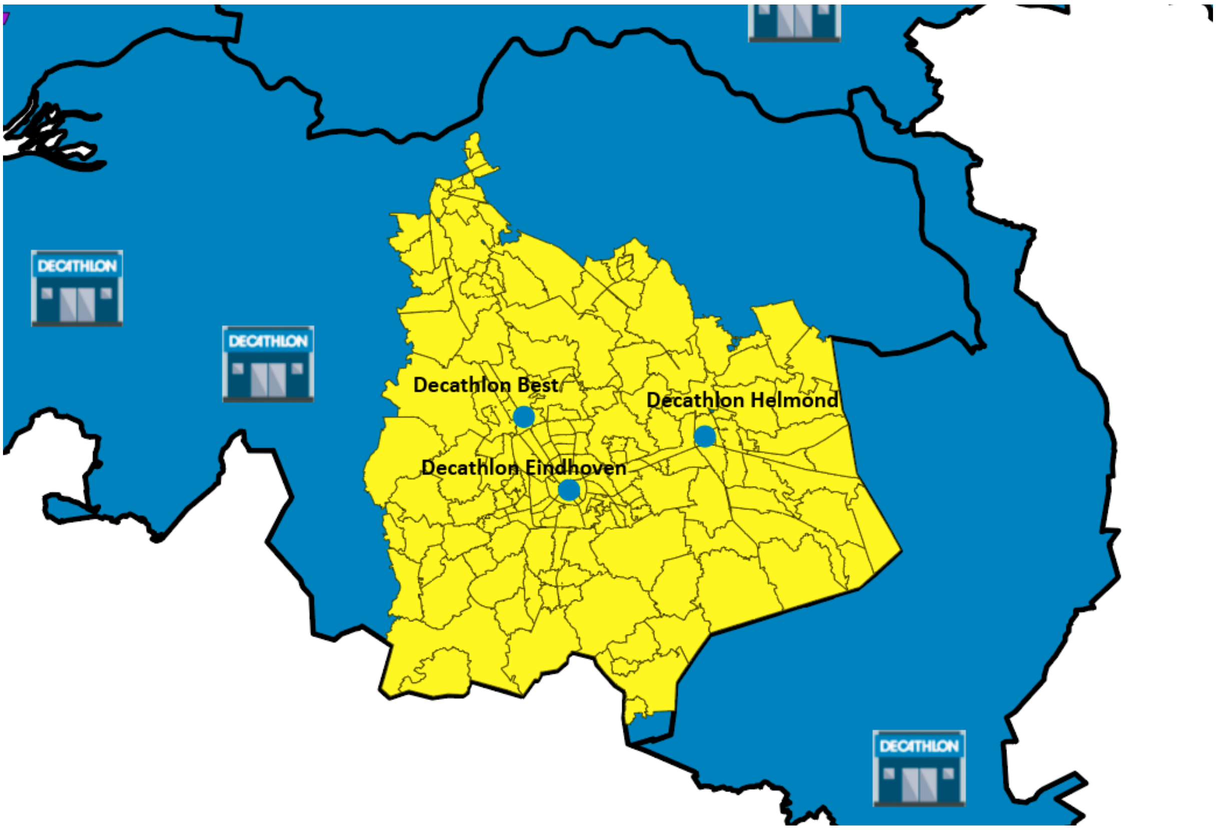

Figure 29 – Location of the physical store in Helmond in the catchment area of Eindhoven

The physical store located in the city center of Helmond will have a size between 1000 m^2 and 3000 m^2 and there are paid parking facilities nearby the physical store. The new physical store will target consumers to visit the store who live too far away from the stores located in Best and Eindhoven and it is expected to cannibalize physical store transactions from the stores in Best and Eindhoven. The cannibalization will happen as the overall utility of the store in Helmond will become higher for consumers living closer to the new store. The goal is to minimize the cannibalization from the stores located in Best and Eindhoven, but maximize the potential for the city of Helmond and attract consumers to the physical store that would have otherwise chosen to make an online transaction or a transaction in a physical store outside the catchment area.

The case study will be analyzed using two different models, namely the estimated ML model that was used for the catchment area of Eindhoven in section 4.2.2 and the estimated ML model for all catchments in section 4.2.8. This will be done to test if the models are robust at predicting changes in consumer behavior. The results from the ML model will be used to predict the changes in consumer behavior when an additional store is added in the catchment area in Helmond. The additional store will result in two more shopping alternatives, namely purchasing a product in the physical store in Helmond and ordering a product online in the physical store in Helmond. With the addition of these two alternatives, the total number of shopping alternatives in the catchment area will increase to eleven shopping alternatives.

To facilitate the prediction of opening a third physical store in the Eindhoven catchment area, a so-called synthetic population is created. The decision was made to create a synthetic population that has six individuals in every 4-digit postal code in the Eindhoven catchment area. In total, there are 164 4-digit postal codes which means that the synthetic population consists of 984 consumers. As the estimated model contains effects on utilities of both the gender (2 groups) and age (3 groups), the six consumers per postal code area represent the six different combinations of age and gender. For each consumer in the synthetic population, the probabilities of choosing each alternative are predicted under two scenarios. In scenario 1, no new store will be opened. This scenario serves to check how well the results predicted for the synthetic population (the third column of probabilities in Table 48) match the predicted probabilities for all consumers in the estimation dataset (the first column of probabilities in Table 48) and the observed probabilities in the dataset of Eindhoven (the second column of probabilities in Table 48). In scenario 2, the opening of a third physical store is simulated. Note that the predicted probabilities for each postal code area are weighted by the number of inhabitants per postal code area to determine the overall probabilities for the entire Eindhoven catchment area.

First, the model estimated for the catchment area of Eindhoven is used to compare the probabilities predicted for each observation in the catchment area of Eindhoven with the probabilities predicted under scenario 1 (Table 48). It is important to investigate whether these sets of probabilities are similar. According to Table 48, the probability of shopping at physical store 1 is about 5%-points lower under scenario 1. This could mean that more transactions have been made by consumers that live closer to physical store 1 in the estimation dataset. On the other hand, the online channel probabilities are somewhat higher under scenario 1. Secondly, the dataset of the catchment area of Eindhoven is used to compare the observed probabilities in the catchment area of Eindhoven with the probabilities predicted under scenario 1 (Table 48). According to Table 48, the probability of shopping at physical store 1 is about 3%-points lower under scenario 1. The observed probability of physical store 1 is slightly lower than the predicted probability for physical store 1 from the estimated model. On the other hand, the online channel probabilities show similar results when comparing them with the predicted probabilities from the model estimation of Eindhoven. Overall both sets of probabilities match rather well.

Table 48 – Scenario 1 compared to probabilities Eindhoven model

Comparing scenario 1 with predicted and observed probabilities	**Predicted probabilities Eindhoven model**	**Observed probabilities Eindhoven dataset**	**Scenario 1**
Purchase in store 1	0.631	0.609	0.579
Purchase in store 2	0.106	0.105	0.115
Purchase in store outside catchment area	0.020	0.027	0.030
Online purchase + home delivery	0.124	0.126	0.143
Online purchase + pickup point delivery	0.039	0.046	0.045
Online purchase + click and collect delivery	0.050	0.051	0.058
Online order in store 1	0.026	0.029	0.024
Online order in store 2	0.004	0.007	0.005
Online order in store outside catchment area	0.001	0.001	0.001

To predict the effects of opening a third store in the catchment area of Eindhoven, five versions of scenario 2 were created, corresponding with five different sizes of the new store in Helmond, ranging from 1000 to 3000 m^2. The results can be seen in Table 49. The probability of the physical store in Best decreases by 8.3% when the physical store in Helmond will have a size of 1000 m^2 (scenario 2a). The larger the physical store in Helmond, the higher the decrease in the probability of the physical store located in Best. The physical store in Eindhoven is also negatively affected by the opening of the physical store in Helmond. The probability of the physical store in Eindhoven decreases by 13% when the physical store in Helmond has a size of 1000 m^2. The probability of the physical store in Eindhoven shows a slightly decreasing linear relationship with the increase in store size of the physical store in Helmond. The physical store located in Helmond will have a probability of 8.6% when the store has a size of 1000 m^2. When the size of the physical store in Helmond becomes larger, the probability that consumers will choose the physical store in Helmond increases. The size of the physical store in Eindhoven is equal to 965 m^2. When the physical store in Helmond has a size of 1000 m^2, the physical store in Eindhoven has a higher probability than the physical store in Helmond while the physical store in Eindhoven is smaller. This can be caused by the demographic characteristics of the catchment area of Eindhoven, as more consumers live around the physical store in the city center of Eindhoven than around the physical store in the city center of Helmond. When consumers live closer to a physical store, the probability is higher that these consumers will shop in that physical store. The new physical store in Helmond competes more with the existing physical store in Eindhoven than the physical store in Best. This makes sense as the distance between Helmond and Eindhoven is shorter than between Helmond and Best.

Additionally, the probabilities of the online channel shopping alternatives decrease when the physical store opens in Helmond. When the store in Helmond has a size of 1000 m^2, the combined probability of the online channel decreases by 8.1% indicating that consumers living close to the physical store in Helmond will substitute shopping online with shopping in a physical store. The same counts for the purchase in physical stores outside the catchment area. When the store in Helmond has a size of 1000 m^2, the probability of shopping in a physical store outside the catchment area decreases by 13%. The probability of shopping in a physical store outside the catchment area and the combined probability of the online shopping alternatives show a growing decreasing trend when the size of the physical store increases. This could indicate that the physical store in Helmond will attract consumers that live relatively far away from the physical store in Eindhoven and Best. These consumers used to have a high probability for online shopping and shopping outside the catchment area. However, the opening of the physical store in Helmond causes these consumers to make a trade-off between online shopping, shopping in a physical store outside the catchment area, or shopping at the physical store in Helmond.

