Master Thesis: Real Estate Studies

The effects of the local central retail area on the purchasing decision;
Determinants of physical and online shopping behavior.

rijksuniversiteit
groningen

Abstract:

The purpose of this paper is to analyze the effect of area specific, consumer specific and geographical variables on the purchasing decisions by local consumers. The paper focuses on the central retail area in relation to the shopping behavior of local consumers. Characteristics such as size of the area, allocation within the area and the quality of the area are indicators were examined. The paper analyzes whether these variables have an effect on the shopping behavior of local consumers, when choosing between purchasing either physically or online. The results indicate several differences among product groups. Moreover, the results indicate that the size of the area has for most size categories no significant effect on the purchasing decision. Relatively large central retail areas, that exceed 60,000 sqm, do lower the odds that a local consumer will purchase online. In addition, a better perception and rating of the area in terms of accessibility and vitality are associated with a decrease in the likelihood that a local consumer will purchase a specific product online. Finally, the allocation of shops and the ratio between shops and Food & Beverage amenities have an effect on the shopping behavior in specific product groups. An allocation of shops that moves away from the mean allocation of central retail areas increases the likelihood that a local consumer will purchase a product online in specific product groups. Finally, when the ratio of F&B amities becomes too high this will increase the likelihood that a local consumer will purchase online.

Keywords: Retail, Consumer behavior, Central retail area, Online shopping.

Colophon

Title:	The effect of the local central retail area on the purchasing decision;
Subtitle:	Determinants of physical and online shopping behavior.

Document:	Master thesis Real Estate Studies
Date:	20-08-2017
Supervisor:	Ass. Prof. dr. X. (Xiaolong) Liu
Second reader:	Dr. M. (Mark) van Duijn

Author:	Justin Arnolli
Student number:	S 2219093
Phone number:	+31 6 55689101
Addres:	Amsterdamsestraatweg 967, Utrecht
E-mail:	j.r.arnolli@student.rug.nl

Rijksuniversiteit Groningen | University of Groningen

Faculteit Ruimtelijke Wetenschappen | Faculty of Spatial Sciences

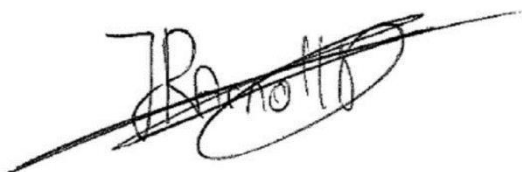Landleven 1, 9749 AD Groningen

Disclaimer:

Preface

This paper is the final project in completing my MSc Real Estate Studies at the University of Groningen. This challenging period has led to this final product titled: '*The effect of the local central retail area on the purchasing decision; Determinants of physical and online shopping behavior'*. The freedom within the MSc program allowed me to combine my studies with an internship at Colliers International, specifically at the Research & Consultancy department. This gave me the opportunity to gain experience in the past two years, in addition to gaining the valuable knowledge about the Real Estate market(s).

Moreover, this combination and the advice from both fields, has led to the current topic of my thesis. The retail market has been subject to several changes in the recent years and will probably remain a relevant topic in the near future. The topic for this thesis and the searching and collecting process for relevant data has been adjusted extensively along the way. However, the introduction of the Koopstromen onderzoek (KSO2016), conducted by I&O Research, offered the solution to the problems concerning the data availability for this analysis. Combined with the data offered by Colliers International, I was able to complete this paper.

I would like to thank my supervisor at the University of Groningen, dr. Xiaolong Liu, for his time, advice and guidance throughout the process. Moreover, I would like to thank the complete R&C team at Colliers International, and Maud van Vlerken and Bart Stek in particular, for their guidance during my internship at Colliers and their advice in the graduation process. Furthermore, I would like to thank Thijs Lenderink from I&O Research for providing the relevant data of the KSO2016 research, that was needed to complete the analyses in this paper. Finally, I would like to thank my family and friends who supported me during my Master's program. Although not everything went according to the original plan during the past two years, they never stopped supporting me.

I hope you will enjoy reading this thesis.

Groningen, August-20-2017.

Justin Arnolli

Executive summary

The purpose of this paper is to analyze the effect of area specific, consumer specific and geographical variables on the shopping behavior of local consumers. The shopping behavior has been subject to several changes in the previous decades, for instance with the introduction and manifestation of the online retail environment as a source of information and possible purchasing location. Moreover, preferences and intentions of consumers are argued to have changed with this transition.

This paper focuses on the central retail area in relation to the shopping behavior of local consumers. Characteristics such as size of and allocation within the area are indicators that are examined. In addition the quality of the area (evaluated by the consumers) is taken into account to answer the central research question of this paper:

What is the effect of the local central retail area on the consumer transaction decision?

This question is answered by introducing three sub-questions and five hypotheses concerning the characteristics of the central retail area. This paper makes use of a discrete choice logistic model to examine the effect of consumer' specific indicators and the state of the local retail area on the shopping decision. With this model, the likelihood that a specific consumer will either buy online or physically is examined. The paper uses a case study for the Randstad area in The Netherlands in combination with data from the KSO2016 research, performed by I&O Research, to test these hypotheses.

The results of the analyses of this paper indicate several differences among product groups. The results indicate that the size of the area has in general no significant effect on the purchasing decision. This indicates that the likelihood that a consumer will purchase a product online or offline is not significantly affected by the size of the area. However, the upper and lower bounds of the categories (central retail areas that exceed 60,000 sqm or are below 5,000 sqm) do influence the odds that a consumer will purchase online. Central retail areas that exceed 60,000 sqm offer a high(er) supply and large(er) variety of shops and shops located in these centers, this lowers the odds that a local consumer will purchase a product online. The opposite is true for the smallest central retail areas with a size under 5,000 sqm.

Moreover, a better perception of the area, expressed as a higher rating by consumers on the accessibility are associated with a decrease in the odds that a local consumer will purchase a product in specific product groups online. These findings give valuable insights to, for instance municipalities, who are able to improve the accessibility of a retail area. In addition, a higher rating on the vitality of a central retail area has a negative effect on the likelihood that a local consumer will purchase a product online for the Electronics and Apparel product groups.

In addition, the results indicate that several allocation variables have a significant effect on the likelihood of purchasing a product online in specific product groups. By implementing strategies that offer a better mix in terms of shops, policy makers could limit the increased probability that a consumer will purchase a product online and improve the competitiveness of the area. While F&B amenities are argued to increase the vitality of the area, a point can be reached where an increase in the ratio of F&B amenities in the central retail area reduces the supply of shops. This dominance of F&B is reached at 50% and increases the likelihood that a local consumer will purchase a product online. In conclusion, the characteristics of the local central retail area, although to a limited extent and for specific product groups, do have an effect on the shopping behavior of local consumers.

Previous empirical and theoretical studies that focus on why consumer make specific choices in the shopping process involve demographic and economic indicators to explain the demand fluctuations and shopping behavior. There is a consensus in the literature that demographic indicators, such as age, gender and possibly income, have an influence on the shopping decision. In addition to demographic variables, the location of a retail area, the distance and associated travel costs are opposed to influence the shopping destination decision process. The results of this paper are in line with this existing literature, with the exception of the effect of distance and gives a clarification on the effect of income on the shopping decision. The specific characteristics of the research area, the Randstad area in The Netherlands, limits the influence of distance on the shopping decision.

The literature on the effect of E-commerce on the demand for retail space are not always supported by quantitative analysis or sufficient data and are mostly survey based (Zhang et al. 2016). According to Weltevreden (2007) time based research is necessary to examine the magnitude of the effects of E-commerce and to overcome the subjective nature of the current literature. Therefore, more quantitative and time-based research is needed to improve the existing literature. The non-existence of this type of literature can be explained by the lack of transparency by data sources, privacy issues and associated costs. Quantitative data concerning, for instance, transaction values are subject to privacy limitations, for both the consumers and the registration companies, and are therefore not widely accessible. This explains why survey-based research is mostly dominant in this line of literature.

Table of contents

1. Introduction

1.1 Motivation

The retail market in the Netherlands has been facing difficulties concerning take-up levels and an increasing vacancy rate in the previous years. Between 2008 and 2016 the overall vacancy rate in the Netherlands increased from approximately 5% in 2008 to over 10% in 2016.[1] Simultaneously the Dutch retail sector was subject to a fast transition, in which the added value of Brick & Mortar (B&M) retail is changing. Traditionally, B&M retailers added value to the shopping process by providing basic information and services to its customers. However, information technology is taking over this role. In addition, the online shopping environment is nowadays a direct competitor to the physical retail stores. The purchasing decision from a consumer to either buy at a B&M store or choose to order online is the underlying determinant. The preferences and needs of the consumers have been modified over time. Between 2008 and 2016 several retailers were not able to adapt to the changing consumers demands, that require a stronger relation between customer and retailer and are mainly focused on the shopping experience. Therefore, these retail chains were unable to distinguish themselves from, for instance, the online alternative.[2] E-commerce, which includes the search for products and gathering information as a whole, or merely the alternative of shopping online has prominent role in the current retail market. Therefore this is nowadays part of the consumer behavior and related to the shopping decision. The spending behavior by consumers has a direct effect on the retailer and indirect implications for the strategy of (foreign) investors, developers and governmental organizations in terms of investment and planning decisions in the (near) future.

1.2 Central research question & contribution to the literature

1.2.1 Scientific relevance
E-commerce offers a new environment that offers information consumers require and offers an alternative purchasing location. Since the introduction of E-commerce within the shopping decision process, there has been a new line of literature concerning shopping behavior. However, the results are still ambiguous and open for debate. The existing literature focusses on the effect of demographic variables, such as age and level of education, and income related variables on the shopping behavior. These factors are partly proven to influence the decision

---

[1] Locatus data 2016

[2] Colliers International (2016): 'Transitie van de Nederlandse Winkelstructuur; Van waarde naar vitaliteit'

to substitute the physical stores for the online environment (Goswami & Khan, 2015). However, the effect of several demographic variables are to a large extent still open for debate. Moreover, a line of literature exists on the consumers' sentiment towards the (online) retail environment. For instance, consumer's trust in the online environment, the experience in (online) shopping of consumers and the marketing strategies adopted by websites are argued to have an effect on the behavior (Sautter et al., 2004; Zhou et al., 2014). On top of that, there is a line of literature that focuses on the possible effect of the local retail area as a determinant for the purchasing decision (Weltevreden & Rietbergen, 2006). However, this existing literature is outdated, since most of the research has been done in the first decade of the 21st century. The online shopping environment was recently introduced during this time period. New research is needed to examine if the results from these papers still hold after the developments in the retail market and the increase in the usage of the internet in the shopping process.

1.2.2 Central question and sub-questions

This paper will focus on the central retail area in relation to the offline and online shopping behavior of local consumers. Characteristics such as size of the area and the quality of the area (evaluated by the consumers) are indicators that could have a significant effect on the shopping process and eventually on the decision where to purchase a product. These variables have not been intensively investigated in the existing literature so far or are measured differently in the existing literature. The central research question of this paper is:

What is the effect of the local central retail area on the consumer decision to buy physically or online?

This paper will analyze the effect by examining five hypotheses that focus on either the effect of demographical characteristics, on the effect of the size of the central retail area, on the effect of the perception of consumers concerning the retail area or on the supply of shops and allocation within the area. The following sub-questions are formed to examine these effects:

1. *What is the effect of consumer specific characteristics on the decision to shop physically or online?*

This question will be answered by examining the existing literature on the effect of demographic characteristics on the shopping behavior and if these results still hold when these variables are included as control variables in the empirical analysis. Existing research on the effect of geographical location and urbanization argue that the demographic characteristics of the local residents are the underlying reason for the difference in consumer behavior. This paper will separately examine the effects of these demographic variables, add to the discussion of the effects of specific characteristics and will examine the effect in the four largest cities in the case study.

2. What is the effect of the size of the area and the perception by consumers *on the decision to shop physically or online*

The size and the quality of amenities in central retail areas could be important factors that possibly affect the decision of a consumer to substitute a physical shopping trip for an online purchase. However, the effects of the size of the central retail area on the decision to shop physically or online is yet to be observed. Arguably, the retail supply and quality of the stores and the retail area itself could influence the decision to undertake a trip towards this retail area. The retail area in the center of a city, town or village is the most common and most important retail area in the Netherlands and in most other western European countries. Therefore, the vitality and accessibility of this local central retail area located near(est) to the consumer is highly relevant for the decision where to shop (Weltevreden, 2007).

3. *What is the effect of the supply and allocation of amenities in the central retail area on the decision to shop physically or online*

The existing literature mainly focuses on the perception of the consumers and neglects to examine the observed quantitative data that is available. The relation between these quantitative measures of a central retail area, such as the diversity in shops, and the perception could help in forming new strategies of the retail area and its retailers in order to adapt to the changing retail market. A possible effect of these indicators, that include for instance the degree to which Food & Beverages (F&B) amnesties are located in the area, could have direct implications for the strategy of the retail area. Moreover, the allocation of shops in terms of sectors determines the variety of shops in the area.

1.2.3 Additional gaps in the literature

Existing literature on the consumer behavior, which includes online shopping, neglects to examine the importance of the consumer's choice in the shopping behavior (Zhang et. al, 2016; Goldmanis et. al., 2009). On the other hand, the literature concerning the consumer shopping decision is more marketing related and mostly focused on physiological and consumer specific characteristics (Goswami & Khan, 2015). So far, merely the line of literature dominated by Weltevreden has proposed the relation of geographical characteristics and consumer perception on the purchasing decision (Weltevreden, 2007; Weltevreden & Rietbergen, 2006; Farag et al., 2006).

This paper adds to the existing literature by Weltevreden in the following ways. Firstly, the existing research has been conducted in the first century of the 21st century. In the last ten years the (online) retail environment has been subject to a strong development, in which the usage of the online retail environment has increased significantly. Moreover, the intentions and preferences of consumers might have been subject to a change.[3] Therefore, new research is needed to ensure that the results from the papers written since the introduction of E-commerce still hold. Secondly, the variables examined in the literature are not aggregated into one model, indicating that the complete picture has not yet been given.

In addition, the effect among different product groups is examined in this paper. Therefore, the results can be attributed to specific product groups which has specific effects on a different set of retailers. Finally, For this paper a new division in terms of size of the local central retail area is used, mainly based on the categories used by the *'Koopstromen onderzoek 2016'* (KSO2016) research performed by I&O Research, to examine if the degree of size of the area itself has an influence on the transaction decision.[4]

1.2.4 Remaining parts of the paper

The empirical research is based on a case study for the Randstad area in the Netherlands. This paper makes use of a discrete choice (logistic) model to examine the effect of consumer' specific indicators and the state of the local retail area on the shopping decision. With this model the likelihood that a specific consumer will either buy online or offline is examined. This method has been used before within the literature on shopping behavior, and specifically on

---

[3] Colliers International (2016): 'Transitie van de Nederlandse Winkelstructuur; Van waarde naar vitaliteit'
[4] I&O Research (2016): 'KSO2016'

online shopping, mostly by Weltevreden et al. (Weltevreden & Rietbergen, 2006; Farag et. al., 2005)).