Table 49 – Average predicted probabilities of synthetic population per store size (Eindhoven model

Average predicted probabilities per store size	Scenario 1	Scenario 2a: 1000 m2	Scenario 2b: 1500 m2	Scenario 2c: 2000 m2	Scenario 2d: 2500 m2	Scenario 2e: 3000 m2
Purchase in store Best	0.579	0.531	0.521	0.509	0.496	0.481
Purchase in store Eindhoven	0.115	0.100	0.098	0.095	0.092	0.089
Purchase in store Helmond	-	0.086	0.104	0.123	0.145	0.170
Purchase in store outside catchment area	0.030	0.026	0.025	0.024	0.023	0.022
Online purchase + home delivery	0.143	0.131	0.129	0.126	0.123	0.120
Online purchase + pickup point delivery	0.045	0.041	0.040	0.040	0.039	0.038
Online purchase + click and collect delivery	0.058	0.054	0.053	0.052	0.051	0.049
Online order in store Best	0.024	0.022	0.021	0.021	0.020	0.020
Online order in store Eindhoven	0.005	0.004	0.004	0.004	0.004	0.004
Online order in store Helmond	-	0.004	0.004	0.005	0.006	0.007
Online order in store outside catchment area	0.001	0.001	0.001	0.001	0.001	0.001

Secondly, the same probabilities are predicted with the model estimated for all catchment areas (except Kerkrade). The main difference with the Eindhoven model is that this model has two additional variables. The first variable is parking which is an additional variable to measure the utility of a physical store by looking if the parking facilities are paid parking or free parking. This variable has a positive effect on the utility indicating that the utility of a physical store is higher when a physical store has free parking than when a physical store has paid parking. The second variable is a variable that measures if there is a preference difference between physical stores located in the city center and physical stores located on the edge of a city. This variable has a negative effect indicating that the utility of a physical store is lower when a physical store is in the city center.

To predict the effects of opening a third store in the catchment area of Eindhoven, five versions of scenario 2 were created, corresponding with five different sizes of the new store in Helmond, ranging from 1000 to 3000 m^2. The results of the predicted probabilities with the model for all the catchment areas under the different scenarios can be seen in Table 50 and the results are similar to those shown in Table 48. The probability of the physical store in Best decreases by 8.8% when the physical store in Helmond will have a size of 1000 m^2 (scenario 2a). The larger the physical store in Helmond, the higher the decrease in the probability of the physical store located in Best. The physical store in Eindhoven is also negatively affected by the opening of the physical store in Helmond. The probability of the physical store in Eindhoven decreases by 10% when the physical store in Helmond has a size of 1000 m^2. The probability of the physical store in Eindhoven shows a slightly decreasing linear relationship with the increase in store size of the physical store in Helmond. The physical store located in Helmond will have a probability of 8.1% when the store has a size of 1000 m^2. When the size of the physical store in Helmond will become larger, the probability that consumers will choose the physical store in Helmond has a growing increasing trend. The size of the physical store in Eindhoven is equal to 965 m^2. When the physical store in Helmond has a size of 2000 m^2, the physical store in Eindhoven and the physical store in Helmond have an equal probability, while the physical store in Eindhoven is smaller. This can be caused by the demographic characteristics of the catchment area of Eindhoven, as more consumers live around the physical store in the city center of Eindhoven than around the physical store in the city center of Helmond. When consumers live closer to a physical store, the probability is higher that these consumers will shop in that physical store. The physical store in Best has a growing decreasing trend and the physical store in Helmond has a growing increasing trend when the physical store in Helmond increases in size. Again, the new physical store in Helmond competes somewhat more with the physical store in Eindhoven than with the physical store in Best.

Additionally, the probability of the online channel shopping alternatives decreases when the physical store opens in Helmond. When the store in Helmond has a size of 1000 m^2, the combined probability of the online channel decreases by 7.2% indicating that consumers living close to the physical store in Helmond will substitute shopping online with shopping in a physical store. The same counts for the purchase in physical stores outside the catchment area. When the store in Helmond has a size of 1000 m^2, the probability of shopping in a physical store outside the catchment area decreases by 16%. The probability of shopping in a physical store outside the catchment area and the combined probability of the online shopping alternatives show a growing decreasing trend when the size of the physical store increases. This could indicate that the physical store in Helmond will attract consumers that live too far away from the physical store in Eindhoven and Best. These consumers used to have a high probability for online shopping and shopping outside the catchment area. However, the opening of the physical store in Helmond causes these consumers to make a trade-off between online shopping, shopping in a physical store outside the catchment area, or shopping at the physical store in Helmond.

Table 50 – Average predicted probabilities of synthetic population per store size (Overall model)

Average predicted probabilities per store size	**Scenario 1**	**Scenario 2a: 1000 m2**	**Scenario 2b: 1500 m2**	**Scenario 2c: 2000 m2**	**Scenario 2d: 2500 m2**	**Scenario 2e: 3000 m2**
Purchase in store Best	0.514	0.469	0.465	0.461	0.456	0.451
Purchase in store Eindhoven	0.110	0.099	0.098	0.097	0.096	0.094
Purchase in store Helmond	-	0.081	0.089	0.097	0.106	0.115
Purchase in store outside catchment area	0.058	0.050	0.050	0.049	0.048	0.048
Online purchase + home delivery	0.175	0.163	0.162	0.160	0.159	0.158
Online purchase + pickup point delivery	0.048	0.044	0.044	0.044	0.043	0.043
Online purchase + click and collect delivery	0.053	0.049	0.049	0.048	0.048	0.047
Online order in store Best	0.032	0.029	0.029	0.029	0.028	0.028
Online order in store Eindhoven	0.007	0.006	0.006	0.006	0.006	0.006
Online order in store Helmond	-	0.005	0.006	0.006	0.007	0.007
Online order in store outside catchment area	0.004	0.003	0.003	0.003	0.003	0.003

The results of the predicted probabilities show that there are quite some similarities between the two models. The model that is estimated using the mixed logit model that included all catchment areas can be seen as a more general model that can be used to predict changing consumer behavior in every catchment area in the Netherlands. This model has two additional variables that are used to increase the accuracy of the predictions. The model that is estimated using the mixed logit model of the catchment area of Eindhoven is a region-specific model that has a higher accuracy to predict the probabilities for the catchment area of Eindhoven. The general model estimated for all catchment areas can also be used as an additional model to test for the robustness of the analysis. In this case, the general model shows that the results from the region-specific model are robust and only small differences exist between the models. The main outcome is that the larger the size of the physical store in Helmond, the larger the competition with the existing stores in the region. Thus, there is a trade-off between the cannibalization of the physical stores in Best and Eindhoven and the number of customers that will be attracted to the physical store in Helmond. Both models perform well at predicting the observed probabilities for the different shopping alternatives. Therefore, both models can be used for predicting shopping behavior when the retail environment changes in a catchment area. It depends on the objective of the prediction which model is most suitable, as both models have their strengths and weakness.

4.4 CONCLUSION

This chapter reported the results of the quantitative study which involves revealed shopping data collected with customer cards in six different regions within the Netherlands. The demographic and transaction analysis were described first to provide a good understanding of the consumers and the transactions they made. The data that has been collected was then analyzed using mixed logit (ML) models and a case study has been presented using the results obtained by the estimated ML model in the catchment area of Eindhoven.

The descriptive analysis has shown that there are some trends among the analyzed customer cards and transactions in the six catchment areas. The most important shopping channel is the physical channel where most of the transactions have taken place and this is similar for every catchment area. On average, the transactions in the physical store have an increasing trend until 2018 and show a decreasing trend in 2019 and 2020. The decreasing trend in 2019 might be due to the increase in online channel transactions and the decreasing trend in 2020 can partly be explained by the Covid-19 pandemic. On the other hand, the transactions in the online channel show an increasing trend since the start of the analysis. The online channel transactions even overtake the physical channel transactions in 2020 in some of the catchment areas. This indicates that the online channel is becoming more important and that this trend should be followed closely in the future. In total, the number of transactions is increasing due to store openings and the increasing popularity of Decathlon in the Netherlands.

The model estimation has shown that the models perform well at predicting observed choices and that there are quite some similarities among the different catchment areas. Every attribute that is included in the model estimation is significant which indicates that all attributes contribute to the utility of the shopping alternatives to a statistically significant extent. The different parameters have the correct signs, indicating that the models are working well and can be used to identify omnichannel shopping behavior in a catchment area.

The model estimations have given some interesting insights into the shopping behavior of consumers. Considering the utility of the different shopping channels, the results of the model estimations comply with the results found in the descriptive analysis in section 4.1. When looking at the shopping channel variables, the physical channel is the shopping channel with the highest overall utility. When looking at the online channel variables, the utility for the online purchase option with home delivery has the highest value. This is consistent across the different catchment areas indicating that home delivery is the most preferred online delivery option. On the other hand, there is no clear preference between the Click and Collect delivery and pickup point delivery method among the different catchment areas. This seems reasonable, as the two delivery methods are quite similar in the way that they both consist of the delivery of an online purchase to a specific location where it can be collected.

Furthermore, the travel distance to a store and the size of the store seems to have the highest impact on the utility of a store. The travel distance to a store has a negative utility value indicating that the further the consumer must travel to a physical store, the lower the utility of that store. The higher the travel distance to the physical store, the more likely the consumer is to adopt online shopping. This is in line with the efficiency hypothesis presented by Farag et al. (2006). The size of the store has a positive utility value indicating that the larger the physical store, the higher the overall utility of that store. This is in line with the findings of Briesch et al. (2009) that consumers will choose larger stores that offer more product categories and more variety of products within a product category. The size of the physical store and the maximum distance that a consumer is willing to travel has a positive correlation. When the size of the store is larger, the distance to the physical store that consumers are willing to travel to becomes higher. On the other hand, when the size of the store is smaller, the distance to the physical store that consumers are willing to travel to becomes shorter. This is in line with the findings of Reilly (1931) who discussed the law of retail gravitation. The law of retail gravitation states that the probability of choosing a retail store is positively related to its size but inversely related to its distance from the shopper's home. The results of this research indicate that the law of retail gravitation still applies nowadays.

Additionally, the effect of age and gender on purchasing products in physical stores have similar results in the different catchment areas. The age and gender variables have positive utility parameters in every catchment area. For the age of the consumer, the age group >55 years tend to have a higher preference for purchasing products in a physical store, and the age group 18 - 35 years old tend to have a lower preference for purchasing products in a physical store than the in-between age group. This is in line with the findings of Farag et al. (2007) and Irawan (2015). For the gender of the consumer, men tend to have higher preferences for shopping in a physical store, and women have lower preferences for shopping in a physical store. This is in line with the findings of Maat and Konings (2018).

The ML model can be used to predict the probability of choosing each shopping alternative using the estimated parameters. The average predicted probability of the shopping alternatives has been calculated for the different catchment areas depending on the number of available shopping alternatives and the residential location of the consumer. Purchasing products online with home delivery has the highest probability in every catchment area indicating that online shopping is preferred when no store is available. Purchasing products in a store outside the catchment area has a high probability, as this is the only option to purchase products in a physical store. The average predicted probabilities change when the number of shopping alternatives changes. The opening of the first physical store in a catchment area has a substitution effect on the other shopping alternatives which is in line with the findings of Joewono et al. (2020). Additionally, when another physical store is added to a catchment area, the combined probability of shopping in a physical store increases, while the individual probabilities of shopping in the existing physical stores decrease indicating a substitution effect of shopping visits between physical stores.

The estimated models can be used to predict the changing omnichannel shopping behavior of consumers when the retail environment changes. A case study has been conducted in the catchment area of Eindhoven to predict the changing shopping behavior when a physical store is added. The case study included a synthetic population of 984 consumers where every 4-digit postal code included six consumers with a different age and gender combination. Using two different models – one specifically estimated for the Eindhoven catchment area and one more general model estimated for all catchment areas – the probabilities of consumers choosing alternative shopping channels were predicted for different scenarios.

The results of these predictions show that there are quite some similarities between the two models. The model that is estimated using the data of all catchment areas can be seen as a more general model that can be used to predict changing consumer behavior in every catchment area in the Netherlands. The model that is estimated using the data of the catchment area of Eindhoven is a region-specific model that has a higher accuracy to predict the probabilities for the catchment area of Eindhoven. The general model shows that the results from the region-specific model are robust and only small differences exist between the models. The probability of choosing to shop at the physical store in Eindhoven or Best decreases when the physical store in Helmond increases in size. This indicates that there is competition between the physical stores. The probability of shopping in a physical store outside the catchment area and the combined probability of the online shopping alternatives show a decreasing trend when the size of the physical store in Helmond increases. This could indicate that the physical store in Helmond will attract consumers that live too far away from the physical store in Eindhoven and Best. The models show that the larger the size of the physical store in Helmond, the larger the competition with the existing stores in the region. Thus, there is a trade-off between the cannibalization of the physical stores in Best and Eindhoven and the number of customers that will be attracted to the physical store in Helmond. Overall, both models can be used for predicting shopping behavior when the retail environment changes in a catchment area.

5. CONCLUSION

Although extensive research has been conducted on omnichannel shopping behavior, almost no studies exist that combine omnichannel shopping behavior with opening new physical channels. Therefore, this research was designed to gain insight into individual decision-making when an additional physical store is opened and into the effects on omnichannel shopping behavior in a region. This chapter will combine the findings from both the literature study and the data analysis to answer the research questions and formulate general conclusions in section 5.1. Section 5.2 describes the managerial recommendations and the limitations and recommendations for future research are discussed in section 5.3.

5.1 GENERAL CONCLUSIONS

The retail industry is changing as the popularity of online shopping is booming and the physical retail stores are under pressure reinforced by the Covid-19 pandemic. Online shopping has increased in popularity because of the possibilities to compare price and quality and the information that is available and easily accessible (Wagner et al., 2013). Online shopping is very convenient and changes the decision-making process of the consumer. Moving between different online and offline channels is becoming the norm for consumers and they expect a seamless experience through all channels (Melero et. al, 2016; Xu & Jackson, 2019; Blázquez, 2014). This research tried to understand the impact of physical stores and online websites to target and position these channels better and to be able to serve current consumers better.

Consumers have more power than ever to search for the best and unique experiences with their products and services which makes it a complex market. Omnipresent retailers need to understand their consumers and they must keep adapting to new channels with the increasing digitization in the retail industry. Online shopping is increasing rapidly, while the number of physical stores is decreasing in the Netherlands. Opening a physical store has an impact on both online and offline channel performance and it is important to gain insight into consumer behavior in different shopping channels. This research tried to predict how consumers use different shopping channels and how they integrate an additional physical store in their shopping behavior. This can be used by retailers to predict the changes in consumer behavior when adding physical stores. The main research question of this research was formulated as follows:

- What is the impact of opening an additional physical store on omnichannel shopping behavior in a region?

Transactional data was used to analyze changing consumer behavior trends over several years. Customer cards of consumers living in six catchment areas in the Netherlands were analyzed using their transactions in different shopping channels. The main shopping channels are the physical channel, the online channel with different delivery options, and online orders in a physical channel. One important notice is that it is not obligatory to use a customer card in a physical store which results that only 20% of the consumers scan their customer card when shopping in a physical store. The physical shopping channel alternatives were described by the characteristics of parking, store size, and distance to the store. Additionally, the age and gender of the consumers were used to investigate if there are demographic differences in shopping channel behavior. In total, 166,212 customer cards including 1,256,696 transactions were analyzed using a mixed logit (ML) model. The ML model is a random utility model and an extension of the MNL model. The model was specified to measure substitution between the physical shopping alternatives. This was useful in this research, as the ML model can take uneven competition between online and offline shopping alternatives in a region into account.

The descriptive analysis has shown that there are some trends among the analyzed customer cards and transactions in the six catchment areas. The most important shopping channel is the physical channel where most of the transactions have taken place. On average, the transactions in the physical store have an increasing trend until 2018 and show a decreasing trend in 2019 and 2020. On the other hand, the transactions in the online channel show an increasing trend since the start of the analysis. The online channel transactions even overtake the physical channel transactions in 2020 in some of the catchment areas. This indicates that the online channel is becoming more important and that this trend should be followed closely in the future. In total, the number of transactions is increasing due to physical store openings and the increase in online shopping.

The model estimations have given some interesting insights into the shopping behavior of consumers. Considering the overall utility of the different shopping channels, the results of the model estimation follow the same results found in the descriptive analysis in section 4.1. When looking at the shopping channel variables, the physical channel is the shopping channel with the highest utility. When looking at the online channel variables, the utility of the online purchase option with home delivery has the highest value. This is consistent across the different catchment areas indicating that an online purchase with home delivery is the most preferred delivery option. The travel distance to a store has a negative utility value indicating that the further the consumer must travel to a physical store, the lower the overall utility of that store. The higher the travel distance to the physical store, the more likely the consumer is to adopt online shopping. This is in line with the efficiency hypothesis presented by Farag et al. (2006). The size of the store has a positive effect indicating that the larger the physical store, the higher the overall utility of that store. This is in line with the findings of Briesch et al. (2009). Additionally, the combination of the travel distance to a store and the size of the store is in line with the law of retail gravitation that was formulated by Reilly (1931). The probability of choosing a retail store is positively related to the size of the store but inversely related to its distance from the shopper’s home. Furthermore, the age and the gender of the individual consumers influence choice probabilities. Consumers in the age group >55 years tend to have a higher preference for purchasing products in a physical store, and that the age group 18 - 35 years old tend to have a lower preference for purchasing products in a physical store than the in-between age group which is in line with the findings of Farag et al. (2007) and Irawan (2015). For the gender of the consumer, men tend to have higher preferences for shopping in a physical store, and women have lower preferences for shopping in a physical store which is in line with the findings of Maat and Konings (2018).

The ML model can be used to predict the probability of choosing each shopping alternative. Purchasing products online with home delivery has the highest probability if there is no physical store in the catchment area. The average predicted probabilities change when the number of shopping alternatives changes. The opening of the first physical store in a catchment area has a substitution effect on the other shopping alternatives which is in line with the findings of Joewono et al. (2020). When an additional physical store is added to a catchment area, the combined probability of shopping in a physical store increases, while the individual probabilities of shopping in the existing physical stores decrease indicating a substitution effect of shopping visits between physical stores.

The estimated models can be used to predict the changing omnichannel shopping behavior of consumers when the retail environment changes. A case study has been conducted in the catchment area of Eindhoven to predict the changing shopping behavior when a physical store is added. The case study included a synthetic population of 984 consumers where every 4-digit postal code included six consumers with a different age and gender combination. Using two different models – one specifically estimated for the Eindhoven catchment area and one more general model estimated for all catchment areas – the probabilities of consumers choosing alternative shopping channels were predicted for different scenarios.

The results of these predictions show that there are quite some similarities between the two models. The model estimated for all catchment areas can be seen as a more general model that can be used to predict changing consumer behavior in every catchment area in the Netherlands. The model estimated for the catchment area of Eindhoven is a region-specific model that has a higher accuracy to predict the probabilities for the catchment area of Eindhoven. The general model shows that the results from the region-specific model are robust and only small differences exist between the models. The models show that the larger the size of the physical store in Helmond, the larger the competition with the existing stores in the region. Thus, there is a trade-off between the cannibalization of the physical stores in Best and Eindhoven and the consumers that will switch from online shopping to offline shopping in the physical store in Helmond. Overall, both models can be used for predicting shopping behavior when the retail environment changes in a catchment area.

The present study confirms previous findings and contributes additional evidence to the current literature. The main findings contribute to the existing literature on how omnichannel shopping behavior changes when additional physical stores open and the present research investigated how these results can be used to predict future store openings.