The rest of the paper is structured as follows: Section two will examine developments of the past years in the retail market. Section three will provide a literature review on the existing theoretical and empirical literature concerning the shopping decision and the changing retail environment. Furthermore it will provide the hypotheses of this research concerning the shopping decision. Section four will examine the data and methodology to conduct the empirical research for this case study. Section five will provide the estimation results of the empirical analysis and test the robustness of these results. Finally, section six will provide conclusions, acknowledgements and list suggestions for further research.

2. Developments in the retail market

To better evaluate the current retail market, this section will provide an explanation of the recent developments in the retail market, specifically in the Netherlands. Moreover, this section will provide the underlying mechanisms behind the changes in the market and the changes caused by the introduction of E-commerce in the market. This section will focus on the effect of the changes on the retailer and the market as a whole, where section three will switch to the perspective of the consumer. Most of the developments and difficulties in the market took place during difficult economic circumstances between 2008 and 2015. There are several trends noticeable in the previous years. According to Agency organization Dynamis the retail markets of 111 municipalities in the Netherlands are currently facing structural problems.[5] Several retailers were not able to adapt to the changes in the retail market and the adjusted preferences of consumers. In addition, the population in specific regions was declining and certain retail areas were unable to offer a suitable set of leisure activities or a desired overall shopping experience. These results are in line with other market reports that argue that there exists a polarization in the market.[6] The average size of shops in general has increased in the Netherlands during this period, while the number of shops has decreased. In terms of allocation the number of F&B amenities has increased in the Netherlands. The ratio of F&B amenities has increased in several retail areas in the last two years, by this take-up of vacant retail units.[7] These developments, combined with the technological developments and changing consumer

---

[5] Dynamis (2016): 'Sprekende cijfers Winkelmarkten'
[6] Colliers International (2016): 'Transitie van de Nederlandse Winkelstructuur; Van waarde naar vitaliteit'
[7] Colliers International (2016):'Transitie van de Nederlandse winkelstructuur'

behavior, should have an effect on the development strategies in local retail areas, on area management and strategies by municipalities.

The increase in information availability (online) has partly led to a change in consumer behavior that requires more personal and specialized advice. In addition, shopping experience is now the most important factor for Dutch consumers in deciding where to consume.[8] As a result, B&M retailers need to add a certain entertainment to the process to improve the in-store shopping experience. In addition, customers frequently consult the internet before they make an in-store purchase and compare prices online (Weltevreden, 2007). E-commerce is defined as searching for information and searching and/or buying goods and services via the internet (Mokhtarian, 2004) and online shopping is the final stage of E-commerce, at which the actual transaction takes place. The incorporation of E-commerce in the shopping behavior affected the level of knowledge and expectations of consumers. Potential customers are therefore more informed about products and prices when they enter the store and have a stronger bargaining position.

The final stage in the shopping process is to choose a transaction method. The introduction of E-commerce has led to an additional method. The decision where to actually purchase the desired products has an effect on the strategy retailers need to adopt. The usage of an online shop could have a competitive advantage over in-store shopping in terms of prices, availability/accessibility and product diversity. Technological innovations has led to the reduction of search costs and has increased the ability to compare between suppliers (Steel et. al. 2013). On the contrary, physical stores are able to offer a more extensive shopping experience that includes direct personal advice and possibly includes leisure activities (Dixon & Martson, 2010). Currently, shopping is one of the top-3 leisure activities in the Netherlands (after "going out" and "being outside"/recreation in nature).[9] Time-efficiency and product availability are less important factors when shopping becomes a leisure activity. Especially the mix of shopping facilities, F&B amenities and entertainment generate the success of a retail area. Therefore, there is a clear tradeoff between choosing a physical or an online transaction method.

Recent market reports indicate that by 2025 the fraction of online transactions could increase to approximately a quarter of the total sales in the Dutch retail market.[10] This increase is expected because the trust that consumers have for the online retailers will increase further. This has already led to an increase of cross-border online shopping worldwide in 2015.[11]

---

[8] Colliers International (2016): 'Transitie van de Nederlandse Winkelstructuur; Van waarde naar vitaliteit'
[9] Colliers International (2016): 'Transitie van de Nederlandse Winkelstructuur; Van waarde naar vitaliteit'
[10] ING Bank (2014): 'Winkelgebied 2025; Eindrapport' (Amsterdam 2014)
[11] Payvision (2016): 'Key business drivers and opportunities in cross-border E-commerce'

Although the fraction of online sales of the total retail market sales has increased in the past few years, the increase is not as severe as expected. Between 2011 and 2016 the total amount of online sales in the Netherlands increased from €4 billion in 2011 to approximately €6.5 billion forecasted in 2016. This indicates that online sales take up less than 7.5% of total retail sales in the Netherlands.[12] Moreover, between 2011 and 2015 total in-store sales only slightly decreased and are expected to increase again in the near future.[13]

E-commerce and physical shopping are also becoming more interwoven and could form as a complementary factor to each other. Insights in the motives of consumers are valuable for retailers in determining their strategy. The developments in the retail market are argued to either reduce the number of trips or to change the nature of the shopping trips.[14] It could influence the duration of trips, transportation mode and shopping destination. Therefore, online search behavior by the consumer before a purchase influences the shopping trip, but might not necessarily eliminate the trip (Farag et al., 2007). Moreover, less shopping trips do not necessarily result in lower in-store spending, but only affects the duration or frequency of shopping trips. On top of that, online activity can also be a complementary factor to B&M shopping, since it could generate shopping trips. For instance, online shopping could increase footfall and provide additional revenue in B&M stores nowadays when the store is combined with for instance a Collection-and-Delivery point (CDP).

According to Weltevreden (2014) the distinction between 'E-tailers' and B&M retailers is likely to diminish further in the future, because online shopping is used by B&M retailers as an additional channel to generate sales. Moreover, there are several companies that were founded on the internet that are opening physical stores nowadays, in order to maximize their performances. (Avery et al., 2012; Pauwels & Neslin, 2011). These retail brands, that were originally only active online, are opening physical stores throughout Europe.[15] These stores are used as CDP's, in-store sell mostly accessories to your online purchase and can be used for consulting purposes. This multichannel sales approach integrates the multiple strategies into one unique strategy.

Multi-channel strategies by retailers that combine an online platform with a physical store are therefore a solution in adapting to the modified consumer needs. The added value of the B&M store changes in this concept, however retaining its added value. According to several studies, the usage of a multichannel strategy increases customer acquisition and customer retention

[12] Deloitte (2016): 'Digital impact on retail in the Netherlands'
[13] Deloitte (2016): 'Digital impact on retail in the Netherlands'
[14] (Dxion & Martson, 2002; Cubukcu, 2001; Dixon & Martson, 2002; Bhat et al., 2003; Corpuz and Peachman, 2003; Tonn and Hemrick, 2004; Esser and Kurte 2005; Ferrell, 2005; Krizek et al., 2005).
[15] ICSC (2015): 'The Socio-Economic contribution of European Shopping Centres'

(Wallace et al., 2004). On top of that, a multichannel strategy has a positive effect on the frequency of shopping trips and the amount spend per trip (Venkatesan et al., 2007). Nowadays it is also possible to combine in-store shopping with home delivery by ordering products in-store. Moreover, the use of mobile devices in retail will cause that the physical and virtual world will become even more interwoven. Currently, 70% to 80% of the online purchases included the usage of mobile devices, varying from searching activities to purchasing products.[16] These developments indicate that also the retail market is still subject to a transition, in which the behavior of consumers is the driving force behind the changes.

3. Literature review

3.1 Determinants of the shopping decision

This section will give an overview of the most relevant literature concerning the effect of consumer characteristics on the shopping decision. Throughout this paper the shopping decision will include the tradeoff by a consumer either buy a desired product physically, specifically in the local central retail area, or online. There has been a rapid increase in the usage of the online environment, that allowed consumers and retailers to communicate with each other. Moreover, the online environment has been adopted by a broad scale of consumer segments for a variety of purposes (Häubl & Trifts, 2000). These purposes include the search for pre-purchase information and the online environment as a substitute for traditional shopping (Alba et al., 1997). According to Häubl and Trifts (2000), interactive tools designed to aid the potential consumer throughout the searching and purchasing process have a positive effect on the quality and efficiency of the shopping decision process. These aids create an environment in which better decisions are made by consumers with less effort. Therefore, the online environment has its own tools to affect the shopping decision of its potential customer base and targets different types of consumers differently.

Previous empirical and theoretical studies that focus on why a consumer makes specific choices in the shopping process involve demographic and economic indicators to explain the demand fluctuations. Although Hernandez et al. (2011) argue that the effects of demographic attributes are insignificant whenever consumers are more experienced e-shoppers, there is still a consensus in the literature that demographic indicators, such as age, gender and possibly income, have an influence on the shopping decision among a larger variety of

---

[16] Payvision (2016): 'Key business drivers and opportunities in cross-border E-commerce'

consumers. This line of literature will be used to answer the first sub-question and leads to the first hypothesis of this paper:

Hypothesis 1: Demographical consumer specific characteristics have an effect on the likelihood that a consumer will purchase a product online

Several demographic indicators are now separately discussed to examine what the possible effect of these characteristics will be. In addition to demographic variables, the location of a retail area, the distance towards the area and the associated travel costs are argued to influence the shopping decision process. These variables will be used as control variables in the analysis.

3.1.1 Gender

Difference in gender is argued to potentially affect the decision-making process of potential consumers, due to differences in the adoption of information and differences in trust and aversion to take certain risks (Ditmar et al, 2004. Sharma et al., 2012). A lack of trust and being more risk averse decreases the likelihood that a consumer will engage in an online transaction (Panda & Swar, 2013). Men are argued to be more utilitarian orientated when it comes to shopping, therefore valuing efficiency and effectiveness higher compared to women (Mattila et al., 2003; Dittmar et al., 2004). Thus, men are in general argued to be more of a task-focused shopper compared to women and are argued to be more functional in the purchase process in terms of accessibility and time consumption. Moreover, the social role theory leads to the fact that men are less risk averse (Walsh et al., 2008). Studies show that women experience more risk from online shopping (Bae & Lee, 2011). The experienced risk is a result of the expected negative utility that women experience whenever an order does not meet up to their expectations and the psychological risk that is formed by the uncertainty created by online shopping (Fosythe & Shi, 2003). The perceived uncertainty and lack of trust is caused by asymmetric information, which indicates that one side of the transaction has more information than necessary in the transaction process, concerning relevant factors of the transaction (Akerlof, 1970). Asymmetric information can therefore eliminate a transaction online when a consumer is not able to distinguish between online web shops that they can trust or not trust (Lee et al., 2005). This is more likely to occur when the consumer is a woman, due to the more risk averse approach in the shopping process. To finalize the gender debate, women are argued to perceive a higher utility from the physical evaluation of products (Ditmar et al., 2004). The combinations of perceived risks and the degree of efficiency suggests that men are more likely than woman to shop online.

3.1.2 Age

Several other person specific indicators will influence (online) spending behavior, such as age, household situation and education level. Mägi (2003) found that consumer characteristics influences the consumer's satisfaction on purchases and shopping experience. The differences in shopping behavior among age groups are a result of the aging process and the accumulated experience throughout an individual's life (Sharma et al., 2012). According to the *Information processing theory*, an increase in a consumers age makes him or her less reliant on additional information in the shopping decision (Ganesan-Lim et al., 2008; Yoon et al., 2005). The life experience on which the consumer can rely has in this sense a positive effect on the decision-making process. The combination of experience from age combined with experience in online shopping has a positive effect on the decision-making process (Cheung et al., 2014. Fang et al., 2016), which suggests that an increase in age will increase the probability that a consumer considers buying products online. However, the online environment is a technological innovation. Specific age groups, for instance above a certain age, are less familiar with the online environment and are unable to adapt to the rapidly changing online environment. Therefore, these age groups are less active and more dependent on the physical retail area. On the contrary, young adults grew up in the internet era and are more reliant on the online platform. On top of that, older consumers are argued to be more affected by the link between satisfaction and repeat purchase behavior (Cheung et al., 2014) while young consumers, who have less experience with the decision making process, are argued to rely more on the judgement of others and the information granted by the seller (Homburg & Giering, 2001). The online shopping environment offers several tools that let customers know about the product experience from previous consumers. Therefore, young consumers, who rely more on the experiences of others, are possibly more drawn to the online shopping environment compared to the physical store, due to the information availability on the web. Panda & Swar (2013) verify that young consumers are most likely to consume online. According to this paper this group experiences more 'ease of use' and 'usefulness' from buying online. In conclusion, the effect of age on the decision-making process in terms of choosing the online environment or the physical store might differ per age group, possibly also in combination with other consumer specific indicators such as the personal income level and the educational level.

3.1.3 Income

The paper by Benjamin, Jud and Winkler (1994) indicates that developments in the retail market are mostly explained by changes in total retail sales. The total spending on retail is directly related to the (disposable) income by consumers in a market. Therefore, a higher

income level will have a positive effect on retail spending. In the previous years, the fraction of total retail spending that was converted from physical spending towards online spending almost doubled between 2011 and 2016.[17] Therefore, a higher income level does not only result in a higher physical spending pattern, but also in an increase in the total online spending. The effect of income on the shopping decision is still ambiguous. Prices are an important distinction between offline and online shopping. In general, online stores are able to charge a lower price, due to the lower operational costs. This would suggest that consumers who are more concerned about prices, for instance due to a lower income level, are more likely to shift to online shopping. However, an increase in income also reduces the perceived risk by the consumers and thus increases the ease of use in the transaction process, which would result in a higher chance that a consumer would shop online (Chau & Hu, 2002; Hubona & Kennick, 1996). According to Hernandez et al. (2011) income does not have a significant effect on whether or not to shop online, whenever shoppers are more experienced in the online environment. On the contrary, Chiang and Dholakia (2003) do not find a significant effect of the income level on the consumer's shopping decision. Therefore, the effect of income on the shopping decision is still ambiguous.

3.1.4 Location and distance

The distance between a consumer and the store is argued to affect the likelihood of a transaction occurring at that store (Darley & Lim, 1999). This follows from the number of alternatives and opportunity costs that arise whenever the distance between a consumer and a store increases (Loudon and Della Bitta, 1993). According to Fox et al. (2004) the time that a consumer must travel to reach its desired products has a consistent negative effect on the expenditure of the consumer. The effects of distance is expressed in terms of travel time and travel costs (Bell et al., 1998). The existence of travel costs is a fundamental difference between the offline and online shopping format. Comparing products and prices and purchasing online reduces or eliminates these costs. In addition to costs related to travelling and time, there exists a third cost component associated with the distance between the consumer and the store, namely psychic costs. Psychic costs are costs that arise from the level of stress that a consumer experiences while purchasing a product. (Lusch & Lusch, 1997). These psychic costs increase whenever the effort by the consumer to reach the store and buy the product increases. Cairncross (2001) argues that the existence on the internet eliminates the need to travel. Therefore, the effect of geography and location could become

---

[17] I&O Research (2016): 'KSO2016'

non-existent in the future and offers consumers in more rural areas a direct alternative for travelling to a central retail area.