5.2 MANAGERIAL IMPLICATIONS

The findings within this research suggest several actions for managers in retail real estate. This research aimed to predict how consumers use different shopping channels and how they integrate an additional physical store in their shopping behavior. Using the ML model, consumer behavior can be analyzed and predicted. The results are relevant for both retailers and managers in retail real estate.

First, the model estimation that has been performed in this research can also be implemented in other catchment areas where stores have been opened. This means that the type of data and the models that were used in this research can be implemented to model shopping channel choice behavior. The ML model can be used to predict shopping channel choice behavior given changes in the catchment area. These changes can, for example, be store openings and store closures in the catchment area, changing parking facilities (free parking changed to paid parking), and a new type of shopping alternative. This provides an insight into the effects of consumer behavior when the retail environment changes in a catchment area.

Secondly, an implication for retailers and managers in retail is that the ML model can predict shopping channel behavior and retailers can take this into account when opening new physical stores or closing existing physical stores. The predictive model that has been developed in this research can be used to predict changes in shopping behavior within a catchment area. The opening of an additional store has an impact on the overall shopping behavior of consumers in a catchment area. The predictive model gives more insight into the changing travel choices of consumers and how the opening of an additional physical store fits in the retail environment of a catchment area. However, one important notice is that this model is not timeless and will need to be updated regularly to keep up with current and future trends. Additionally, the model has some limitations that need to be considered when retailers use this model for predicting future store openings which are described in the next section. Thus, the current model is not a model that can replace all the other models that are used by a retailer, but it should be an addition to the existing models that are being used.

Furthermore, the increasing trend of online channel transactions indicates that the role of the physical store is changing. For managers in physical retail stores, it is important to offer convenience and focus on the consumer experience. Consumers are always online and have fast access to information using different channels to compare prices and evaluate products. Therefore, it is important to strengthen the role of the physical retail store by using online channels. Consumers must be attracted by online channels to visit the physical store. For instance, promoting the Click and Collect delivery to physical stores will increase the store visits of consumers. Another way is to attract consumers through targeted online advertising to persuade online-minded consumers to visit the physical store. To be successful, it is important to manage online channels and use targeted advertisements on online channels.

5.3 LIMITATIONS AND FUTURE RESEARCH

Research regarding changing omnichannel shopping behavior when opening new physical stores is scarce. This research adds some further insights into the preferences for shopping alternatives. However, this research has some limitations giving opportunities for future research. First, the results of this model might not be generalizable to other retailers and might not be able to predict shopping channel behavior in other countries outside of the Netherlands. However, the methodology and the data preparation and model estimation are generalizable and can be used to predict changes in shopping behavior when a physical store is opened or closed in a certain catchment area.

The physical channel and online channel transactions follow certain trends that have been determined in this research. In general, online channel transactions are increasing, and physical channel transactions have seen a decrease in the last few years. These trends have not explicitly been considered in this research. This means that the utility values that have been determined using the ML model are average values for the physical channel and online channel transactions, not taking into account ongoing trends. The average utility value for the online channel transactions may therefore be less accurate, as there is an increasing trend of online channel transactions in almost every catchment area. It would be interesting to take this into account in future research and to research the impact of the growing online trend and how this would affect the future of retailing.

Another issue that needs attention in predicting future store openings is the number of new consumers and the increase in the number of transactions that a physical store opening generates. The current research focuses on the consumer behavior of consumers already being a consumer. It is unclear how many new consumers that have never shopped at Decathlon will be attracted by opening new channels. The current model cannot predict the number of new consumers that the opening of a store generates and how many additional transactions that it generates. The model focuses on the current consumers and predicts what decisions the consumers in a catchment area will make. It would be interesting to research the increase in consumers when a new store opens, where they live, and how they behave. Next to this, it would be interesting to research the increase in the number of transactions per shopping alternative when a new physical store opens focusing on both new and existing consumers.

Furthermore, the current research measures the changes in the shopping behavior of consumers in a catchment area but does not consider the changes in revenues across the different shopping alternatives. The addition of a financial component in the model will create a full, practical model that can be used by retailers. The financial component will make it possible to predict the impact of revenues on the new physical store and on the existing shopping alternatives making it a complete model. It would be interesting to take this into account in future research.

Additionally, an issue that needs attention is the Covid-19 period that has a huge impact on omnichannel shopping behavior and might change the way that consumers shop in the future.

The number of online channel transactions overtook the number of physical channel transactions for the first time in 2020 due to Covid-19. It is important to identify the actual impact that Covid-19 has on omnichannel shopping behavior, but this is only possible when all the Covid-19 restrictions are removed. Therefore it is recommended to repeat this research roughly one year after all Covid-19 restrictions vanished to identify whether the results in this research are still relevant.

Despite these limitations, this research has contributed to the understanding of omnichannel shopping behavior. It shows how omnichannel shopping behavior changes when additional physical stores open and how these results can be used to predict the effects of future store openings.

LITERATURE

Aiello, L. M., Quercia, D., Schifanella, R., & Del Prete, L. (2020). Tesco Grocery 1.0, a large-scale dataset of grocery purchases in London. *Scientific Data*, **7**(1), 1-11. https://doi.org/10.1038/s41597-020-0397-7

Avery, J., Steenburgh, T. J., Deighton, J., & Caravella, M. (2012). Adding Bricks to Clicks: Predicting the Patterns of Cross-Channel Elasticities over Time. *Journal of Marketing*, **76**(3), 96-111. https://doi.org/10.1509/jm.09.0081

Bedgood, L. (2016). 90 Percent of Social Media Users Reach Out To Retailers. Why Social Media Is Your Secret Weapon. Retrieved from https://v12data.com/blog/90-percent-social-media-users-reach-out-retailers-why-social-media-your-secret-weapon/

Blázquez, M. (2014). Fashion Shopping in Multichannel Retail: The Role of Technology in Enhancing the Customer Experience. *International Journal of Electronic Commerce,* **18**(4), 97-116. https://doi.org/10.2753/jec1086-4415180404

Böttcher, M., Spott, M., Nauck, D., & Kruse, R. (2009). Mining changing customer segments in dynamic markets. *Expert Systems with Applications,* **36**(1), 155–164. https://doi.org/10.1016/j.eswa.2007.09.006

Briel, F. V. (2018). The future of omnichannel retail: A four-stage Delphi study. *Technological Forecasting and Social Change*, **132**, 217-229. https://doi.org/10.1016/j.techfore.2018.02.004

Briesch, R. A., Chintagunta, P. K., & Fox, E. J. (2009). How Does Assortment Affect Grocery Store Choice? *Journal of Marketing Research*, **46**(2), 176–189. https://doi.org/10.1509/jmkr.46.2.176

Brynjolfsson, E., Hu, Y.J., Rahman, M.S. (2013). Competing in the Age of Omnichannel Retailing. MIT Sloan Management Review, 54, 23-29.

Calderwood, E., & Freathy, P. (2014). Consumer mobility in the Scottish isles: The impact of internet adoption upon retail travel patterns. *Transportation Research Part A: Policy and Practice*, **59**, 192-203. https://doi.org/10.1016/j.tra.2013.11.012

CBS (2019a). Internet: access, usage, and facilities. Retrieved from https://opendata.cbs.nl/statline/#/CBS/nl/dataset/83429NED

CBS (2019b). In tien jaar tijd ruim 11 procent minder winkels. Retrieved from https://www.cbs.nl/nl-nl/nieuws/2019/51/in-tien-jaar-tijd-ruim-11-procent-minder-winkels

CBS (2020a). Detailhandel; omzetontwikkeling, index 2015=100. Retrieved from https://opendata.cbs.nl/#/CBS/nl/dataset/83868NED

CBS (2020b). Population, age, and 4-digit postal code information. Retrieved from https://opendata.cbs.nl/statline/#/CBS/nl/dataset/83502NED

Chen, X., Wang, X., & Jiang, X. (2016). The impact of power structure on the retail service supply chain with an O2O mixed channel. *Journal of the Operational Research Society*, **67**(2), 294-301. https://doi.org/10.1057/jors.2015.6

Childers, T. L., Carr, C. L., Peck, J., & Carson, S. (2001). Hedonic and utilitarian motivations for online retail shopping behavior. *Journal of Retailing*, **77**(4), 511-535. doi:10.1016/s0022-4359(01)00056-2

Clarke, G., Thompson, C., & Birkin, M. (2015). The emerging geography of e-commerce in British retailing. *Regional Studies, Regional Science,* **2**(1), 371–391. https://doi.org/10.1080/21681376.2015.1054420

De Blasio, G. (2008). Urban-Rural Differences in Internet Usage, e-Commerce, and e-Banking: Evidence from Italy. *Growth and Change*, **39**(2), 341–367. https://doi.org/10.1111/j.1468-2257.2008.00422.x

Doherty, N. F., & Ellis-Chadwick, F. (2010). Internet retailing: The past, the present and the future. *International Journal of Retail & Distribution Management*, **38**(11/12), 943-965. https://doi.org/10.1108/09590551011086000

Dunne, P., & Carver, J. (2013). *Retailing* (8th ed.). South-Western Cengage Learning.

Edelman, D. C., & Singer, M. (2015). Competing on customer journeys. *Harvard Business Review,* 88–100.