The existing literature suggests that price and convenience are important factors for a consumer in the determining where to shop (Burke, 1997; Lee et al., 2005; Peterson et al., 1997; Panda & Swar, 2013). The reduction in search/travel costs that the online environment offers could arguably outweigh the preference of the consumer to shop physically whenever this reduction in costs is considered large enough by the consumer. Therefore, when the distance towards the retail area increases the probability that a consumer would buy online increases as well (Farag et al. 2006). Especially, when it comes to the consumers that prefer efficient shopping behavior (run shopping), the distance between its living environment and the amenities/products it requires is positively related to the chance that the consumer will shop online. In terms of fun shopping, the maximum distance that a local consumer is willing to bridge in order to reach a (retail) environment that suits the requirements for fun shopping is larger.

3.2 Differences among product groups

Certain product groups are more suitable for online shopping, such as products from the entertainment industry (music, movies and books), electronics, home and living products and trips offered by traveling agencies. Therefore, the consumer behavior differs between these groups. According to Shen et al. (2016) product groups are either more *Utilitarian* or *Hedonic*, indicating that products are more practical and functional or more experiential and create enjoyment, respectively. Product categories such as books and DVD's as well as office supplies and computing equipment are considered more utilitarian in their nature, since these products are more homogeneous, straightforward and the products can be specified to a high level of detail (Shen et al., 2016; Lee et al., 2005). The more a product is considered to be a utilitarian product, the higher the chance that online shopping will form an alternative location to purchase the product. On top of that, Shen et al. (2016) argue that the complexity of the product affects the degree to which a product is suitable for online purchases. Product complexity is defined as the degree to which specific knowledge and expertise is necessary to correctly evaluate a product (McQuiston, 1989). More complex goods require more explanation and expertise to reduce the risks and uncertainty that come with the purchasing decision (McQuiston, 1989). Therefore, product categories that are more complex are in theory less likely to be purchased online, because offline retailing enables the consumer to inspect the product and directly ask for advice face-to-face (Shen et al., 2016).

3.2.1   Unique consumer and shopping characteristics

According to several studies, there is also an interaction effect between demographic indicators of consumers and different product groups (Lee & Johnson, 2002; Xu & Paulins, 2005). According to these papers young and more educated consumers are more likely to choose the online platform to buy products in the apparel category. Goswami & Khan (2015) argue that this group of consumers is more fashion conscious than other age and educational groups and that this group is in general more likely to engage in online shopping. In addition, there might also be a different effect between product groups for distance between a consumer and a store and travel time/costs (Darley & Lim, 1999). The willingness to travel a longer distance varies between products, where some products might attract customers from a relatively far distance almost eliminating the effect (Hawkins et al., 1998). This willingness is likely to increase whenever product groups are more orientated towards 'fun shoppers' and if the retail area offers more amenities than just retail stores. Therefore, it is likely that consumers are willing to travel longer and spent more on making the trip to for instance a larger city with an historic city center. Bell et al. (1998) express this by dividing the travel costs into fixed and variable costs. The fixed costs are associated with the direct travel costs which result from the trip towards the store. Variable costs are dependent on the consumers' 'shopping list', the products a consumer is intending to buy or bought during the trip, and the loyalty the consumer has towards the specific store or area. When the fraction of the variable costs of the total costs increases, consumers are generally more willing to travel further distances. Moreover, the experienced loyalty and other similar factors positively influences a consumer's level of utility, which have a positive effect on the willingness to travel further. Because certain product groups have on average a higher price per product, the variable costs of the shopping list are not only affected by the number of products on it, but mostly by the total costs of purchasing the list. This paper will separate the product groups within the case study and examine the effects on these groups individually.

3.3 Characteristics of the local central retail area

This paper will focus on the effects of the most used physical retail area in the Netherlands, the city center in the Randstad area and surrounding municipalities. The city center can be defined as 'an area, central to the city as a whole, in which the main land uses are commercial' (Guy, 1994, p. 14). More than 40% of all retail shops in the Netherlands are located in the central retail areas and central retail areas are also a common form of retail area throughout

Western Europe[18]. Therefore, the center of a city or town is one of the most important locations for consumers to shop. Next to the commercial use, the central retail area has a cultural and business purpose for the society (Guy, 1994). Arguably, a strong central area, in terms of size and diversity, nearby a local consumer could have an effect on the shopping decision of a local consumer. The size of the area resembles the potential of the area as a shopping, leisure, cultural and business location. The size of this area is subject to urban planning and natural boundaries, such as water, and is therefore relatively fixed in the short run. Retail activity in the Randstad area in the Netherlands, which includes the four largest cities in the Netherlands, is expected to increase in the near future (Bouman, 2012). In addition, there is a subsequent group of mid-sized cities that will also be likely to experience a slight growth in the coming years.[19] However, there are multiple studies that also argue that online shopping is a more urban phenomenon (Innovation-diffusion hypothesis: Farag et al. 2006; Farag et al. 2007; Weltevreden et al, 2005). This would indicate an increased likelihood that a consumer would purchase a product online when living in a more urban environment, even though the shop accessibility is relatively high in these areas. Therefore, there is no consensus on whether the size of a city and its central retail area has an effect on the shopping behavior by local consumers. The following hypothesis will be tested to determine the effect of the size of the local central retail area on the purchase decision of a consumer:

Hypothesis 2: The size of a central retail area in a town or city has an effect on the likelihood that a consumer will purchase a product online.

This hypothesis is dependent on the different categories in terms of size of the area. This paper will mostly follow the categories used by the KSO2016 research. The four largest cities (Amsterdam, Rotterdam, The Hague and Utrecht) of the Netherlands and cities with a total retail stock in the city center exceeding 60,000 sqm will enhance their competitiveness.[20] Therefore, the probability that a local consumer will purchase a product physically is expected to increase when the size of the central retail area exceeds 60,000 sqm. A small-town central retail area is mainly focused on daily necessities. Therefore, central locations in relatively small-towns (<10,000 sqm) are mainly suited for providing the local demand. However, not all central locations offer enough facilities to consume a certain set of products among specific product categories. Therefore, the low supply of local shops in these categories could increase

---

[18] Locatus 2016
[19] Syntrus Achmea (2015): 'De Nederlandse Winkelmarkt'
[20] I&O Research (2016): KSO2016

the probability that local consumers will shop online, due to the increased effort in terms of travel time, increased costs and the low attractiveness, relatively to larger retail areas (*Efficiency hypothesis:* Farag et al., 2006; Gillespie et al., 2001; Dixon et al., 2005).

Mid-sized centers (10,000 to 40,000 sqm) and town centers exceeding 10,000 sqm will be less competitive in the near future.[21] The competitiveness of centers between 40,000 sqm and 60,000 sqm is still ambiguous, however, this group is potentially comparable to the central retail areas between 10,000 sqm and 40,000 sqm. The reason why these centers are less competitive is due to the number of amenities and the shopping experience a central retail area can offer. The larger central retail areas offer a large variety of shops, activities and F&B amenities, which makes these centers more suitable for 'fun shopping'. The cities with a size within the upper and lower bound categories are argued to be too small to offer variety, but too large for only a focus on daily necessities. Thus, these central retail areas have a low variety of shops, mainly dominated by large retail chains, and not the atmosphere that attracts 'fun shoppers'. Moreover, these cities often contain small district centers within neighborhoods, that eliminate the need to travel towards the city center to buy the daily necessities. The probability that a local consumer will consume online might therefore increase when the size of the central retail area is between 10,000 sqm and 60,000 sqm.

In order to further distinguish between small central areas in these villages, this category is split up at the 5,000 sqm mark. Moreover, to better spread the observations over the categories the 10,000 to 40,000 sqm is split at the 20,000 sqm mark. Thus, six categories in terms of size of the central retail area are considered (namely, <5,000 sqm, between 5,000 sqm and 10,000 sqm, between 10,000 sqm and 20,000 sqm, between 20,000 sqm and 40,000 sqm, between 40,000 sqm and 60,000 sqm and finally >60,000 sqm).

The variety and diversity in terms of shops and characteristics of the area are frequently mentioned to be one of the key factors in the determining the attractiveness of a retail area. For instance, a retail area needs to be easily accessible to attract customers. (Rotem-Minndali & Salomin, 2007). Moreover, according to Darley & Lim (1999) the decision to travel to a specific store is a tradeoff between the inconvenience of the distance towards the store and the attractiveness of the store. A positive evaluation of the store increases the willingness of the consumer to increase the time to travel and spent money on the trip towards the store. Arguably this is not limited to only one specific store, as a large part of the consumers combine their shopping errands or shop for entertainment purposes. Therefore, the evaluation of the entire retail area is highly relevant for the shopping decision and a vital part of a store's attractiveness (Farag et al. 2006; Weltevreden & Rietbergen, 2006; Weltevreden et al, 2005).

---

[21] I&O Research (2016): KSO2016

To examine whether the attractiveness of the area influences the shopping behavior of local consumers, the following hypothesis is tested:

Hypothesis 3: A better perception of the local central retail area decreases the likelihood that a consumer will purchase online.

In order to examine if the perceptions of the consumers match the actual situation in the area and to introduce quantitative variables that offer valuable insights for policy makers, also the actual ratio/degree of shopping variety and degree of F&B amenities will be examined. The variety of shops in a central retail area could partly be explained by the allocation of shops in different sector. The more a sector becomes dominant, naturally reducing the supply of shops in different sectors , the more this could limit the variety of the supply of products in the area. In order to test if a different allocation of sectors in a central retail area influences the choice of a consumer whether or not to visit and buy at that area, the following hypothesis is tested:

Hypothesis 4: The diversity in shop allocation in the local central retail area influences the likelihood that a consumer will purchase online.

In addition, the supply within a central retail area does not only consist of shopping facilities. Also F&B amenities contribute to the attractiveness of an area. Municipalities nowadays frequently allow F&B amenities to take-up vacant retail location in order to reduce the overall vacancy rate and improve the attractiveness of the area.[22] Therefore, a higher degree of F&B amenities in a central retail area is expected to have a positive effect on the shopping experience, which would influence the shopping behavior of local consumers. Therefore the following hypothesis is tested in order to measure if the ratio of F&B amenities negatively influences the probability of a consumer' purchasing online:

Hypothesis 5: A higher ratio of F&B amenities in the local central retail area decreases the likelihood that a consumer will purchase online.

---

[22] Colliers International (2016): 'Transitie van de Nederlandse Winkelstructuur; Van waarde naar vitaliteit'

4   Data, model and descriptive statistics

4.1 Data

4.1.1 Data collection

This paper will conduct a case study in a diverse region in the Netherlands, namely the three provinces near the Randstad area and surrounding municipalities. The entire research area covers the three provinces of Zuid-Holland, Noord-Holland and Utrecht and several municipalities in the provinces of Noord-Brabant, Zeeland and Gelderland. For simplicity, the research area will be addressed as the Randstad area throughout this paper. The Randstad area includes the four largest cities in the Netherlands (Amsterdam, Rotterdam, The Hague and Utrecht), several mid-sized cities for Dutch standards and more rural areas. The analysis uses the data from the KSO2016 research, made available by I&O Research. The KSO2016 is a large-scale survey based research conducted every five years among (local) consumers in the Randstad area in the Netherlands. The KSO2016 research in the Randstad area was first conducted in 2006, also by I&O Research. Due to the large differences in the line of questioning it is unfortunately not possible to correctly compare the KSO2016 research with the research done in either 2011 or 2006. The aim of the KSO2016 research is to determine the developments in retail market by examining the recent spending behavior of consumers and the national and local economic developments. The research offers insights in the shopping behavior of (local) consumers in the Netherlands, specifically in the Randstad area. The research consists of over 100,000 surveys being conducted in over 120 municipalities between August and November 2016. To have a sufficient number of observations per region, a minimum of 385 per village/city was taken and a minimum of 210 surveys per retail area. In the survey conducted by I&O Research, respondents are asked to answer whether they most recently bought a product, in a specific product group, at a physical store, and if so in which shopping area or via an alternative method, such as online. This line of questioning was preferred by I&O compared to asking respondents where they usually buy these products. The latter method mostly results in a limited set of answers, while the preferred method does not exclude smaller retail areas or the online transaction method and is less affected by the respondent's perception and interpretation. The scale of the KSO2016 research is argued to be sufficient to adopt this method. Unfortunately, the questionnaire does not ask which products were purchased or any details concerning the price and information on whether this product was purchased in combination with other products. Moreover, the collection and distribution of the data among the different groups makes it difficult to integrate the results into one data set. In addition, the statistical programs available and used for the analyses would

be insufficient in using such a large dataset. Because the literature states that there is a consensus on the differences among product groups, the data is not integrated into one pooled set for the purpose of this research. Therefore, this paper will examine the product groups separately.

As a measure for the size of the center, the central retail area is used expressed as the size in square meters of units with a commercial space within the central area. This includes all shops, the F&B amenities and the vacant units. Therefore, this measure resembles the total potential size of the center that can be used for commercial use. The KSO2016 research offers insights on how consumers value the (central) retail area by asking the respondents to rate the area on different categories. To fully examine the effects of the perception of the central retail area by consumers, categories such as diversity in retail shops, the atmosphere and appearance, cleanliness and safety, reachability and accessibility, supply of F&B amenities and condition of other leisure facilities are considered. The data on the performance of the central retail area comes from an average ranking of the retail area given by the respondents of the KSO2016. Each respondent is asked to review the retail area for both the Apparel sector and a combined rating for the Home & living sector, which includes the other product groups examined in this paper. This paper uses these average rankings per product category as a general perception of the local central retail area by consumers. Subsequently, these ratings are linked to the local retail area near(est to) the respondent, regardless of the purchasing choice.

The survey used for the KSO2016 research splits up several product groups, such as daily necessities, luxury & fashion, and several product groups in the home & living sector, such as electronics. For the scope of this research the daily necessities sector and several other Home & Living sectors will be excluded, due to the still relatively low online penetration in the Netherlands of these sectors and because the type of real estate and locations required for these forms of retail are not easily comparable to the demand and supply in the central retail area. Respondents are also asked to answer questions concerning demographic characteristics, such as age, income, level of education and household composition. Appendix 1 shows a translation of the relevant question out of the questionnaire. The questions are identical for each of the product groups.[23]

The data on the size, supply and location of the central retail area, together with data on the allocation of shops and F&B amenities are obtained from Locatus. The location of the area is linked with the residential location of the respondent (on a six figures postal code scale) and distances between the respondents' residential location and the central retail area are

---

[23] The full survey can be consulted in Dutch at: http://www.kso2016.nl/downloads/overige-bijlagen/KSO2016-Vragenlijst.pdf

calculated using the 'Distance Calculator' tool ('AfstandBerekenen') in Excel. This Excel tool calculates the travel distance and time between two locations and the associated radius between the two locations. To analyze the effect of the four largest cities in general, four dummy variables will be introduced to eliminate any leftover effect of this specific group of cities.

4.1.2 Research area: Randstad area in The Netherlands

The region in which the research was conducted was the Randstad area in the Netherlands. The research area consists of all municipalities in the provinces of Noord-Holland, Zuid-Holland and Utrecht (plus the municipality of Nijkerk). To investigate the spillover effects of the area also the bordered regions were added. The research area is a mix of relatively large and urban municipalities (such as the four largest cities in the Netherlands) and relatively small and more rural areas. However, in general the research area can be considered to be relatively urbanized. The Dutch average population density is approximately 520 inhabitants per squared kilometer, while the density in the provinces of Utrecht, Noord-Holland and Zuid-Holland on average exceeds 1,000 inhabitants per squared kilometer.[24] Distances between villages and cities are relatively low compared to other Western or developed countries. Figure 1 shows the area in which the KSO2016 surveys were conducted. The orange regions are the key Randstad regions examined, the green regions are the bordered regions to control for spillover effects. The questions asked at both types of regions are identical. Therefore, both the 'Key' region and the 'Spillover region' can be used for the analysis.
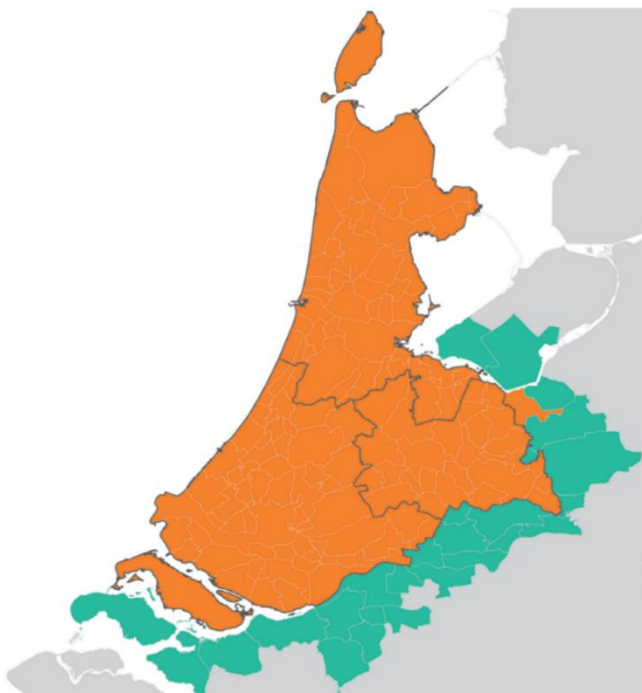
Figure 1: Research area KSO2016 (*Orange area is the key region; green is the bordered/spillover region):*

Source: Koopstromen onderzoek 2016 (I&O Research)

---

[24] CBS Statline 2017

4.2 Descriptive Statistics

Table 1 shows the descriptive statistics of the original KSO2016 research survey.

Table 1: Descriptive statistics of the KSO2016 survey research

Surveys per region		
Province	**Frequency**	**Percentage**
Noord-Holland	30,695	30%
Zuid-Holland	48,860	48%
Utrecht	18,866	19%
Spillover region	3,359	3%
Total	**101,780**	**100%**
Educational level		
Level of education	**Frequency**	**Percentage**
Primary & Lower Secondary & Upper first phase	25,057	25%
Post Lower Secondary & Upper Secondary second phase	32,993	32%
Bachelor	27,720	27%
Master/Doctoral	14,668	14%
Unknown	1,342	1%
Total	**101,780**	**100%**
Gender distribution		
Gender	**Frequency**	**Percentage**
Male	46,237	45%
Female	55,383	54%
Unknown	160	0%
Total	**101,780**	**100%**
Household composition		
Household composition	**Frequency**	**Percentage**
One-person	23,721	23%
Single parent	4,308	4%
Two-person	44,441	44%
Family with kids	25,443	25%
With parent(s)/legal guardian(s)	2,155	2%
Other	1,712	2%
Total	**101,780**	**100%**
Income level		
Net Income per month per household	**Frequency**	**Percentage**
< €950	3,001	3%
between € 951 & € 1,300	6,910	7%
between € 1,301 & € 1,900	13,911	14%
Between € 1,901 & € 3,150	30,509	30%
> € 3,150	26,240	26%
Unknown	21,209	21%
Total	**101,780**	**100%**

Table 1 continued: Descriptive statistics of the KSO2016 survey research

Age distribution		
Between ages	**Frequency**	**Percentage**
15-24	3,219	3%
25-34	9,333	9%
35-44	12,407	12%
45-54	19,772	19%
55-64	24,574	24%
65-74	24,384	24%
>74	7,500	7%
Unknown	591	1%
Total	**101,780**	**100%**

The survey was successfully conducted among 101,780 respondents. The survey was sent to a random sample of potential respondents and it was possible to fill in the survey in three ways. Respondents were contacted and asked to submit the survey either by mail, online or by phone. Approximately 37% of the surveys were submitted by mail, 56% of the surveys were submitted online and approximately 7% of the surveys were conducted by phone.

Age groups are for this paper divided into categories of ten years. For the surveys the minimum age requirement was fifteen years. The respondents rate is higher among relatively older people that were asked to fill in the survey. Therefore, compared to the average allocation in terms of ages in the Netherlands, the sample contains a relatively large fraction of persons between the ages 55 and 75. The sample has a relatively low fraction of persons between 0 and 25 and over 75, however these age groups are argued to be less active in the retail market. Nevertheless, the sample is somewhat biased to relatively older consumers. Compared to the average level of education, the sample is relatively high educated. Compared to the average levels in the Netherlands, the fraction of at least Bachelor graduates is higher (+12%) and the fraction of high school graduates as highest form of education is lower (-12%).[25] This can partly be explained by the relatively old sample, which lowers the number of respondents still in the process of completing a study. This issue is also noticeable in the income distribution as older respondents often receive a relatively high income. In terms of household composition, the fraction of one-person households in the survey is lower than the average in the Netherlands (38%) and the fraction of two-person households exceeds the national average (29%).[26]

The product groups are categorized by the KSO2016 research and are individually analyzed in this paper. Table 2 shows the different product categories that are being evaluated, including

---

[25] 'Enquête Beroepsbevolking (EBB)' (Survey of Working population), Onderwijs in cijfers

[26] CBS Statline 2016

several examples of products that belong to these groups and the associated number of observations that are used in the empirical analysis. The number of correct surveys are the surveys that were filled in sufficiently to use in the analyses and are a selection out of the total number of 101,780 surveys conducted in the KSO2016 research. Appendix 2 contains more detailed descriptive statistics tables specifically for each of the examined product groups mentioned in table 2. Notably, the ratio between transactions being made either online or physically already suggests a larger influence of the online retail environment on specific product groups. The Apparel and Sports groups show the least transactions online, while this figures increase for Electronics and even more for products in the Media product group. To ensure the most optimal results for the likelihood ratios, only complete surveys were included in the model estimation. Therefore, when an answer regarding the demographic characteristics or purchasing location was unclear or missing, the respondent was taken out of the final sample.

Table 2 Different product categories of KSO2016

Category	Includes	Number of correct surveys
Electronics	*TV/radio/computers, Refrigerators etc.*	67,273
Apparel	*Clothes and luxury items (such as jewelry)*	70,885
Media	*Books, DVD's, CD's and hobbies*	59,781
Sports	*Sports equipment, bicycles*	52,065

4.3 Model

Equation 1 shows the base model used for the analysis of this paper. A non-linear, discrete choice model is used to determine the probability of occurrence. For the purpose of this research the respondents who never buy products in this specific product category, respondents that do not remember where and when the last purchase took place or respondents who used an alternative method such as on purchasing a product on TV are excluded from the analysis for the specific product group. This results in a binary output of a product being purchased either physically or online. The line of questioning makes the data suitable for a logistic regression model, in which the output of the dependent variable is either physical or online. In the analysis Y has a value of 1 if a consumer purchased a product online and 0 if the product was bought physically.

A logistic regression determines the impact of independent variables presented simultaneously to predict whether the dependent variable belongs to either one or the other category (Moore & McCabe, 2014). The regression uses the maximum likelihood methodology, which maximizes the probability of allocating the observed data into the right category of the

dependent variable. The probability is calculated using the following formula (odds ratio) (Moore & McCabe, 2014):

$$Log\left(\frac{p}{1-p}\right) = \alpha + \beta_i x_i \tag{2}$$

Or,

$$P = \frac{e^{(\alpha + \beta_i x_i)}}{1 + e^{(\alpha + \beta_i x_i)}} \tag{3}$$

Where

P = probability of Y=1.

Exponent (e) = base of natural logarithms

α = constant

β = vector of regression coefficients

x = vector of independent variables

Multiple regressions will be done, for each of the product groups individually, to distinguish between the product groups

$$Y^* = \alpha + \beta_i * [Size_i] + \beta_j * [R_j] + \beta_k * \ln(D_{con}) + \beta_l * [CS_m] + \beta_m * [I_m] + \beta_n * \ln(RS_n) + \varepsilon \tag{1}$$

The Y^*- variable is a binary dependent variable, limited in range to have either a value of 0 or 1. The variable *Size* resembles a vector of the categories for the size of the central retail area that were determined in order to answer the first hypothesis. *R* is a vector of the ratings of the retail area to answer the second hypothesis. Respondents were asked to rate the central location in which they last bought a product in either the Apparel sector or the Home & Living sector, which consists of the other product groups examined in this paper. Respondents had to score the area on several categories varying from atmosphere/vibe and cleanliness, safety and accessibility, variety of retail supply and other facilities on a 10-point scale. The descriptive statistics concerning the ratings can be found in Appendix 2. Higher ratings of the retail area on the categories are expected to decrease the probability to shop online.

D is the natural logarithm of the distance between the consumer's residential location and the nearest central area that offers enough facilities to consume a product in the analyzed categories. Several small-town locations do not offer suitable facilities to consume for instance products in the apparel and Home & Living product categories. To correctly measure the

distance between a consumer and a suitable central retail location, several towns were excluded and the distance variable is adjusted to the nearest suitable central retail area. Appendix 3 provides the list with the adjustments that were made. The distance is measured from the residential location of the consumer to the (newly assigned) central retail area. The variable is transformed to the natural logarithm of the distance in meters to increase the normality of the distribution of the variable.

CS is a vector of the consumer specific control variables: Income, gender, age and educational level. *I* is a vector of several quantitative variables that relate to the central retail area, such as the deviation from average shop allocation (Stand. Dev.) and degree of F&B amenities (F&B). The Stand. Dev. variable is the standard deviation of a retail area based on the average allocation of number shops in supply in a central retail area. Due to differences in allocation between city sizes in general, the average allocation was calculated for different size categories (<10,000 sqm, between 10,000 sqm and 40,000 sqm, between 40,000 sqm and 60,000 sqm and >60,000 sqm) and is based on the entire country. When the size of a center increases, the fraction of shops in the apparel and leisure sectors in the center increases, while the number of shops in the home & living sector, transport sector, daily necessities sector and services sector decreases. The fraction of the number of facilities in the leisure sector remains relatively constant among the different size groups.

The variable F&B estimates the percentage of number of F&B amenities in the central retail area to check whether the degree of F&B amenities increase the experience factor of shopping and has a negative effect in the likelihood that a consumer will purchase online. RS is a control variable for the additional retail stock in the city, town or village in which the respondent is living, defined as the total retail stock in a town or city in sqm minus the stock of the central retail area in sqm. As discussed in previous sections, the central retail area is the most common form and most used form of retail area by Dutch consumers. However, several locations, mostly the relatively larger cities, do offer different types of retail areas, such as district centers and/or large-scale shopping centers. Although most of the transactions reported by the respondents were made either at a central retail location or online, part of the observations took place at either a district center or large-scale shopping center. To control for these alternative forms of shopping locations, the additional retail stock (not including the central area) is added as a variable. The variable is transformed to the natural logarithm of the additional supply in sqm to increase the normality of the distribution of the variable. Finally, the α is a constant and ε the error term.

4.4 Model analysis

4.4.1 Odds ratios

The regression output of a logistic regression differs from a more generally known regression analysis, such as an Ordinary Least Squares regression. The Logistic regression model estimates coefficients that are Odds ratios. Odds ratios determine the likelihood ratio of two variables instead of general coefficients and should therefore be interpreted differently. The Odds ratio can be expressed as the increased likelihood that the dependent variable will occur by a unit shift in the independent variable. When the odd ratio is less than one, the likelihood of y=1 occurring decreases, while an Odds ratio above one increases the likelihood of y=1 occurring. When the Odds ratio is equal to 1, there is no relation between the independent and dependent variable. As an example, when the Odds ratio estimate has a value of 2 the situation in y is two times more likely to occur with a one unit increase in x, ceteris paribus, or the odds of occurring increased by 100%. In the case of a binary independent variable, instead of a continuous variable, the Odds ratio estimates the increased likelihood that y=1 will occur over the base case. In order to calculate a probability the Odds ratio needs to be divided by (1 + Odds ratio). In the example case the probability of y=1 is approximately 0.6667.

4.4.2 Analysis and introduction of the independent variables

The perception of consumers is not always in line with reality. Consumers might experience the variety of shops in a retail area or the existence of F&B amenities differently from the actual situation. This is likely to be due to a mismatch in quality and quantity of these factors. Therefore, two variables were introduced to distinguish between the difference of the perception of the consumers, expressed as the rating given by respondents, and the actual allocation of the supply of shops in the central retail area. Table 3 shows the correlation estimations between two of these variables to check if the perception of the consumers match the actual situation in the retail areas of the sample. An increase in the standard deviation from the mean allocation of shop sectors is expected to be negatively correlated with the rating for variety in shops. Furthermore, a higher percentage of F&B amenities in the supply of the retail area is expected to be positively correlated with the rating given by consumers on the supply of F&B amenities. The correlations of both the shops allocation and the supply of F&B have the expected sign. However, in both cases the correlations are relatively low. In the case of F&B amenities this might be caused by the perception of the consumer in rating the F&B amenities. Consumers might not increase or decrease their ratings because of the number of amenities, but on the quality of them and the ability to stand out in the area. The correlation between the number of F&B amenities and the ratio of the F&B of the total supply are relatively

equal. The low correlation of the standard deviation of the sector allocation in shops might not resemble the utopic distribution that consumers intend. Logically, this distribution is also consumer specific and changes over time.

Tabel 3: Correlations of Consumers' perception and actual supply/allocation

Food & Beverages amenities				Sector allocation shops		
	Number of F&B	% F&B of total supply	Rating F&B		$\sigma\_Allocation$	Rating variety in supply
Number of F&B	1,000			$\sigma\_Allocation$	1,000	
% F&B of total supply	0,559	1,000		Rating variety in supply	-0,116	1,000
Rating F&B	0,266	0,277	1,000			

To exclude any possible chances of multicollinearity among the independent variables, a correlation matrix was calculated for the demographic control variables in the model (Appendix 4). The correlation matrix does not show any concerns for multicollinearity among the demographic variables in the model. However, a correlation matrix among the rating categories does show significantly high correlations among each other (Table 4). In addition, the central retail areas that exceed 60,000 sqm are relatively heavily correlated with the ratings for Accessibility and Diversity in the supply of shops. A larger central area might be argued to be more difficult to reach, more subject to congestion and parking becomes more difficult. Furthermore, the total supply of shops in a larger central area is higher. A higher supply might offer more exclusive shops and in general might increase the perception of variety.