Eurostat (2020). Impact of Covid-19 crisis on retail trade. Retrieved from https://ec.europa.eu/eurostat/statistics-explained/index.php?title=Impact_of_Covid-19_crisis_on_retail_trade

Farag, S., Weltevreden, J., Rietbergen, T. V., Dijst, M., & Oort, F. V. (2006). E-Shopping in the Netherlands: Does Geography Matter? *Environment and Planning B: Planning and Design*, **33**(1), 59-74. https://doi.org/10.1068/b31083

Farag, S., Schwanen, T., Dijst, M., & Faber, J. (2007). Shopping online and/or in-store? A structural equation model of the relationships between e-shopping and in-store shopping. *Transportation Research Part A: Policy and Practice*, **41**(2), 125-141. https://doi.org/10.1016/j.tra.2006.02.003

Gensler, S., Leeflang, P., & Skiera, B. (2012). Impact of online channel use on customer revenues and costs to serve: Considering product portfolios and self-selection. *International Journal of Research in Marketing*, **29**(2), 192–201. https://doi.org/10.1016/j.ijresmar.2011.09.004

Green, M. A., Watson, A. W., Brunstrom, J. M., Corfe, B. M., Johnstone, A. M., Williams, E. A., & Stevenson, E. (2020). Comparing supermarket loyalty card data with traditional diet survey data for understanding how protein is purchased and consumed in older adults for the UK, 2014–16. *Nutrition Journal*, **19**(1), 1–10. https://doi.org/10.1186/s12937-020-00602-3

Grewal, D., Roggeveen, A. L., & Nordfält, J. (2017). The Future of Retailing. *Journal of Retailing*, **93**(1), 1-6. https://doi.org/10.1016/j.jretai.2016.12.008
Grewal, D., & Roggeveen, A. L. (2020). Understanding Retail Experiences and Customer Journey Management. *Journal of Retailing*, **96**(1), 3-8. https://doi.org/10.1016/j.jretai.2020.02.002
Groß, M. (2015). Mobile shopping: A classification framework and literature review. *International Journal of Retail & Distribution Management,* **43**(3), 221-241. https://doi.org/10.1108/ijrdm-06-2013-0119
Haas, A., & Kenning, P. (2014). Utilitarian and Hedonic Motivators of Shoppers' Decision to Consult with Salespeople. *Journal of Retailing*, **90**(3), 428-441. https://doi.org/10.1016/j.jretai.2014.05.003
Hagberg, J., Sundstrom, M., & Egels-Zandén, N. (2016). The digitalization of retailing: an exploratory framework. *International Journal of Retail & Distribution Management*, **44**(7), 694–712. https://doi.org/10.1108/ijrdm-09-2015-0140
Hägerstrand, T. (1970). What about people in Regional Science? *Papers of the Regional Science Association*, **24**(1), 6–21. https://doi.org/10.1007/bf01936872
Hensher, D. A., Rose, J. M. & Green, W. H. (2015). *Applied Choice Analysis*. Cambridge: Cambridge University Press.
Hernant, M., & Rosengren, S. (2017). Now what? Evaluating the sales effects of introducing an online store. *Journal of Retailing and Consumer Services*, **39**, 305-313. https://doi.org/10.1016/j.jretconser.2017.08.010
Hosseini, S. M. S., Maleki, A., & Gholamian, M. R. (2010). Cluster analysis using data mining approach to develop CRM methodology to assess the customer loyalty. *Expert Systems with Applications*, **37**(7), 5259–5264. https://doi.org/10.1016/j.eswa.2009.12.070
Hoyer, W., MacInnes, D., & Pieters, R. (2016). *Consumer Behavior.* Cengage Learning.
Huang, Z., & Benyoucef, M. (2017). The effects of social commerce design on consumer purchase decision-making: An empirical study. *Electronic Commerce Research and Applications*, **25**, 40–58. https://doi.org/10.1016/j.elerap.2017.08.003
Irawan, M. Z. (2015). Understanding the Effect of Online Shopping Behavior on Shopping Travel Demand through Structural Equation Modeling. *Journal of the Eastern Asia Society for Transportation Studies*, **11**, 614-625. https://doi.org/10.11175/easts.11.614.
Joewono, T. B., Rizki, M., Belgiawan, P. F., & Irawan, M. Z. (2020). Why do shoppers keep making online shopping trips? Learning from evidence in Bandung, Indonesia. *Asian Transport Studies*, **6**, 100016. https://doi.org/10.1016/j.eastsj.2020.100016
Kirby-Hawkins, E., Birkin, M., & Clarke, G. (2018). An investigation into the geography of corporate e-commerce sales in the UK grocery market. *Environment and Planning B: Urban Analytics and City Science*, **46**(6), 1148–1164. https://doi.org/10.1177/2399808318755147
Lee, R. J., Sener, I. N., & Handy, S. L. (2015). Picture of Online Shoppers. *Transportation Research Record: Journal of the Transportation Research Board*, **2496**(1), 55-63. https://doi.org/10.3141/2496-07
Lee, R. J., Sener, I. N., Mokhtarian, P. L., & Handy, S. L. (2017). Relationships between the online and in-store shopping frequency of Davis, California residents. *Transportation Research Part A: Policy and Practice*, **100**, 40-52. https://doi.org/10.1016/j.tra.2017.03.001
Lemon, K. N., & Verhoef, P. C. (2016). Understanding Customer Experience Throughout the Customer Journey. *Journal of Marketing,* **80**(6), 69-96. https://doi.org/10.1509/jm.15.0420
Levy, M., & Weitz, B. A. (2009). *Retailing Management*. McGraw-Hill Education.
Liang, T., & Lai, H. (2002). Effect of store design on consumer purchases: An empirical study of online bookstores. *Information & Management*, **39**(6), 431-444. https://doi.org/10.1016/s0378-7206(01)00129-x
Magestore. (2020). What is Omnichannel Retailing: Omnichannel Retail Guide: Magestore. Retrieved from https://www.magestore.com/retail-resources/omnichannel-retail-guide/what-is-omnichannel-retail/
Melero, I., Javier Sese, F., & Verhoef, P. C. (2016). Recasting the customer experience in today's omnichannel environment. *Universia Business Review*, **2016**(50), 18–37. https://doi.org/10.3232/UBR.2016.V13.N2.01
Melis, K., Campo, K., Breugelmans, E., & Lamey, L. (2015). The Impact of the Multi-channel Retail Mix on Online Store Choice: Does Online Experience Matter? *Journal of Retailing*, **91**(2), 272-288. https://doi.org/10.1016/j.jretai.2014.12.004
Mirzabeiki, V., & Saghiri, S. S. (2020). From ambition to action: How to achieve integration in omnichannel? *Journal of Business Research*, **110**, 1-11. https://doi.org/10.1016/j.jbusres.2019.12.028
Mokhtarian, P. L. (2008). Telecommunications and Travel: The Case for Complementarity. *Journal of Industrial Ecology*, **6**(2), 43–57. https://doi.org/10.1162/108819802763471771
Newcom (2020). Social media onderzoek 2020. Retrieved from https://www.newcom.nl/socialmedia2020?page=socialmedia2020
Nierop, E. V., Leeflang, P. S., Teerling, M. L., & Huizingh, E. K. (2011). The Impact of the Introduction and Use of an Informational Website on Offline Customer Buying Behavior. *SSRN Electronic Journal*. Published. https://doi.org/10.2139/ssrn.1772785

OECD (2020). COVID-19 and the retail sector: Impact and policy responses. Retrieved from http://www.oecd.org/coronavirus/policy-responses/covid-19-and-the-retail-sector-impact-and-policy-responses-371d7599/
Pantano, E., & Priporas, C. (2016). The effect of mobile retailing on consumers purchasing experiences: A dynamic perspective. *Computers in Human Behavior,* **61**, 548-555. https://doi.org/10.1016/j.chb.2016.03.071
Pantano, E., & Viassone, M. (2015). Engaging consumers on new integrated multichannel retail settings: Challenges for retailers. *Journal of Retailing and Consumer Services*, **25**, 106–114. https://doi.org/10.1016/j.jretconser.2015.04.003
Patel, J. D., Bhatt, N., Shukla, Y., & Gadhavi, D. (2015). Antecedents of rural and urban consumers' propensity to outshop and product specific outshopping behaviour. *Journal of Retailing and Consumer Services*, **26**, 97–103. https://doi.org/10.1016/j.jretconser.2015.05.011
Pauler, G., & Dick, A. (2006). Maximizing profit of a food retailing chain by targeting and promoting valuable customers using Loyalty Card and Scanner Data. *European Journal of Operational Research*, **174**(2), 1260–1280. https://doi.org/10.1016/j.ejor.2005.03.028
Pauwels, K., Leeflang, P. S., Teerling, M. L., & Huizingh, K. E. (2011). Does Online Information Drive Offline Revenues? *Journal of Retailing*, **87**(1), 1-17. https://doi.org/10.1016/j.jretai.2010.10.001
Puccinelli, N. M., Goodstein, R. C., Grewal, D., Price, R., Raghubir, P., & Stewart, D. (2009). Customer Experience Management in Retailing: Understanding the Buying Process. *Journal of Retailing*, **85**(1), 15-30. https://doi.org/10.1016/j.jretai.2008.11.003
Reilly, W. J. (1931). *The Law of Retail Gravitation*. New York: Knickerbocker Press
Reinartz, W., Wiegand, N., & Imschloss, M. (2019). The impact of digital transformation on the retailing value chain. *International Journal of Research in Marketing*, **36**(3), 350-366. https://doi.org/10.1016/j.ijresmar.2018.12.002
Reynolds, J., & Sundström, M. (2014). Digitalisation, retail transformation and change : what will European consumers want from their future shopping centre experience? Stockholm School of Economics, 5-6.
Scott, D. M., & He, S. Y. (2012). Modeling constrained destination choice for shopping: a GIS-based, time-geographic approach. *Journal of Transport Geography*, **23**, 60–71. https://doi.org/10.1016/j.jtrangeo.2012.03.021
Shobeirinejad, M., Veitch, T., Smart, J.C.R., Sipe, N. & Burke, M. (2013). Destination choice decisions of retail travellers: Results from discrete choice modelling in Brisbane. Australasian Transport Research Forum, ATRF 2013 - Proceedings.
Siebert, A., Gopaldas, A., Lindridge, A., & Simões, C. (2020). Customer Experience Journeys: Loyalty Loops Versus Involvement Spirals. *Journal of Marketing*, **84**(4), 45–66. https://doi.org/10.1177/0022242920920262
Sinai, T., & Waldfogel, J. (2004). Geography and the Internet: is the Internet a substitute or a complement for cities? *Journal of Urban Economics,* **56**(1), 1–24. https://doi.org/10.1016/j.jue.2004.04.001
Sprout Social (2020). Sprout Social Index 2020. Retrieved from https://sproutsocial.com/insights/data/index/
Stein, A., & Ramaseshan, B. (2016). Towards the identification of customer experience touch point elements. *Journal of Retailing and Consumer Services*, **30**, 8–19. https://doi.org/10.1016/j.jretconser.2015.12.001
Stirista (2020). The Ultimate Guide to the Customer Journey: Definition, Stages, Mapping, and an Example. Retrieved from https://www.stirista.com/ultimate-guide-customer-journey-definition-stages-mapping-example
Ström, R., Vendel, M., & Bredican, J. (2014). Mobile marketing: A literature review on its value for consumers and retailers. *Journal of Retailing and Consumer Services,* **21**(6), 1001-1012. https://doi.org/10.1016/j.jretconser.2013.12.003
Verhoef, P. C., Neslin, S. A., & Vroomen, B. (2007). Multichannel customer management: Understanding the research-shopper phenomenon. *International Journal of Research in Marketing*, **24**(2), 129-148. https://doi.org/10.1016/j.ijresmar.2006.11.002
Verhoef, P. C., Kannan, P., & Inman, J. J. (2015). From Multi-Channel Retailing to Omni-Channel Retailing. *Journal of Retailing*, **91**(2), 174-181. https://doi.org/10.1016/j.jretai.2015.02.005
Wagner, G., Schramm-Klein, H., & Steinmann, S. (2013). Effects of cross-channel synergies and complementarity in a multichannel e-commerce system – an investigation of the interrelation of e-commerce, m-commerce and IETV-commerce. *The International Review of Retail, Distribution and Consumer Research*, **23**(5), 571-581. https://doi.org/10.1080/09593969.2013.832697
Weltevreden, J. W., & Rotem-Mindali, O. (2009). Mobility effects of b2c and c2c e-commerce in the Netherlands: A quantitative assessment. *Journal of Transport Geography*, **17**(2), 83-92. https://doi.org/10.1016/j.jtrangeo.2008.11.005
Xi, G., Zhen, F., Cao, X., & Xu, F. (2018). The interaction between e-shopping and store shopping: Empirical evidence from Nanjing, China. *Transportation Letters*, **12**(3), 157-165. https://doi.org/10.1080/19427867.2018.1546797