Table 4 Correlations among Ratings

	RatingAtmospere	RatingFeeling	RatingAccess	RatingOther	RatingF&B
RatingAtmospere	1				
RatingFeeling	0.2845	1			
RatingAccess	-0.1517	0.6301	1		
RatingOther	0.8464	0.0212	-0.2426	1	
RatingF&B	0.6825	0.0021	-0.3505	0.8078	1
RatingSupply	0.7437	0.0392	-0.314	0.5936	0.5697

In order to reduce the possible serial correlation between these variables, the ratings need to be transformed. Firstly, the ratings for Atmosphere, Feeling Other facilities and activities are taken together to form one rating. The same is done for F&B and Variety in the supply of shops and the Accessibility ratings. Correlations between the variables and these ratings show sufficiently low scores with these new rating variables.

To further investigate the possibility of multicollinearity, a Variance Inflation Factor (VIF) was estimated for all variables in the regression. The model excluding the ratings shows relatively

low VIF scores among the variables, with the exception of certain control variables and the ratings of the central retail area. The VIF scores for the demographic variables do not exceed any critical levels. Values exceeding ten usually indicate possible multicollinearity issues (Long & Freese, 2004). VIF scores of the continuous variables additional supply, distance of the central retail area high and exceed this critical level. Because only two of the control variables exceed the critical value of ten, the risk of multicollinearity in the model is argued to be reduced sufficiently to correctly interpret the coefficients and standard errors in the models.

In the second set of models, the newly formed ratings do show relatively high VIF scores that exceed the critical value of 10. To decrease the risk of multicollinearity a Principal Component Analysis (PCA) was used to analyze the possibility of reducing the number of rating variables. A PCA analysis eliminates the correlation problem by linearly transforming the set of rating variables, creating a smaller set of variables that are uncorrelated with each other, while remaining the maximum variance among the original variables (Weltevreden & Rietbergen, 2007). The goal of the analysis is to extract the most relevant information, while reducing the size of the total data set. Moreover, the analysis simplifies the dataset by computing new variables named principal components (linear combinations), which resemble the original variables of the dataset. (Abdi & Williams, 2010). Although using a PCA analysis eliminates the multicollinearity issues, it reduces the usefulness of the different ratings on different detailed categories given by the respondents. Therefore, the newly formed variables are only able to resemble a set of ratings instead of the detailed separate ratings given in different categories. Appendix 5 shows the output tables and the Screeplot of the analysis. The first two components show eigen values significantly above the value of 1 and are therefore used in the calculation of the new predicted scores. As expected, the table shows a clear separation between ratings on the accessibility of the area and ratings on the vitality of the area. The new predicted values of these two components are used as the new rating variables in the second model. Appendix 6 shows the output of the VIF analysis for each specific product group models after the PCA.

5. Model analysis, empirical analysis and results

5.1 General results of the models

Table 5 shows the empirical results of the analysis per product group for the Logistic regression models. The first model per product group does not include the ratings variables, since not all of the central areas examined in this sample include a rating. In the second set of models in

Table 5 the full model, including the rating for vitality, is the included. Including ratings slightly reduces the total number of observations. The table shows the Odds ratios and the corresponding standard errors of the coefficients. The base case is a female consumer, between the ages of 55 and 65, with a Two-person household composition, a household income between €1,900 and €3,150 per month, educated at a 'Mid'-level (finished high school or equivalent), living in a town with a central retail area between 5,000 sqm and 10,000 sqm. Results of the Odds ratios are in relation to the base case.

The Pseudo R² of all of the product group models are relatively low, varying between 0.03 and 0.06. These relatively low R²'s are not uncommon in this line of literature (Weltevreden, 2007; Weltevreden & Rietbergen, 2006; Farag et al., 2006). The low explanatory power of these models can be explained by the high degree of uncertainty in the purchasing process, personal customer behavior and the specific questions asked in the KSO2016 survey. The purchasing decision is subject to a high degree of personal preferences, that cannot be fully captured by the demographic variables in the model. For instance, the purchasing decision is highly depending on the experiences at different shopping locations by the consumer, including online. These experiences create a certain expectation about the purchasing location (either a physical location or online) and strengthens the preferences of the specific consumers (Frambach et al, 2007). Therefore, experience in shopping physically and online has a significant effect on the purchasing decision (Weltevreden & Rietbergen, 2007). In addition, the choice to either purchase a product online or physically is subject to personal preferences concerning the attractiveness of the area and the preference of the purchasing method. The average ratings do probably not fully capture these preferences (Weltevreden & Rietbergen, 2007). Furthermore, the survey asks the respondent to state the last location where the respondent has purchased a product within a product group. However, from the survey results it is unclear which product was purchased and details concerning the price and information on whether this product was purchased in combination with other products is not available. Moreover, the setting of the shopping trip, such as a quick errant or fun shopping, and the motive of the respondent are not given. Nevertheless, all of the product group models are able to, to a limited extent, explain the determinants of the purchasing decision. The test of the complete model against a model containing only a constant is statistically significant for all product group models.

Table 5: Model estimation output per product group

Product group	Electronics		Electronics		Apparel		Apparel	
Model #	(1)		(2)		(1)		(2)	
Variables	Odds Ratio	Std. Err.	Odds Ratio	Std. Err.	Odds Ratio	Std. Err.	Odds Ratio	Std. Err.
Central area size								
<5000 sqm	0.870***	(0.032)	0,936	(0.053)	1.087**	(0.043)	0,925	(0.055)
10,000sqm - 20,000 sqm	0,962	(0.031)	0,973	(0.037)	1.065*	(0.038)	1,008	(0.042)
20,000 sqm - 40,000 sqm	0,955	(0.028)	0,938	(0.029)	1,000	(0.032)	0,990	(0.034)
40,000 sqm - 60,000 sqm	1.063*	(0.038)	0,986	(0.040)	1,023	(0.040)	1,010	(0.047)
>60,000 sqm	0,965	(0.032)	0.850***	(0.039)	0.903***	(0.033)	0.876***	(0.044)
Rating Vitality			0.964***	(0.009)			0.975**	(0.011)
Rating Accessibility			1.013**	(0.006)			0.984***	(0.006)
Amsterdam	1.259**	(0.107)	1,033	(0.098)	1.210**	(0.112)	1,159	(0.121)
Utrecht	1,100	(0.065)	1,030	(0.065)	1.130*	(0.072)	1.125*	(0.079)
Rotterdam	1,026	(0.066)	0,991	(0.067)	1,116	(0.076)	1.148*	(0.085)
The Hague	1.125**	(0.058)	1,051	(0.060)	1.201***	(0.067)	1.202***	(0.074)
Deviation from mean allocation	1,356	(0.652)	5.056**	(3.449)	0,541	(0.283)	0,326	(0.244)
Distance	0,992	(0.008)	0,985	(0.009)	0.934*	(0.009)	0.981*	(0.010)
% F&B								
40% - 50%	1,009	(0.025)	1,008	(0.028)	1,041	(0.028)	0,977	(0.030)
>50%	1.126***	(0.035)	1.076**	(0.037)	1.104***	(0.035)	1,051	(0.042)
Age group								
< 25	2.296***	(0.135)	2.324***	(0.146)	1.700***	(0.105)	1.734***	(0.116)
25 - 34	2.270***	(0.071)	2.308***	(0.078)	2.253***	(0.075)	2.219***	(0.080)
35 - 44	2.081***	(0.062)	2.089***	(0.068)	2.132***	(0.068)	2.053***	(0.073)
45 - 54	1.422***	(0.039)	1.409***	(0.041)	1.382***	(0.041)	1.366***	(0.045)
65 -74	0.652***	(0.018)	0.646***	(0.019)	0.699***	(0.022)	0.690***	(0.024)
>75	0.381***	(0.018)	0.372***	(0.020)	0.616***	(0.030)	0.601***	(0.033)
Wage level								
< €950	1,039	(0.054)	1,041	(0.058)	0.782***	(0.043)	0.787***	(0.047)
€950 - €1,300	1,032	(0.038)	1,060	(0.042)	0,955	(0.037)	0.930*	(0.040)
€1,300 - €1,900	1,002	(0.027)	1,008	(0.030)	0,990	(0.028)	0,999	(0.032)
> €3,150	1.154***	(0.025)	1.146***	(0.027)	1.107***	(0.026)	1.102***	(0.029)
Household composition								
Single-person	1.077***	(0.027)	1.073***	(0.029)	1.284***	(0.035)	1.274***	(0.038)
Single-parent	1.159***	(0.051)	1.060***	(0.055)	1.167***	(0.055)	1.211***	(0.062)
Family with kids	1.202***	(0.029)	1.008***	(0.032)	1.151***	(0.030)	1.157***	(0.034)
Living with Parents/Legal guardian	1,135	(0.090)	1,146	(0.089)	1.412***	(0.108)	1.267***	(0.109)
Level of Education								
Low	0.889***	(0.023)	0.896**	(0.025)	1,008	(0.027)	1,006	(0.031)
Bachelor	1.127***	(0.025)	1.134**	(0.028)	0,984	(0.024)	0,994	(0.027)
Doctoral	1.294***	(0.035)	1.280**	(0.037)	0.925***	(0.028)	0.939*	(0.031)
Gender								
Male	1.271***	(0.023)	1.257***	(0.025)	0.904***	(0.018)	0.924***	(0.020)
Additional Supply	0.970***	(0.008)	0,990	(0.012)	0.982**	(0.008)	0.951***	(0.020)
Constant	0.406***	(0.048)	0.338***	(0.054)	0.2764***	(0.034)	0.418***	(0.076)
Log likelihood =	-39322,211		-32910		-35285		-29095	
N =	67.273		56,368		70,885		59.154	
LR chi² =	4822,530		4105,95		2828		2256	
Prob > chi² =	0,000		0.0000		0,0000		0,0000	
Pseudo R² =	0,058		0,059		0,039		0,0373	

Dependent variable: Online, a binary variable on whether a consumer purchased a product in the product group online (Y=1) or physically (Y=0).

The standard errors are given in the parentheses behind the corresponding Odds ratios.

* significant at a 10% level, **significant at a 5% level, *** significant at a 1% level.

The base case is a female consumer, between the ages of 55 -65, with aTwo-person household composition, a household income between €1,900 and €3,150 per month, educated at a 'Mid'-level (finished highschool or equivalent), living in a town with a central retail area between 5,000 sqm and 10,000 sqm.

Table 5 continued: Model estimation output per product group

Product group	Media		Media		Sport		Sport	
Model #	(1)		(2)		(1)		(2)	
Variables	Odds Ratio	Std. Err.	Odds Ratio	Std. Err.	Odds Ratio	Std. Err.	Odds Ratio	Std. Err.
Central area size								
<5000 sqm	1.389***	(0.042)	1,003	(0.056)	1.103**	(0.055)	1,074	(0.082)
10,000sqm - 20,000 sqm	1.080**	(0.035)	1.110***	(0.042)	1,013	(0.045)	0,968	(0.050)
20,000 sqm - 40,000 sqm	0,980	(0.029)	0,998	(0.031)	0,984	(0.040)	0,998	(0.042)
40,000 sqm - 60,000 sqm	1.203***	(0.044)	1.208***	(0.050)	1,006	(0.051)	1,044	(0.059)
>60,000 sqm	0.943*	(0.032)	0,971	(0.045)	0.917*	(0.042)	0,946	(0.060)
Rating Vitality			1,008	(0.009)			0,988	(0.013)
Rating Accessibility			0.981***	(0.006)			0,991	(0.008)
Amsterdam	1.337***	(0.114)	1.341***	(0.127)	1.333**	(0.155)	1.401**	(0.182)
Utrecht	1,089	(0.008)	1,100	(0.070)	1.168**	(0.091)	1.210**	(0.103)
Rotterdam	1.152**	(0.075)	1,109	(0.077)	1,050	(0.094)	1,103	(0.106)
The Hague	1.127**	(0.059)	1.119*	(0.065)	1.498***	(0.103)	1,565***	(0.121)
Deviation from mean allocation	1,142	(0.025)	1,099	(0.765)	3.342*	(2.191)	1,772	(1.688)
Distance	0,999	(0.031)	0,999	(0.009)	0,996	(0.012)	0,991	(0.013)
% F&B								
40% - 50%	1,021	(0.025)	1,009	(0.028)	1.104***	(0.036)	1.035***	(0.039)
>50%	1.105***	(0.033)	1.117***	(0.039)	1.071*	(0.044)	0,986	(0.048)
Age group								
< 25	1.471***	(0.089)	1.521***	(0.099)	3.086***	(0.227)	3.228***	(0.255)
25 - 34	1.771***	(0.056)	1.749***	(0.060)	2.744***	(0.111)	2.759***	(0.122)
35 - 44	1.660***	(0.050)	1.630***	(0.054)	2.128***	(0.085)	2.164***	(0.095)
45 - 54	1.409***	(0.038)	1.404**	(0.041)	1.533***	(0.058)	1.550***	(0.065)
65 -74	0.522***	(0.014)	0.521***	(0.015)	0.591***	(0.026)	0.599***	(0.029)
>75	0.283***	(0.014)	0.279***	(0.015)	0.331***	(0.033)	0.337***	(0.037)
Wage level								
< €950	0.704***	(0.038)	0.691***	(0.040)	0.590***	(0.042)	0.593***	(0.051)
€950 - €1,300	0.841***	(0.032)	0.841***	(0.035)	0.866***	(0.065)	0.878**	(0.052)
€1,300 - €1,900	0.895***	(0.025)	0.913***	(0.028)	0.928*	(0.035)	0,938	(0.040)
> €3,150	1.177***	(0.025)	1.188***	(0.028)	1.111***	(0.095)	1.109***	(0.035)
Household composition								
Single-person	1.260***	(0.031)	1.242***	(0.034)	1.170***	(0.042)	1.140***	(0.045)
Single-parent	1,008	(0.045)	0,999	(0.049)	1,109	(0.065)	1,055	(0.069)
Family with kids	1,015	(0.025)	1.022***	(0.028)	1.074**	(0.035)	1.069*	(0.038)
Living with Parents/Legal guardian	1.354***	(0.114)	1,265	(0.113)	0,996	(0.095)	0,897	(0.096)
Level of Education								
Low	0.890***	(0.024)	0.884***	(0.026)	0.908**	(0.035)	0.892***	(0.038)
Bachelor	1,009	(0.023)	1,004	(0.025)	1,008	(0.030)	0,998	(0.033)
Doctoral	1,002	(0.027)	0,978	(0.029)	0,972	(0.035)	0,974	(0.038)
TestGender								
Male	1.297***	(0.024)	1.245***	(0.026)	1.206***	(0.026)	1.178***	(0.032)
Additional Supply	0.982**	(0.008)	0,993	(0.013)	0.982*	(0.011)	0.959**	(0.017)
Constant	0.548***	(0.065)	0.492***	(0.079)	0.147***	(0.023)	0.210***	(0.047)

Log likelihood	=	-37.888	-31.566	-22869	-18959
N	=	59.781	50.265	52.065	43.661
LR chi²	=	4096,06	3410,71	2446,46	2018,69
Prob > chi²	=	0,000	0,000	0,000	0,000
Pseudo R²	=	0,052	0,051	0,051	0,051

Dependent variable: Online, a binary variable on whether a consumer purchased a product in the product group online (Y=1) or physically (Y=0).

The standard errors are given in the parentheses behind the corresponding Odds ratios.

* significant at a 10% level, **significant at a 5% level, *** significant at a 1% level.

The base case is a female consumer, between the ages of 55 -65, with aTwo-person household composition, a household income between €1,900 and €3,150 per month, educated at a 'Mid'-level (finished highschool or equivalent), living in a town with a central retail area between 5,000 sqm and 10,000 sqm.

5.2 Results concerning the central retail area

For a better division of the categories in terms of central retail area size, the categories between 0 sqm and 10,000 sqm and between 10,000 sqm and 40,000 sqm were split at 5,000 sqm and 20,000 sqm, respectively. Most of the coefficients are statistically insignificant, which becomes even more frequent when the ratings for the area are introduced. The results of the separate models indicate that the specific sizes of the central retail areas specifically for those product groups do not significantly influence the likelihood that a consumer will purchase a product either physically or online. However, several product group models do indicate a significant decrease in the likelihood that a consumer will shop online, specifically for that product group, when the size of the central retail is above 60,000 sqm. For both the Apparel and Electronics group the odds that a consumer, who lives near a center that exceeds 60,000 sqm, decreases by 12% and 15% respectively, compared to a consumer who lives near a center between 5,000 sqm and 10,000 sqm. Moreover, the odds that a local consumer will purchase a product online increases by 8.7%, for the Apparel group, to almost 40%, for the Media group, when the size of a local central retail area drops under 5,000 sqm. These findings are in line with the hypotheses stated by Farag and Weltevreden (Farag et al., 2006). For consumers living close to a central retail area between 10,000 sqm and 20,000 sqm and between 40,000 sqm and 60,000 sqm the odds of purchasing Media products online increases by approximately 11% and 21%, respectively, compared to the base case. These products are often argued to be more suitable for online shopping, due to the degree of homogeneity of the products (Weltevreden & Rietbergen, 2006). This makes the products in this group less suitable for fun shopping and more suitable for online searching behavior, which allows the consumer to compare prices more easily. Therefore, to a certain extent, the second hypothesis which states that the size of the retail area significantly influences the shopping decision of local consumers, cannot be rejected.

For several product groups, consumers living in the four largest cities in the Netherlands are expected to be more likely to purchase a product online. The probability that a consumer will purchase a product online when living in The Hague increases by approximately 1.5 times compared to a consumer who does not live in one of the four largest cities in the country in the Sports product group. Several cities show significant odds ratios above 1 for specific product group models. These findings are in line with the expectation that consumers who relate more to a urban type of consumer are more likely to engage in the process of online searching and/or online shopping (Weltevreden & Rietbergen, 2006).

For the Apparel and Media product groups, the perception of the accessibility of a central retail area, expressed as the predicted scores by the PCA analysis for the accessibility variables,

significantly influences the purchasing decision. A higher rating on the accessibility is considered to decrease the odds that a consumer will purchase a product online by approximately 1.6% and 1.9% respectively. These results are in line with the previous literature (Farag et al., 2006; Weltevreden & Rietbergen, 2006). The sign for the Electronics product group shows a counterintuitive, positive relationship in the likelihood of ordering online, when the rating for accessibility increases. This could be explained by an additional effect specific for this product group. The purchase prices of products in this group are usually relatively high and/or the products are large in terms of size. This might influence the preferences of consumers in choosing where to purchase. For this category other retail locations which are, by definition, better accessible in terms of parking and generally offer larger shop sizes are more suitable for part of the products in this product group. This might also be expressed by the significant effect of the additional supply in the market for this product group in the first model. Moreover, online shopping eliminates the travel efforts, since products can be delivered to the consumer's home address. The rating for the vitality is significant for the Electronics and Apparel product groups. A higher rating decreases the odds for a base case consumer to buy a product in these groups online by approximately 2.5% and 3.6% respectively. The insignificance specifically in the other models are probably because average ratings are taken as a proxy, instead of consumer specific ratings. There is arguably a larger variance among individuals in these ratings, compared to the accessibility. Moreover, these products require less comparison (Media) or are either more specialized and offered by specialized shops of which less shops are located in the central retail area. Therefore, the third hypothesis which argues that the rating of the area by local consumers significantly influences the likelihood that a local consumer will shop cannot be rejected for several specific product groups.

The variable for distance between a consumer's residential location and the central retail area is either insignificant, or the Odds ratio has the wrong sign in all specific models. These findings are contrary to the existing literature, which indicated that in other countries an increase in distance from the retail area has a negative effect on the willingness to travel towards a retail area. The difficulties that occur with this variable are possibly caused by the high urban density in the research area. The maximum distance between a consumer's residential location and a central research area is less than ten kilometers. Arguably, a minimum distance between a consumer and a retail area is needed before the effect of distance becomes clear. Moreover, several different central areas that offer sufficient facilities, shops and amenities are located within a relatively small distance of each consumer, indicating a high variety in options a consumer has in choosing where to shop. This reduces the dependency of consumers on either one retail location or the online retail environment. In addition, the willingness to travel increases when consumers consider shopping as a leisure activity.

The variable for the deviation in the sector allocation of shops has a significant effect on the purchasing decision for the Electronics product group. When the deviation from the mean allocation of sectors increases by one unit, indicating a more skewed allocation, the odds that a consumer will purchase online increases by 5 times. This relates to the possible reduction in shops that offer products belonging to the Electronics sector. Figures on the allocation of sectors in a central retail area indicate that when a central area increases in size, the percentage of shops in the Apparel industry increases and the percentage of shops in the Home & Living sector and Services sector decreases.[27] As mentioned, the Electronics sector is more reliant on preparation and comparing behavior of consumers. The ability to achieve this is reduced when the supply of shops within an area is reduced or spread out over a larger area. This makes purchasing online a relevant alternative and could possibly explain the significant effect in only this specific sector. Therefore, the fourth hypothesis, which states that the allocation of shops in the area has an effect on the shopping decision cannot be rejected for the specific case of the Electronics product group.

A high percentage of F&B amenities in a central retail area is often considered as a positive development.[28] The ideal mix of shopping and leisure could increase the attractiveness of a retail area (Gillespie et al., 2001). The ideal mix is subject to a certain ratio between the two activities. The results of the models of several product groups indicate a significant positive effect on the likelihood that a consumer will purchase a product online when the percentage of F&B in a central retail area is too large and exceeds 50%. This is possibly explained by a lower degree of supply in a central retail area for shopping purposes. The central retail area is more of a leisure location in these cases, instead of a retail location, reducing its attractiveness as a shopping location. The odds that a consumer will purchase a product online increases by for instance 11.7% for the Media product group, when the ratio of F&B amenities exceeds 50% of all units in the area. Therefore, hypothesis five is rejected, as it was argued that F&B amenities would increase the probability that a local consumer will consume physically. On the contrary, when the ratio is too high, the effect is the opposite from this expectation.

5.3 Results concerning the control variables

The estimated effects of different age groups in determining whether a consumer is more likely to consume online are in line with the existing literature by Panda & Swar (2013). The different age groups all have a significant effect on the likelihood that a consumer will purchase online.

[27] Locatus data 2017
[28] Colliers International (2016):'Transitie van de Nederlandse winkelstructuur'

Different age groups have different preferences in determining whether to shop physically or online. Younger consumers, starting from an age of 25, are more likely to choose online and this effect decreases with age shifting to a higher probability that a consumer will choose to purchase physically. These results are in line with the findings by Farag et al. (2006).

The estimation of the models in individually indicates in general that a higher income increases the likelihood that a consumer will purchase online, which is in line with the findings of Chau & Hu (2002), Farag et al. (2006) and Hubona & Kennick (1996). The results contradict with the expectation that the lowest incomes would also consume online more frequently, due to the lower prices offered online.

Different households compositions lead to different preferences when choosing between purchasing either online or physically. In general, compared to a consumer that lives in a household that consists of two persons, a person living either alone (one-person household), living with his/her parents/legal guardian or has a family of his/her own is more likely to purchase a product online. The odds of purchasing a product online increases for instance with 27.4%, 15.7% and 26.7% for these compositions respectively for the Apparel sector, compared to a consumer living in a two-person household. This could be explained by the hedonic and utilitarian principle (Shen et al., 2016). This might indicate that a couple is more likely to initiate a shopping trip in the form of a leisure activity, while families cope with a lack of time and single-person households and consumers still living with their parents might not experience as much of a boost in their utility when initiating a shopping trip.

In general, the effect of the level of education is ambiguous. However, several models individually do show that the odds that a consumer will purchase a product online decreases by for instance approximately 11.4% and 11.6% for the Electronics and Media sector respectively when a consumer has a low level of education, compared to a high school level of education. This can possibly be explained by the knowledge a consumer holds and the ability to process a possibly complex retail environment such as the online retail environment. However, this effect is only significant for specific educational categories.

In addition also gender has a significant effect on the purchasing decision. For all product groups models individually in the Home & Living sector male consumers are more likely to purchase a product online. This is in line with the existing literature on consumption behavior among different gender (Farag et al. 2006, Ditmar et al., 2004). However, the probability that a consumer will purchase a product in the apparel group decreases when the respondent is a male consumer. The odds that a male consumer will purchase a product online relatively to a female consumer decreases by approximately 7.6%. This indicates that male consumers, compared to female consumers, might be less confident in shopping for clothes and luxury

items online in comparison to more straightforward products such as in the Home & Living sector. This is in line with the statements by Mattila et al. (2003) and Dittmar et al. (2004). In general the results concerning the effects of demographical characteristics of local consumers on the shopping decision are in line with the existing literature. In addition, specific income levels and household compositions do have an effect on the tradeoff between shopping physically or online for specific product groups. Therefore, the first hypothesis, which stated that demographical consumer specific characteristics have an effect on the likelihood that a consumer will purchase a product online, cannot be rejected.

6. Conclusion, discussion & limitations

6.1 Conclusion

The purpose of this paper was to analyze the effect of area specific, consumer specific and geographical variables on the purchasing decisions by local consumers. The shopping behavior of consumers has been subject to several changes in the previous decades, for instance with the introduction and manifestation of the online retail environment as source of information and purchasing location. Moreover, preferences and intentions of consumers are argued to have changed with this transition. Therefore, more research in general and more extensive recent research is needed in order to analyze the possible changes in consumer behavior.

In order for retailers or municipalities to correctly determine their strategy, it is important to determine what the effects are of specific (central) retail area characteristics and the perception of the area by consumers. Therefore, this paper focuses on these possible effects and controlled for consumer specific and geographical characteristics. The analysis focuses on the purchasing decision by consumers choosing either to shop at a physical store or online. In order to analyze this binary dependent variable, the paper uses the results of the KSO2016 research conducted by I&O Research. The paper focuses on the central retail area near a consumer's residential location, since this is the most important and most frequently used retail area in many Western economies, and more specifically in The Netherlands.

Firstly, the results on the demographic variables are in line with the expectations and provide an answer to the first sub-question of the paper. The results indicate that the demographical characteristics to a significant extent influences the shopping decision of consumers. Specific consumers are more inclined to shop either physically or online in all product group models,

ceteris paribus. In addition, they offer some valuable insight in terms of differences between product groups. For instance, while male consumers are in general more likely to shop online, this is not the case for products in the apparel industry. Moreover, several households compositions were not yet proven to have a significant impact on the likelihood that a consumer will purchase online. A two-person household is the least likely to purchase a product online, while the probability of an online purchase differs in magnitude between different product groups for the other household compositions. The results on the demographic variables give valuable insights to, for instance retailers and municipalities, because the actual figures of these variables are (publicly) available or known for the local case.

The second and third sub-question are also answered by the empirical analysis. The results indicate that the size of the area, which resembles the potential of the area, only has a significant effect on the purchasing decision for specific sizes. This indicates that the likelihood that a consumer will purchase a product online or offline is not significantly affected by this variable in general. This holds for nearly all different size categories, with the exception of the category of which the central retail area exceeds 60,000 sqm and the smallest area size under 5,000 sqm. The first group offers a high supply and large variety of shops. The odds of a consumer purchasing online are decreased due to these characteristics. Moreover, when excluding the rating variables, the odds of a consumer purchasing online the odds increase for the lowest category in terms of size for specific product models. This is partly resembled by the efficiency theory used by Farag et al. (2006).

The results on the rating variables show the importance of the perception of accessibility by consumers. When the threshold between leaving the house and initiate a shopping trip is taken away as far as possible, this decreases the likelihood that a consumer will choose for the online alternative. These findings give valuable insights to for instance municipalities who are able to improve the accessibility of a retail area. In addition the vitality of a central retail area has a negative effect on the likelihood that a local consumer will purchase a product online for the Electronics and Apparel product groups. Therefore, the size and perception of a central retail area do influence the shopping decision of local consumers in specific cases in determining whether to shop physically or online.

Finally, the results indicate that the allocation variables have a significant effect on specific product groups on the likelihood of purchasing a product online, which provides an answer to the third sub-question. By implementing strategies that offer a better mix in terms of shops, policy makers could limit the increased probability that a consumer will purchase a product in the Electronics product group. While F&B amenities are argued to increase the vitality of the area, a too high degree of F&B amenities in a central retail area reduces the supply of shops

too far. This dominance of F&B in a central retail area, above 50% of all units, increases the likelihood that a local consumer will purchase a product online. In conclusion, the characteristics of the local central retail area, although to a limited extent and for specific product groups, do have an effect on the likelihood that a consumer will switch to either online or physical shopping behavior.

6.2 Discussion, limitations and suggestions for further research

There are several characteristics of the Dutch case study that need to be acknowledged and might limit the transferability of the results to other markets. The Dutch retail market is structured differently compared to other Western and developed countries. The market is characterized by a larger number of stores per capita and stores are compared to other markets relatively small in terms of size. Moreover, the influence of the city center, being the most important retail location in the country, and strict restrictive planning regulations have limited the growth of large-scale shopping centers in the country. As mentioned in the previous sections the distances between villages and cities in the Randstad area are small and the degree of urbanization in the area is relatively high.

From this paper it follows that personal preferences have an important effect on the personal actions. Therefore, it is difficult to model for instance why a consumer behaves in certain ways. Specifically for the dataset used in this analyses the ratings are not on a personal level, but on a aggregated average level. This limits the variable as it expresses a general rating on every aspect, where it would be more useful to interpret the personal ratings that come from personal experiences. Therefore, future research should incorporate data measured on a personal level and including preferences of consumers individually. In general the dataset is not a perfect fit for the model used in this paper. However, the analyses does offer useful insights that could contribute to refueling the discussion on the effects of online shopping and the difficulties physical retailers are facing now or in the future.

The literature on E-commerce is not always supported by quantitative analysis or sufficient data and is mostly survey based (Zhang et al. 2016). According to Weltevreden (2007) time based research is necessary to examine the magnitude of the effects of E-commerce and to overcome the subjective nature of the current literature. Therefore, more quantitative research examining the variables over time is needed to improve the existing literature. The non-existence of this type of literature is explained by the lack of transparency by possible data sources. Quantitative data concerning for instance transaction values are subject to privacy limitations, for both the consumers and the registration companies, and are not widely

accessible. This explains why survey-based research is mostly dominant in this line of literature.