Xu, X., & Jackson, J. E. (2019). Examining customer channel selection intention in the omni-channel retail environment. *International Journal of Production Economics*, **208**, 434-445. https://doi.org/10.1016/j.ijpe.2018.12.009
Zhen, F., Du, X., Cao, J., & Mokhtarian, P. L. (2018). The association between spatial attributes and e-shopping in the shopping process for search goods and experience goods: Evidence from Nanjing. *Journal of Transport Geography*, **66**, 291-299. https://doi.org/10.1016/j.jtrangeo.2017.11.007
Zhu, W., Timmermans, H., & Wang, D. (2006). Temporal Variation in Consumer Spatial Behavior in Shopping Streets. *Journal of Urban Planning and Development*, **132**(3), 166–171. https://doi.org/10.1061/(asce)0733-9488(2006)132:3(166)

APPENDIX

APPENDIX A: VARIABLE LIST

ID	Label	Value Name	Value Numeric
Cust_ID	Customer ID number	Numeric	
ID_Pur	Transaction ID number	Numeric	
Nalt	Number of alternatives	No physical stores in the catchment area	5
		1 physical store in the catchment area	7
		2 physical stores in the catchment area	9
		3 physical stores in the catchment area	11
		Covid-19 Lockdown	2
Index	Alternative	Purchase in physical store 1	1
		Purchase in physical store 2	2
		Purchase in physical store 3	3
		Purchase in store outside the catchment area	4
		Online purchase + Home delivery	5
		Online purchase + Pickup point delivery	6
		Online purchase + Click and Collect delivery	7
		Online order in physical store 1	8
		Online order in physical store 2	9
		Online order in physical store 3	10
		Online order in physical store outside the catchment area	11
Chosen	Choice	Not chosen	0
		Chosen	1
Pur_Tot	Total purchases by customer	Numeric	
Age	Age	Unknown	-1
		<35 years	0
		35-44 years	0
		45-54 years	0
		>55 years	1
Gender	Gender	Female	-1
		Unknown	0
		Male	1

Parking	Parking facilities	No parking facilities	0
		Paid parking facilities	1
		Free parking facilities	2
Size	Size of store (m^2)	Numeric	
Distance	Distance to store (km)	Numeric	
OnlH	Is the online purchase with home delivery?	Yes	1
		No	0
OnlP	Is the online purchase with pickup point delivery?	Yes	1
		No	0
OnlCC	Is the online purchase with Click and Collect delivery?	Yes	1
		No	0
Del	Is the purchase an online order in a physical store?	Yes	1
		No	0
ST	Is the purchase in a store?	Yes	1
		No	0
Dist3	Is the purchase in a store outside the catchment region?	Yes	1
		No	0
CST	Is the purchase in a store located in a city center?	Yes	1
		No	0
AST	The recoded variable Age multiplied by the dummy variable ST	18 – 35 years old * purchasing in a physical store	-1
		36-55 years old * purchasing in a physical store	0
		>55 years old * purchasing in a physical store	1
GST	The recoded variable Gender multiplied by the dummy variable ST	Women * purchasing in a physical store	-1
		Other/unknown gender * purchasing in a physical store	0
		Men * purchasing in a physical store	1

APPENDIX B: RESULTS MODEL ESTIMATIONS

Catchment area Eindhoven

Multinomial logit model

```
g'(-H)^-1g = 0.000204
successive function values within tolerance limits

Coefficients :
          Estimate  Std. Error  z-value  Pr(>|z|)
SizeST  4.5999e-04  1.5480e-06  297.145 < 2.2e-16 ***
DistST -7.9784e-02  5.3202e-04 -149.965 < 2.2e-16 ***
Dist3  -3.5789e+00  1.3280e-02 -269.500 < 2.2e-16 ***
ST      1.4260e+00  1.2417e-02  114.840 < 2.2e-16 ***
Del    -3.1900e+00  9.5930e-03 -332.531 < 2.2e-16 ***
OnlH    8.9085e-01  9.4638e-03   94.133 < 2.2e-16 ***
OnlP   -2.6739e-01  1.2044e-02  -22.202 < 2.2e-16 ***
AST     2.3744e-01  8.4189e-03   28.203 < 2.2e-16 ***
GST     1.2230e-01  4.5432e-03   26.919 < 2.2e-16 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Log-Likelihood: -389700
```

Mixed logit model

```
g'(-H)^-1g = 5.15E-08
gradient close to zero

Coefficients :
          Estimate  Std. Error  z-value  Pr(>|z|)
SizeST  4.7801e-04  1.7570e-06  272.061 < 2.2e-16 ***
DistST -1.1924e-01  6.8404e-04 -174.313 < 2.2e-16 ***
Dist3  -4.2102e+00  1.4173e-02 -297.060 < 2.2e-16 ***
ST      2.3555e+00  1.4069e-02  167.430 < 2.2e-16 ***
Del    -3.1900e+00  9.6066e-03 -332.063 < 2.2e-16 ***
OnlH    8.9085e-01  9.5888e-03   92.906 < 2.2e-16 ***
OnlP   -2.6739e-01  1.2186e-02  -21.943 < 2.2e-16 ***
AST     3.4051e-01  1.1364e-02   29.963 < 2.2e-16 ***
GST     1.6660e-01  5.6198e-03   29.645 < 2.2e-16 ***
sd.ST   1.8596e+00  9.3509e-03  198.870 < 2.2e-16 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Log-Likelihood: -362910

random coefficients
   Min.  1st Qu.   Median     Mean  3rd Qu. Max.
ST -Inf 1.101251 2.355541 2.355541 3.609832  Inf
```

Catchment area Amsterdam

Mixed logit model

```
g'(-H)^-1g = 7.47E-07
gradient close to zero

Coefficients :
          Estimate  Std. Error   z-value  Pr(>|z|)
SizeST  3.8055e-04  7.3989e-06   51.4327 < 2.2e-16 ***
DistST -1.1494e-01  9.0891e-04 -126.4591 < 2.2e-16 ***
Dist3  -3.7539e+00  1.4847e-02 -252.8445 < 2.2e-16 ***
ST      2.3417e+00  2.4310e-02   96.3259 < 2.2e-16 ***
Del    -2.8597e+00  1.1682e-02 -244.7896 < 2.2e-16 ***
OnlH    1.3812e+00  9.9323e-03  139.0610 < 2.2e-16 ***
OnlP    3.6640e-02  1.2368e-02    2.9626  0.003051 **
AST     3.5374e-01  1.1782e-02   30.0231 < 2.2e-16 ***
GST     2.0532e-01  5.7471e-03   35.7265 < 2.2e-16 ***
sd.ST  -1.7487e+00  9.5531e-03 -183.0498 < 2.2e-16 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Log-Likelihood: -284030

random coefficients
   Min. 1st Qu.   Median     Mean  3rd Qu. Max.
ST -Inf 1.16221 2.341691 2.341691 3.521171  Inf
```

Catchment area Utrecht

Mixed logit model

```
g'(-H)^-1g = 6E-08
gradient close to zero

Coefficients :
         Estimate  Std. Error   z-value  Pr(>|z|)
SizeST  1.5981e-04  8.6339e-06   18.5101 < 2.2e-16 ***
DistST -1.1774e-01  1.7610e-03  -66.8617 < 2.2e-16 ***
Dist3  -1.8563e+00  1.8892e-02  -98.2616 < 2.2e-16 ***
ST      2.3922e+00  3.6455e-02   65.6210 < 2.2e-16 ***
Del    -1.9101e+00  1.2447e-02 -153.4631 < 2.2e-16 ***
OnlH    1.3637e+00  1.3743e-02   99.2330 < 2.2e-16 ***
OnlP    2.6300e-01  1.6219e-02   16.2159 < 2.2e-16 ***
AST     1.1292e-01  1.6778e-02    6.7299 1.698e-11 ***
GST     1.9973e-01  7.8149e-03   25.5580 < 2.2e-16 ***
sd.ST   1.5410e+00  1.2127e-02  127.0746 < 2.2e-16 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Log-Likelihood: -153430

random coefficients
   Min.  1st Qu.   Median     Mean  3rd Qu. Max.
ST -Inf 1.352766 2.392182 2.392182 3.431598  Inf
```

Catchment area Rotterdam

Mixed logit model

```
g'(-H)^-1g = 2E-07
gradient close to zero

Coefficients :
         Estimate  Std. Error   z-value  Pr(>|z|)
SizeST  2.4176e-04  3.8259e-06   63.1917 < 2.2e-16 ***
DistST -1.7357e-01  9.7851e-04 -177.3860 < 2.2e-16 ***
Dist3  -3.0209e+00  1.3588e-02 -222.3271 < 2.2e-16 ***
ST      2.4905e+00  2.0209e-02  123.2388 < 2.2e-16 ***
Del    -2.7136e+00  1.1919e-02 -227.6762 < 2.2e-16 ***
OnlH    1.2179e+00  1.0103e-02  120.5422 < 2.2e-16 ***
OnlP   -6.8171e-02  1.2680e-02   -5.3762 7.608e-08 ***
AST     2.5470e-01  1.2256e-02   20.7813 < 2.2e-16 ***
GST     1.8443e-01  6.0552e-03   30.4581 < 2.2e-16 ***
sd.ST   1.5587e+00  9.3736e-03  166.2919 < 2.2e-16 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Log-Likelihood: -265810

random coefficients
   Min. 1st Qu.   Median     Mean  3rd Qu. Max.
ST -Inf 1.43916 2.490519 2.490519 3.541878  Inf
```

Catchment area Kerkrade

Mixed logit model

```
g'(-H)^-1g = 3.89E-07
gradient close to zero

Coefficients :
        Estimate Std. Error  z-value  Pr(>|z|)
DistST -0.1134002  0.0010793 -105.064 < 2.2e-16 ***
Dist3  -5.1656651  0.0231682 -222.964 < 2.2e-16 ***
ST      9.3822800  0.0976358   96.095 < 2.2e-16 ***
Del    -3.4437039  0.0137514 -250.427 < 2.2e-16 ***
OnlH    5.3353403  0.0960292   55.559 < 2.2e-16 ***
OnlP    3.8459947  0.0968216   39.722 < 2.2e-16 ***
AST     0.9240371  0.0164525   56.164 < 2.2e-16 ***
GST     0.1624846  0.0085310   19.046 < 2.2e-16 ***
sd.ST   1.9914087  0.0148215  134.359 < 2.2e-16 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Log-Likelihood: -122020

random coefficients
   Min.  1st Qu.  Median    Mean  3rd Qu. Max.
ST -Inf 8.039095 9.38228 9.38228 10.72546  Inf
```

Catchment area Apeldoorn

Mixed logit model

```
g'(-H)^-1g = 4.85E-07
gradient close to zero

Coefficients :
          Estimate  Std. Error  z-value  Pr(>|z|)
SizeST  5.2675e-04  1.5759e-05   33.426 < 2.2e-16 ***
DistST -1.0054e-01  1.0483e-03  -95.912 < 2.2e-16 ***
Dist3  -4.2964e+00  1.9337e-02 -222.181 < 2.2e-16 ***
ST      2.8536e+00  5.8057e-02   49.151 < 2.2e-16 ***
Del    -3.0934e+00  1.2458e-02 -248.312 < 2.2e-16 ***
OnlH    9.9444e-01  1.4707e-02   67.617 < 2.2e-16 ***
OnlP   -7.3109e-01  2.1572e-02  -33.891 < 2.2e-16 ***
AST     2.4215e-01  1.6535e-02   14.645 < 2.2e-16 ***
GST     1.1750e-01  8.4053e-03   13.979 < 2.2e-16 ***
sd.ST   1.7991e+00  1.4030e-02  128.232 < 2.2e-16 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Log-Likelihood: -140400

random coefficients
   Min.  1st Qu.   Median     Mean  3rd Qu. Max.
ST -Inf 1.640082 2.853559 2.853559 4.067035  Inf
```

All catchment areas

Mixed logit model

```
g'(-H)^-1g = 7.8E-07
gradient close to zero

Coefficients :
           Estimate  Std. Error   z-value  Pr(>|z|)
ParkST12  9.4884e-02  1.3514e-02    7.0212 2.199e-12 ***
SizeST    2.2287e-04  4.3550e-06   51.1760 < 2.2e-16 ***
DistST   -1.0376e-01  9.6293e-04 -107.7571 < 2.2e-16 ***
Dist3    -3.2831e+00  1.5153e-02 -216.6714 < 2.2e-16 ***
ST        3.1584e+00  2.2034e-02  143.3435 < 2.2e-16 ***
Del      -2.7730e+00  1.1836e-02 -234.2932 < 2.2e-16 ***
OnlH      1.1991e+00  1.1728e-02  102.2418 < 2.2e-16 ***
OnlP     -9.8665e-02  1.4805e-02   -6.6644 2.658e-11 ***
CST      -8.1033e-01  1.2324e-02  -65.7508 < 2.2e-16 ***
AST       3.0790e-01  1.3816e-02   22.2865 < 2.2e-16 ***
GST       1.4666e-01  6.8411e-03   21.4376 < 2.2e-16 ***
sd.ST     1.7950e+00  1.0900e-02  164.6786 < 2.2e-16 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Log-Likelihood: -216540

random coefficients
   Min. 1st Qu.   Median     Mean  3rd Qu. Max.
ST -Inf 1.94767 3.158386 3.158386 4.369102  Inf
```

APPENDIX C: TURNING POINT PHYSICAL STORE AND ONLINE SHOPPING

The turning point can be calculated when the estimated utility of shopping in a physical store overtakes the utility of purchasing products online. This can be done for every store located in the catchment area and will be explained per catchment area. The online shopping alternative with the highest utility is online purchase with home delivery. This variable will be used for the calculation as this is the point where the customer will have the same preference for shopping online (with home delivery) and going to a physical store. The estimated sum of utility of purchasing in a physical store includes three variables, namely the variables purchase in a physical store (ST), size of the store (SizeST), and distance to the store (DistST). Furthermore, the age (AST), and gender (GST) of every customer influence the utility of the physical store. The calculations for every store located in every catchment area can be seen in the tables below.

Turning point Eindhoven (Results in section 4.2.2)

The turning point for the store located in Best has a value between 29 and 37 kilometers. This means that when a customer has to travel more than 29-37 km (depending on age and gender) to the store in Best, the customer would on average prefer ordering online (with home delivery) rather than going to the store. For the store in Eindhoven, this would be the case if the customer has to travel more than 12 to 20 kilometers. The reason why customers want to travel shorter distances to the Eindhoven store is that this store is considerably smaller than the one in Best. The calculations for the store in Best and for the store in Eindhoven can be seen in the tables below.

Turning point store Best

Turning point store Best	Age 18 - 34 Gender female	Age 35 - 54 Gender female	Age > 55 Gender female	Age 18 - 34 Gender male	Age 35 - 54 Gender male	Age > 55 Gender male
ST	2.356	2.356	2.356	2.356	2.356	2.356
SizeST (5193 m2)	2.482	2.482	2.482	2.482	2.482	2.482
AST	-0.341	0	0.341	-0.341	0	0.341
GST	-0.167	-0.167	-0.167	0.167	0.167	0.167
Total	4.331	4.672	5.012	4.664	5.005	5.345
OnlH	0.891	0.891	0.891	0.891	0.891	0.891
Total - OnlH	3.440	3.781	4.121	3.774	4.114	4.455
DistST	-0.119	-0.119	-0.119	-0.119	-0.119	-0.119
Turning point (km)	28.9	31.7	34.6	31.6	34.5	37.4

Turning point store Eindhoven

Turning point store Eindhoven	Age 18 - 34 Gender female	Age 35 - 54 Gender female	Age > 55 Gender female	Age 18 - 34 Gender male	Age 35 - 54 Gender male	Age > 55 Gender male
ST	2.356	2.356	2.356	2.356	2.356	2.356
SizeST (965 m2)	0.461	0.461	0.461	0.461	0.461	0.461
AST	-0.341	0	0.341	-0.341	0	0.341
GST	-0.167	-0.167	-0.167	0.167	0.167	0.167
Total	2.310	2.651	2.991	2.643	2.984	3.324
OnlH	0.891	0.891	0.891	0.891	0.891	0.891
Total - OnlH	1.419	1.760	2.100	1.753	2.093	2.434
DistST	-0.119	-0.119	-0.119	-0.119	-0.119	-0.119
Turning point (km)	11.9	14.8	17.6	14.7	17.6	20.4

Turning point Amsterdam (Results in section 4.2.3)

The turning point for the store Amsterdam Arena has a value between 13 and 23 kilometers. This means that when a customer has to travel more than 13-23 km to Amsterdam Arena, the customer would on average prefer ordering online (with home delivery) rather than going to the store. For the store in Amsterdam Noord, this would be the case if the customer has to travel more than 14 to 23 kilometers. For the store Amsterdam Kinkerstraat, this would be the case if the customer has to travel more than 9 to 18 kilometers. The reason why customers want to travel shorter distances to the Amsterdam Kinkerstraat store is that this store is considerably smaller than Amsterdam Arena and Amsterdam Noord. The calculations for the three stores located in Amsterdam can be seen in the tables below.