An improvement to this paper would be the introduction of more consumer specific preferences and characteristics concerning the consumer's behavior. Moreover, the used dataset did not offer sufficient data on the online retail environment that has been used in previous literature. Therefore, a complete model that includes next to the physical retail area characteristics, also the online retail environment characteristics and consumer experience, has yet to be introduced.

7. References

Abdi, H. and Williams, L. J. (2010). 'Principal component analysis'. *WIREs Comp Stat, 2: 433–459. doi:10.1002/wics.101*

Akerlof, G.A. (1970). 'The Market for "Lemons": Quality Uncertainty and the Market Mechanism.' *The Quarterly Journal of Economics Vol. 84, No. 3 (Aug., 1970), pp. 488-500.*

Alba, J., Lynch, J., Weitz, B., Janiszewski, C., Lutz, R., Sawyer, A. and Wood, S. (1997). 'Interactive home shopping: Consumer, retailer, and manufacturer incentives to participate in electronic marketplaces.' *Journal of Marketing, 61(3), 38–53.*

Avery, J., Steenburgh, T. J., Deighton, J. and Caravella, M. (2012). 'Adding bricks to clicks: Predicting the patterns of cross-channel elasticities over time.' *Journal of Marketing, 76(3), 96–111.*

Bae, S. and Lee, T. (2011). 'Gender differences in consumers' perception of online consumer reviews'. *Electron Commer Res (2011) 11: 201–214.*

Bell, D.R., Ho, T. and Tang, S. (1998). 'Determining Where to Shop: Fixed and Variable Costs of Shopping.' *Journal of Marketing Research, Vol. 35, No. 3 (Aug., 1998), pp. 352-369.*

Bainbridge, R.E., MAI (2014). 'The Stormy Future of Retail Real Estate' *The Appraisal Journal, Spring 2014, pp. 155-160.*

Benjamin, J.D., Jud, G.D. and Winkler, D.T. (1994).'The Supply Adjustment Process in Retail Space Markets'. *The Journal of Real Estate Finance and Economics, vol. 10, pp 161-168.*

Bhat, C.R., Sivakumar and A., Axhausen, K.W. (2003).'An analysis of the impact of information and communication technologies on non-maintenance shopping activities.' *Transportation Research B 37 (10), pp. 857–881.*

Bouman, W. (2012). 'gevolgen van retail-revolutie voor winkelvastgoed.' *Tijdschrift voor Huurrecht bedrijfsruimte, juni 2012, nr. 3, pp. 148-150.*

Burke, R.R. (1998). 'Do you see what I see? The future of virtual shopping.' *Journal of the Academy of Marketing Science, 25,352-360.*

Chau, P.Y.K. and Hu, P.J. (2002). 'Examining a Model of Information Technology Acceptance by Individual Professionals: An Exploratory Study.' *Journal of Management Information Systems, Vol. 18, No. 4, pp. 191-229.*

Cheung, C., Xiao, B. and Liu, I. (2014). 'Do Actions Speak Louder Than Voices? The Signaling Role of Social Information Cues in Influencing Consumer Purchase Decisions.' *Decisions Support Systems, Vol. 65, pp. 50-58.*

Chiang, K. and Dholakia, R.R. (2003). 'Factors Driving Consumer Intention to Shop Online: An Empirical Investigation.' *Journal of Consumer Psychology, 13(1&2), 177-183.*

Corpuz, G. and Peachman, J., 2003. 'Measuring the impacts of Internet usage on travel behavior in the Sidney Household Travel Survey.' *Presented at the 26th Australian Transport Research Forum Conference, 1–3 October 2003, Wellington, New Zealand. papers/2003/18-Corpuz.pdfS.*

Cubukcu,, K.M. (2001). 'Factors affecting shopping trip generation rates in metropolitan areas.' *Studies in Regional and Urban Planning 9, 51–67.*

Darley, W.K. and Lim, J. (1999). 'Effects of store image and attitude toward secondhand stores on shopping frequency and distance traveled.' *International Journal of Retail & Distribution Management, Vol. 27 Issue: 8, pp.311-318.*

Dittmar, H., Long, K. and Meek, R. (2004). 'Buying on the Internet: Gender Differences in On-line and Conventional Buying Motivations.' *Sex Roles, Vol. 50, Nos. 5/6, March 2004.*

Dixon, T. and Martson, A. (2010).'U.K. Retail Real Estate and the Effects of Online Shopping.' *Journal of Urban Technology, 9:3, 19-47.*

Dixon, T., Thompson, B., McAllister, P., Marston, A.D., Snow, J. (2005). 'Real Estate and the New Economy; The Impact of Information and Communications Technology.' *Blackwell Publishers, Oxford.*

Eppli, M.J. and Laposa, S.P. (1997). 'A Descriptive Analysis of the Retail Real Estate Markets at the Metropolitan Level'. *Journal of Real Estate Research, Vol. 14, No. 3, pp. 321-338.*

Esser, K. and Kurte, J. (2005), 'B2C-Handel: Qualitative und Quantitative Analyse des Konsumenten- und Anbieterverhaltens und deren Verkehrliche Auswirkungen in

Ballungsräumen (B2C E-commerce: A Qualitative and Quantitative Analysis of the Consumer and Supplier Relationships and their Implications for Urban Transport).' *KE-Consult, Cologne.*

Fang, J., Wen, C., George, B. and Prybutok, V.R. (2016). 'Consumer Heterogeneity, Perceived Value, and Repurchase Decision-Making in Online Shopping.' *Journal of Electronic Commerce Research, VOL 17, NO 2, 2016.*

Farag, S., Weltevreden, J., van Rietbergen, T. and & Dijst, M. (2006). 'E-shopping in the Netherlands: does geography matter?' *Environment and Planning B: Urban Analytics and City Science, Vol 33, Issue 1, pp. 59 - 74*

Farag, S., Schwanen, T., Dijst, M. and 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 A 41 (2), 125–141.*

Ferrell, C.E. (2005). 'Home-based teleshopping and shopping travel: where do we find the time?' *CD Proceedings of the 84th Annual Meeting of the Transportation Research Board, Transportation Research Board, Washington, DC.*

Fox, E.J., Montgomery, A.L. and Lodish, L.M. (2004). 'Consumer Shopping and Spending across Retail Formats'. *The Journal of Business, Vol. 77, No. S2 (April 2004), pp. S25-S60.*

Frambach, R. T., Krishnan, T. V. and Roest, H. C. (2007). 'The impact of consumer internet experience in channel preference and usage intentions across the different stages of the buying process.' *Journal of interactive marketing, 21(02), 26-41.*

Ganesam-Lim, C., Russel-Bennet, R. and Dagger, T. (2008). 'The Impact of Service Contact Type and Demographic Characteristics on Service Quality Perceptions.' *Journal of Services Marketing, Vol. 22, No.7, pp. 550-561.*

Gillespie, A., Marvin, S. and Green, N. (2001). 'Bricks versus clicks: planning for the digital economy.' *In: Wilsdon, J. (Ed.), Digital Futures: Living in a Dot-Com World. Earthscan, London, pp. 200–218.*

Goldmanis, M., Hortac̦su, A., Syverson and C. Emre, Ö. (2009). 'E-commerce and the market structure of retail industries'. *The Economic Journal, June, pp. 651-682.*

Goswami, S. and Khan, S. (2015). 'Impact of Consumer Decision-making Styles on Online Apparel Consumption in India.' *Vision 19(4) 303–311, 2015.*

Guy, C.M. (1994). 'The Retail Development Process'. *Routledge, London.*

Häubl, G. andTrifts, V. (2000). 'Consumer Decision Makin in Online Shopping Environments: The Effects of Interactive Decision Aids.' *Marketing Science (2000), Vol. 19, No. 1, Winter 2000, pp. 4-21.*

Hawkins, D.I., Best, R.J. and Coney, K.A. (1998). 'Consumer Behavior: Building Marketing Strategy.' *Irwin/McGraw-Hill, Boston, MA.*

Hernandez, T., Gomez-Insausti and R., Biasiotto, M. (2001). 'Non-store retailing and shopping centre vitality.' *Journal of Shopping Centre Research 8 (2), 58–81.*

Hernández, B., Jiménez and J., Martín, M.J. (2011). 'Age, gender and income: do they really moderate online shopping behaviour?' *Online Information Review, Vol. 35 Issue: 1, pp.113-133.*

Homburg, C. and Giering, A. (2001). 'Personal Characteristics as Moderators of the Relationship between Customer Satisfaction and Loyalty: An Empirical Analysis.' Psychology & Marketing, Vol. 18, No. 1, pp. 43-66.

Hubona, G.S. and Kennick, E. (1996). 'The influence of external variables on information technology usage behavior.' *Proceedings of HICSS-29: 29th Hawaii International Conference on System Sciences, Wailea, HI, 1996, pp. 166-175 vol.4.*

Krizek, K.J., Li, Y. and Handy, S.L. (2005). 'ICT as a substitute for non-work travel: a direct examination.' *CD Proceedings of the 84th Annual Meeting of the Transportation Research Board, Transportation Research Board, Washington, DC.*

Lee, B., Ang, L. and Dubelaar, C. (2005). 'Lemons on the Web: A signaling approach to the problem of trust in Internet commerce.' *Journal of Economic Psychology 26 (2005) 607–623*

Lee, M. and Johnson, K.K.P. (2002). 'Exploring differences between internet apparel purchasers, browsers and nonpurchasers.' *Journal of Fashion Marketing and Management, 6(2), 146–157.*

Long, J. S. and Freese, J. (2004). 'Regression Models for Categorical Dependent Variables Using Stata'. *1st Edition, College Station, Texas, Stata Press.*

Loudon, D.L. and Della Bitta, A.J. (1993). 'Consumer Behavior: Concepts and Applications.' *McGraw-Hill, New York, NY.*

Lusch, R.F. and Lusch, V.N. (1987). 'Principles of Marketing.' *Kent Publishing Company, Boston, MA.*

Mägi, A.W. (2003). 'Share of wallet in retailing: the effects of customer satisfaction, loyalty cards and shopper characteristics.' *Journal of Retailing 79 (2003) 97–106.*

Mattila, A.S., Grandey, A.A. and Fisk, G.M. (2003). 'The Interplay of Gender and Affective Tone in Service Encounter Satisfaction.' *Journal of Service Research, Volume 6, No. 2, November 2003 136-143.*

McQuiston, D.H. (1989). 'Novelty, Complexity, and Importance as Causal Determinants of Industrial Buyer Behavior.' *Journal of Marketing Vol. 53, No. 2 (Apr., 1989), pp. 66-79.*

Mokhtarian, P.L. (2004). 'A conceptual analysis of the transportation impacts of b2c e-commerce.' *Transportation 31 (3), 257–284.*

Moore, D. S., McCabe, G. P. and Craig, B. A. (2014). Introduction to the practice of statistics. *8th Edition, New York: W.H. Freeman and Company.*

Panda, R. and Swar, B.N. (2013). 'Online Shopping: An Exploratory Study to Identify the Determinants of Shopper Buying Behaviour.' *IJBIT / Volume 7 / Issue 1 / October 2013 - March 2014.*

Pauwels, K. and Neslin, S. A. (2011). 'Building with bricks and mortar: The revenue impact of opening physical stores in a multichannel environment.' *Journal of Retailing 91 (2, 2015) 182–197.*

Peterson, R.A., Balasubramanian S.A. and Bronnenberg, BJ (1997). 'Exploring the Implications of the Internet for Consumer Marketing.' *Journal of the Academy of Marketing Science, Vol. 25, No. 4: 329-346, 1997.*

Sautter, P., Hyman, M.R. and Lukošius V. (2004). 'E-Tail Atmospherics: A Critique of the Literature and Model.' *Journal of Electronic Commerce.*

Sharma, P., Chen, I.S.N. and Luk, S.T.K. (2012). 'Gender and Age as Moderators in the Service Evaluation Process.' *Journal of Services Marketing, Vol. 26 Issue: 2, pp.102-114.*

Shen, K.N., Cai and Y., Guo, Z. (2016). 'When Do Online Consumers Shop in an Offline Store: The Moderating Effects of Product Characteristics.' *Journal of Marketing Channels, 23:3, 129-145.*

Steel, W., Daglish, T., Marriott, L., Gemmell and N., & Howell, B. (2013). 'E-Commerce and its effect upon the Retail Industry and Government Revenue.' *New Zealand Institute for the study of competition and regulation (pp. 22591) (http://www. iscr. org. nz/f859)*

Tonn, B.E. and Hemrick, A. (2004). 'Impacts of the use of e-mail and the Internet on personal trip-making behavior.' *Social Science Computer Review 22 (2), 270–280.*

Venkatesan, R., Kumar, V. and Ravishanker, N. (2007). 'Multichannel shopping: Causes and consequences.' *Journal of Marketing, 71(2), 114–132.*

Wallace, D. W., Giese, J. L. and Johnson, J. L. (2004). 'Customer retailer loyalty in the context of multiple channel strategies.' *Journal of Retailing, 80(4), 249–263.*

Walsh, G., Evanschitzky, H. and Wunderlich, M. (2008). 'Identification and analysis of moderator variables; Investigating the customer satisfaction-loyalty link.' *European Journal of Marketing Vol. 42 No. 9/10, 2008 pp. 977-1004.*

Weltevreden, J.W.J. (2007). 'Substitution or Complementarity? How the Internet changes city Centre shopping.' *Journal of Retailing and Consumer Services, 14 , 192-207.*

Weltevreden, J.W.J. and Rietbergen, T. van (2007). 'E-shopping versus city Centre shopping: the role of perceived city Centre attractiveness.' *Journal of Economic & Social Geography (TESG), 98 (1), 68-85*

Weltevreden, J.W.J. (2014). 'The digital challenge for the high street: insights from Europe.' *Evolving High Streets: Resilience & Reinvention, University of Southampton, E.S.R.C.*

Xu, Y. and Paulins, V.A. (2005). 'College students' attitudes toward shopping online for apparel products: Exploring a rural versus an urban campus.' *Journal of Fashion Marketing and Management, 9(4), 420–433.*

Yoon, C., Laurent, G., Fung, H. and Gonzalez, R. (2005). 'Cognition, Persuasion and Decision Making in Older Consumers.' *Marketing Letters, Vol. 16, No. ¾, pp. 429-441.*

Zhang, D., Zhu, P. and Ye, Y. (2016). 'The effects of E-commerce on the demand for commercial real estate.' *Cities, 51, pp.106-120.*

Zhou, Z., Jin, X. and Fang, Y. (2014). 'Moderating role of gender in the relationships between perceived benefits and satisfaction in social virtual world continuance.' *Decision Support Systems 65 (2014) 69–79.*

8. Appendices

Appendix 1: KSO2016 Survey form translation of one of the product groups.

30A At which location did you most recently buy a product in the Media sector (a.o. Books, DVD's, CD's and hobby related products)?

One answer

In <place/postal code>   → Vraag 30C

Outside of <place/postal code>   → Vraag 30B

Online (via internet)   → Vraag 30D

Other (TV/catalogue)   → Next

I never buy products in this category   → Next

(answers include the retail area at which the product was bought)

Ask if answer at 30A was online

30D   At what type of shop did you buy this product?