Turning point store Amsterdam Arena

Turning point store Amsterdam Arena	Age 18 - 34 Gender female	Age 35 - 54 Gender female	Age > 55 Gender female	Age 18 - 34 Gender male	Age 35 - 54 Gender male	Age > 55 Gender male
ST	2.342	2.342	2.342	2.342	2.342	2.342
SizeST (2919 m2)	1.111	1.111	1.111	1.111	1.111	1.111
AST	-0.354	0	0.354	-0.354	0	0.354
GST	-0.205	-0.205	-0.205	0.205	0.205	0.205
Total	2.893	3.247	3.601	3.304	3.658	4.012
OnlH	1.381	1.381	1.381	1.381	1.381	1.381
Total - OnlH	1.512	1.866	2.220	1.923	2.277	2.630
DistST	-0.115	-0.115	-0.115	-0.115	-0.115	-0.115
Turning point (km)	13.2	16.2	19.3	16.7	19.8	22.9

Turning point store Amsterdam Noord

Turning point store Amsterdan Noord	Age 18 - 34 Gender female	Age 35 - 54 Gender female	Age > 55 Gender female	Age 18 - 34 Gender male	Age 35 - 54 Gender male	Age > 55 Gender male
ST	2.342	2.342	2.342	2.342	2.342	2.342
SizeST (3084 m2)	1.174	1.174	1.174	1.174	1.174	1.174
AST	-0.354	0	0.354	-0.354	0	0.354
GST	-0.205	-0.205	-0.205	0.205	0.205	0.205
Total	2.956	3.310	3.664	3.367	3.721	4.074
OnlH	1.381	1.381	1.381	1.381	1.381	1.381
Total - OnlH	1.575	1.929	2.283	1.986	2.339	2.693
DistST	-0.115	-0.115	-0.115	-0.115	-0.115	-0.115
Turning point (km)	13.7	16.8	19.9	17.3	20.4	23.4

Turning point store Amsterdam Kinkerstraat

Turning point store Amsterdam Kinkerstraat	Age 18 - 34 Gender female	Age 35 - 54 Gender female	Age > 55 Gender female	Age 18 - 34 Gender male	Age 35 - 54 Gender male	Age > 55 Gender male
ST	2.356	2.356	2.356	2.356	2.356	2.356
SizeST (225 m2)	0.086	0.086	0.086	0.086	0.086	0.086
AST	-0.341	0	0.341	-0.341	0	0.341
GST	-0.167	-0.167	-0.167	0.167	0.167	0.167
Total	1.934	2.275	2.615	2.268	2.608	2.949
OnlH	0.891	0.891	0.891	0.891	0.891	0.891
Total - OnlH	1.044	1.384	1.725	1.377	1.717	2.058
DistST	-0.115	-0.115	-0.115	-0.115	-0.115	-0.115
Turning point (km)	9.1	12.0	15.0	12.0	14.9	17.9

Turning point Utrecht (Results in section 4.2.4)

The turning point for the store Utrecht The Wall has a value between 12 and 17 kilometers. This means that when a customer has to travel more than 12-17 km to the store in Utrecht The Wall, the customer would on average prefer ordering online (with home delivery) rather than going to the store. For the store in Utrecht Vredenburg, this would be the case if the customer has to travel more than 10 to 16 kilometers. The calculations for the two stores located in Utrecht can be seen in the tables below.

Turning point store Utrecht the Wall

Turning point store Utrecht the Wall	Age 18 - 34 Gender female	Age 35 - 54 Gender female	Age > 55 Gender female	Age 18 - 34 Gender male	Age 35 - 54 Gender male	Age > 55 Gender male
ST	2.392	2.392	2.392	2.392	2.392	2.392
SizeST (4170 m2)	0.666	0.666	0.666	0.666	0.666	0.666
AST	-0.113	0	0.113	-0.113	0	0.113
GST	-0.200	-0.200	-0.200	0.200	0.200	0.200
Total	2.745	2.858	2.971	3.145	3.258	3.371
OnlH	1.364	1.364	1.364	1.364	1.364	1.364
Total - OnlH	1.382	1.495	1.607	1.782	1.895	2.007
DistST	-0.118	-0.118	-0.118	-0.118	-0.118	-0.118
Turning point (km)	11.7	12.7	13.7	15.1	16.1	17.0

Turning point store Utrecht Vredenburg

Turning point store Utrecht Vredenburg	Age 18 - 34 Gender female	Age 35 - 54 Gender female	Age > 55 Gender female	Age 18 - 34 Gender male	Age 35 - 54 Gender male	Age > 55 Gender male
ST	2.392	2.392	2.392	2.392	2.392	2.392
SizeST (3177 m2)	0.508	0.508	0.508	0.508	0.508	0.508
AST	-0.113	0	0.113	-0.113	0	0.113
GST	-0.200	-0.200	-0.200	0.200	0.200	0.200
Total	2.587	2.700	2.813	2.987	3.100	3.213
OnlH	1.364	1.364	1.364	1.364	1.364	1.364
Total - OnlH	1.224	1.337	1.449	1.624	1.737	1.849
DistST	-0.118	-0.118	-0.118	-0.118	-0.118	-0.118
Turning point (km)	10.4	11.4	12.3	13.8	14.7	15.7

Turning point Rotterdam (Results in section 4.2.5)

The turning point for the store Rotterdam Coolsingel has a value between 11 and 16 kilometers. This means that when a customer has to travel more than 11-16 km to the store Rotterdam Coolsingel, the customer would on average prefer ordering online (with home delivery) rather than going to the store. For the store in Rotterdam Alexandrium, this would be the case if the customer has to travel more than 7 to 12 kilometers. The calculations for the two stores located in Rotterdam can be seen in the tables below.

Turning point store Rotterdam Coolsingel

Turning point store Rotterdam Coolsingel	Age 18 - 34 Gender female	Age 35 - 54 Gender female	Age > 55 Gender female	Age 18 - 34 Gender male	Age 35 - 54 Gender male	Age > 55 Gender male
ST	2.491	2.491	2.491	2.491	2.491	2.491
SizeST (4361 m2)	1.054	1.054	1.054	1.054	1.054	1.054
AST	-0.255	0	0.255	-0.255	0	0.255
GST	-0.184	-0.184	-0.184	0.184	0.184	0.184
Total	3.105	3.360	3.615	3.474	3.729	3.984
OnlH	1.218	1.218	1.218	1.218	1.218	1.218
Total - OnlH	1.887	2.142	2.397	2.256	2.511	2.766
DistST	-0.174	-0.174	-0.174	-0.174	-0.174	-0.174
Turning point (km)	10.9	12.3	13.8	13.0	14.5	15.9

Turning point store Rotterdam Alexandrium

Turning point store Rotterdam Alexandrium	Age 18 - 34 Gender female	Age 35 - 54 Gender female	Age > 55 Gender female	Age 18 - 34 Gender male	Age 35 - 54 Gender male	Age > 55 Gender male
ST	2.491	2.491	2.491	2.491	2.491	2.491
SizeST (1821 m2)	0.440	0.440	0.440	0.440	0.440	0.440
AST	-0.255	0	0.255	-0.255	0	0.255
GST	-0.184	-0.184	-0.184	0.184	0.184	0.184
Total	2.492	2.746	3.001	2.860	3.115	3.370
OnlH	1.218	1.218	1.218	1.218	1.218	1.218
Total - OnlH	1.274	1.528	1.783	1.643	1.897	2.152
DistST	-0.174	-0.174	-0.174	-0.174	-0.174	-0.174
Turning point (km)	7.3	8.8	10.3	9.5	10.9	12.4

Turning point Kerkrade (Results in section 4.2.6)

The turning point for the store located in Kerkrade has a value between 26 and 45 kilometers. This means that when a customer has to travel more than 26-45 km to the store Kerkrade, the customer would on average prefer ordering online (with home delivery) rather than going to the store. The calculations for the store located in Kerkrade can be seen in the table below.

Turning point store Kerkrade

Turning point store Kerkrade	Age 18 - 34 Gender female	Age 35 - 54 Gender female	Age > 55 Gender female	Age 18 - 34 Gender male	Age 35 - 54 Gender male	Age > 55 Gender male
ST	9.382	9.382	9.382	9.382	9.382	9.382
SizeST (3416 m2)	0.000	0.000	0.000	0.000	0.000	0.000
AST	-0.924	0	0.924	-0.924	0	0.924
GST	-0.162	-0.162	-0.162	0.162	0.162	0.162
Total	8.296	9.220	10.144	8.620	9.544	10.468
OnlH	5.335	5.335	5.335	5.335	5.335	5.335
Total - OnlH	2.961	3.885	4.809	3.285	4.209	5.133
DistST	-0.113	-0.113	-0.113	-0.113	-0.113	-0.113
Turning point (km)	26.2	34.4	42.6	29.1	37.2	45.4

Turning point Apeldoorn (Results in section 4.2.7)

The turning point for the store located in Apeldoorn has a value between 33 and 40 kilometers. This means when a customer has to travel more than 33-40 km to the store Apeldoorn, the customer would on average prefer ordering online (with home delivery) rather than going to the store. The calculations for the store located in Apeldoorn can be seen in the table below.

Turning point store Apeldoorn

Turning point store Apeldoorn	Age 18 - 34 Gender female	Age 35 - 54 Gender female	Age > 55 Gender female	Age 18 - 34 Gender male	Age 35 - 54 Gender male	Age > 55 Gender male
ST	2.854	2.854	2.854	2.854	2.854	2.854
SizeST (3431 m2)	1.807	1.807	1.807	1.807	1.807	1.807
AST	-0.242	0	0.242	-0.242	0	0.242
GST	-0.118	-0.118	-0.118	0.118	0.118	0.118
Total	4.301	4.543	4.786	4.536	4.778	5.021
OnlH	0.994	0.994	0.994	0.994	0.994	0.994
Total - OnlH	3.307	3.549	3.791	3.542	3.784	4.026
DistST	-0.101	-0.101	-0.101	-0.101	-0.101	-0.101
Turning point (km)	32.9	35.3	37.7	35.2	37.6	40.0