One answer

Full webshop

Webshop of a chain of stores also located in physical retail areas

Webshop of local entrepreneur
Marktplaats (second hand products)

Other,

I don't know

30E   With what method did the package arrive at your residential location?

One answer

Home delivery

 Pick up at a physical location of the retailer

Pick-up point

Pick up point at local retailers

Other

I don't know

Questions concerning the rating of the physical retail area

Asked if bought at a physical retail location

37   What rating would you give the retail area on the following categories, ranging from 0 to 10. (Does not apply is also an option)

38   Total Rating

Shop variety and supply
Atmosphere
Safety
Cleanliness /maintenance
Parking facilities car

Parking tariffs car
Parking facilities bike
Accessibility and reachability car
Accessibility and reachability public transportation
Accessibility and reachability bike
Supply of F&B
Events

Facilities (such as toilets, Wi-Fi etc.)

A1 Highest form of education (finished)?

Elementary school
First part of high school (low level)
Second part of high school (high level) or equivalent
First year of post high school
Bachelor
Master/Doctored

A2 What is your current household composition?

Single-Person
Single parent
Two-Person
Family (including kids)
With parents/legal guardian(s)
Other

A3   What is the net income of your household?

< € 950
Between € 951 & € 1.300
Between € 1.301 & € 1.900
Between € 1.901 & € 3.150
> € 3.150
No answer

A4   Gender

☒   Male   ☐   Female

A5   Age   _____

A6 Postal code residential location   _____

51

Appendix 2: Descriptive statistics of different product groups

Apparel sector:

Descriptive statistics Apparel sector

Purchase Online	Frequency	Percentage
No	75,749	79%
Yes	20,142	21%
Total	95,891	100%
Central area	**Frequency**	**Percentage**
<10,000 sqm	31,382	33%
Between 10,000 & 40,000 sqm	34,144	36%
Between 40,000 & 60,000 sqm	8,190	9%
> 60,000 sqm	22,174	23%
Total	95,891	100%
Gender distribution	**Frequency**	**Percentage**
Male	42,414	44%
Female	53,339	56%
Unknown	138	0%
Total	95,891	100%
Level of education	**Frequency**	**Percentage**
Primary & Lower Secondary & Upper first phase	23,172	24%
Post Lower Secondary & Upper Secondary second phase	31,251	33%
Bachelor	26,388	28%
Master/Doctoral	13976	15%
Unknown	1,104	1%
Total	95,891	100%
Household composition	**Frequency**	**Percentage**
One-person	21,766	23%
Single parent	4,102	4%
Two-person	41,890	44%
Family with kids	24,579	26%
With parent(s)/legal guardian(s)	1,969	2%
Other	1,560	2%
Unknown	25	0%
Total	95,891	100%
Income distribution	**Frequency**	**Percentage**
< €950	2,591	3%
between € 951 & € 1,300	6,196	6%
between € 1,301 & € 1,900	12,923	13%
Between€ 1,901 & € 3,150	29,017	30%
> € 3,150	25,393	26%
Unknown	19,771	21%
Total	95,891	100%

Descriptive statistics Apparel sector continued

Age distribution	Frequency	Percentage
15-24	3,032	3%
25-34	9,020	9%
35-44	11,937	12%
45-54	18,926	20%
55-64	23,230	24%
65-74	22,572	24%
>74	7,174	7%
Total	95,891	100%

Descriptive statistics Apparel industry continued

Variable	N	Mean	Std. Dev	Min	Max
Rating Other	74,999	7.18	0.57	4.20	8.20
Rating Access	74,999	7.48	0.66	5.70	8.90
Rating Atmosphere	74,999	7.38	0.30	6.20	8.20
Age	95,373	54.68	15.20	15.00	95.00
Additional supply	95,891	93,025.17	170,149.00	0.00	888,143.00
ln(Additional supply)	94,891	9.98	2.34	0.00	1.370
Radius	95,490	4,324.12	2,776.28	1.00	9,840.00
ln(Radius)	95,490	8.05	1.01	0.00	9.19
Travel time (sec)	95,490	1,261.95	524.67	1.00	2,050.00
ln(Travel time)	95,490	6.91	0.94	0.00	7.63
Dev Shop Allocation	89,771	0.0418	0.042	0.01	0.200
F&B	90,638	0.448	0.077	0	0.833

Electronics sector:

Descriptive statistics Household appliances sector

Purchase Online	Frequency	Percentage
No	62,418	69%
Yes	27,677	31%
Total	90,095	100%
Central area	**Frequency**	**Percentage**
<10,000 sqm	29,623	33%
Between 10,000 & 40,000 sqm	32,205	36%
Between 40,000 & 60,000 sqm	7,632	8%
> 60,000 sqm	20,635	23%
Total	90,095	100%
Gender Distribution	**Frequency**	**Percentage**
Male	42,191	47%
Female	47,775	53%
Unknown	129	0%
Total	90,095	100%
Level of education	**Frequency**	**Percentage**
Primary & Lower Secondary & Upper first phase	21,148	23%
Post Lower Secondary & Upper Secondary second phase	29,055	32%
Bachelor	25,245	28%
Master/Doctoral	13,647	15%
Unknown	1,000	1%
Total	90,095	100%
Household composition	**Frequency**	**Percentage**
One-person	20,242	22%
Single parent	3,737	4%
Two-person	39,941	44%
Family with kids	23,218	26%
With parent(s)/legal guardian(s)	1,530	2%
Other	1,405	2%
Unknown	22	0%
Total	90,095	100%
Income distribution	**Frequency**	**Percentage**
< €950	2,250	2%
between € 951 & € 1,300	5,514	6%
between € 1,301 & € 1,900	12,004	13%
Between€ 1,901 & € 3,150	27,834	31%
> € 3,150	24,639	27%
Unknown	17,854	20%
Total	90,095	100%

Descriptive statistics Electronics appliances sector continued

Age distribution	Frequency	Percentage
15-24	2,473	3%
25-34	8,134	9%
35-44	11,166	12%
45-54	17,849	20%
55-64	22,133	25%
65-74	21,620	24%
>74	6,720	7%
Total	90,095	100%

Descriptive statistics Electronics sector continued

Variable	N	Mean	Std. Dev	Min	Max
Rating Other	70.809	6,91	0,94	0,0	7,7
Rating Access	70.809	7,37	0,49	5,5	8,4
Rating Atmosphere	70.809	7,37	0,24	6,3	8,4
Age	89.617	55,01	14,97	15	95
Additional supply	90.095	91.739,89	168.570,00	0,00	888.143,00
ln(Additional supply)	90.095	9,83	2,33	0,00	1,37
Radius	90.041	4.330,75	2.779,33	1,00	9.840,00
ln(Radius)	90.041	8,05	1,01	0,00	9,19
Traveltime	90.041	1.261,17	524,77	1,00	2.050,00
ln(Traveltime)	90.041	6,91	0,94	0,00	7,63
Dev Shop Allocation	84,369	0.0418	0.042	0.01	0.200
F&B	85,172	0.448	0.077	0	0.833

Media sector:

Descriptive statistics Media sector

Purchase Online	Frequency	Percentage
No	49,207	62%
Yes	30,513	38%
Total	79,720	100%
Central area	**Frequency**	**Percentage**
<10,000 sqm	26,013	33%
Between 10,000 & 40,000 sqm	28,607	36%
Between 40,000 & 60,000 sqm	6,515	8%
> 60,000 sqm	18,585	23%
Total	79,720	100%
Gender Distribution	**Frequency**	**Percentage**
Male	36,261	45%
Female	43,357	54%
Unknown	102	0%
Total	79,720	100%
Level of education	**Frequency**	**Percentage**
Primary & Lower Secondary & Upper first phase	16,475	21%
Post Lower Secondary & Upper Secondary second phase	25,431	32%
Bachelor	23,634	30%
Master/Doctoral	13,383	17%
Unknown	797	1%
Total	79,720	100%
Household composition	**Frequency**	**Percentage**
One-person	17,936	22%
Single parent	3,330	4%
Two-person	34,616	43%
Family with kids	21,025	26%
With parent(s)/legal guardian(s)	1,503	2%
Other	1,293	2%
Unknown	17	0%
Total	79,720	100%
Income distribution	**Frequency**	**Percentage**
< €950	1,988	2%
between € 951 & € 1,300	4,604	6%
between € 1,301 & € 1,900	10,017	13%
Between€ 1,901 & € 3,150	24,499	31%
> € 3,150	22,998	29%
Unknown	15,614	20%
Total	79,720	100%

Descriptive statistics Media appliances sector continued

Age distribution	Frequency	Percentage
15-24	2,377	3%
25-34	7,609	11%
35-44	10,289	15%
45-54	16,007	23%
55-64	19,294	28%
65-74	18,462	27%
>74	5,682	8%
Total	69,009	100%

Descriptive statistics Media appliances sector continued

Variable	N	Mean	Std. Dev	Min	Max
Rating Other	62,878	7.38	0.49	5.5	8.4
Rating Access	62,878	7.44	0.63	5.7	8.9
Rating Atmosphere	62,878	7.37	0.32	6.3	8.4
Age	79,305	54.47	15.07	15	95
Additional supply	79,720	92,744.25	169,740.50	0.00	888,143.00
ln(Additional supply)	79,720	9.84	2.33	0.00	1.37
Radius	79,673	4,326.43	2,777.97	1.00	9,840.00
ln(Radius)	79,673	8.05	1.01	0.00	9.19
Traveltime	79,673	1,262.80	524.53	1.00	2,050.00
ln(Traveltime)	79,673	6.92	0.94	0.00	7.63
Dev Shop Allocation	74,676	0.0418	0.042	0.01	0.200
F&B	75,373	0.448	0.077	0	0.833

Sports sector

Descriptive statistics Sport(s) industry

Purchase Online	Frequency	Percentage
No	57,060	82.68%
Yes	11,949	17.32%
Total	69,009	100%
Central area	**Frequency**	**Percentage**
<10,000 sqm	22,834	33%
Between 10,000 & 40,000 sqm	24,879	36%
Between 40,000 & 60,000 sqm	5,521	8%
> 60,000 sqm	15,775	23%
Total	69,009	100%
Gender Distribution	**Frequency**	**Percentage**
Male	31,600	46%
Female	37,322	54%
Unknown	87	0%
Total	69,009	100%
Level of education	**Frequency**	**Percentage**
Primary & Lower Secondary & Upper first phase	13,817	20%
Post Lower Secondary & Upper Secondary second phase	22,267	32%
Bachelor	20,663	30%
Master/Doctoral	11,619	17%
Unknown	643	1%
Total	69,009	100%
Household composition	**Frequency**	**Percentage**
One-person	13,546	20%
Single parent	3,136	5%
Two-person	28,809	42%
Family with kids	21,098	31%
With parent(s)/legal guardian(s)	1,330	2%
Other	1,076	2%
Unknown	14	0%
Total	69,009	100%
Income distribution	**Frequency**	**Percentage**
< €950	1,601	2%
between € 951 & € 1,300	3,565	5%
between € 1,301 & € 1,900	8,169	12%
Between€ 1,901 & € 3,150	21,353	31%
> € 3,150	21,138	31%
Unknown	13,183	19%
Total	69,009	100%

Descriptive statistics Sports sector continued

Age distribution	Frequency	Percentage
15-24	2,179	3%
25-34	7,292	11%
35-44	10,186	15%
45-54	15,138	22%
55-64	16,254	24%
65-74	14,522	21%
>74	3,438	5%
Total	69,009	100%

Descriptive statistics Sports sector continued

Variable	N	Mean	Std. Dev	Min	Max
Rating Other	54241	7.38	0.49	5.5	8.4
Rating Access	54241	7.45	0.62	5.7	8.9
Rating Atmosphere	54241	7.37	0.32	6.3	8.4
Age	68643	52.85	14.73	15	95
Additional supply	69009	90,302.52	165,985.00	0.00	888,143.00
ln(Additional supply)	69009	9.82	2.33	0.00	1.37
Radius	68964	4,326.19	2,778.55	1.00	9,849.00
ln(Radius)	68964	8.05	1.00	0.00	9.19
Traveltime	68964	1,261.26	524.75	1.00	2,050.00
ln(Traveltime)	68964	6.01	0.94	0.00	7.63
Dev Shop Allocation	64,630	0.0418	0.042	0.01	0.200
F&B	65,213	0.448	0.077	0	0.833

Appendix 3: List of adjusted central locations.

List of adjusted Central locations for distance calculation

Town	New Central location	Town	New Central location
Aalsmeerderbrug	Aalsmeer	Maasdijk	Naaldwijk
Aalst (Gld.)	Wijk en Aalburg	Maasland	Maassluis
Aarlanderveen	Alphen aan den Rijn	Marken	Monnickendam
Aartswoud	Heerhugowaard	Maurik	Wijk bij Duurstede
Abbekerk	Hoorn	Meerkerk	Nieuwegein
Abbenbroek	Spijkenisse	Melissant	Middelharnis
Abbenes	Nieuw-vennep	Meteren	Middelharnis
Abcoude	Amsterdam Zuid-Oost	Middelie	Purmerend
Achterveld	Barneveld	Middenbeemster	Purmerend
Achterveld	Barneveld	Middenmeer	Medemblik
Achterveld gem. Leusden	Barneveld	Midwoud	Hoorn
Achthuizen	Roosendaal	Mijnsheerenland	Oud-Beijerland
Acquoy	Nieuwegein	Moerdijk	Zevenbergen
Aerdenhout	Heemstede	Moerkapelle	Zoetermeer
Akersloot	Castricum	Molenaarsgraaf	Dordrecht
Almkerk	Raamsdonkveer	Mookhoek	Dordrecht
Alphen	Alphen aan den Rijn	Moordrecht	Gouda
Ameide	Nieuwegein	Muiderberg	Muiden
Amerongen	Veenendaal	Nederhemert	Wijk en Aalburg
Ammerstol	Gouda	Nieuwveen	Nieuwkoop
Amstelhoek	Amstelveen	Nieuw-Vossemeer	Steenbergen
Andel	Gorinchem	Nigtevecht	Weesp
Andijk	Bovenkarspel	Nijkerkerveen	Nijkerk
Ankeveen	Hilversum	Noordbeemster	Purmerend
Arkel	Gorinchem	Noordeinde (Nh.)	Purmerend
Asperen	Nieuwegein	Noordeloos	Gorinchem
Austerlitz	Driebergen-Rijsenburg	Noorden	Mijdrecht
Avenhorn	Hoorn	Noordgouwe	Zierikzee
Baambrugge	Amsterdam Zuid-Oost	Noordhoek	Zevenbergen
Barsingerhorn	Schagen	Noord-Scharwoude	Heerhugowaard
Bavel	Breda	Noordwelle	Renesse
Beesd	Nieuwegein	Numansdorp	Oud-Beijerland
Beets (Nh.)	Hoorn	Obdam	Heerhugowaard
Beinsdorp	Hillegom	Ochten	Rhenen
Bennebroek	Heemstede	Odijk	Driebergen-Rijsenburg
Benningbroek	Hoorn	Ommeren	Rhenen
Benschop	Ijsselstijn	Ooltgensplaat	Steenbergen
Benthuizen	Zoetermeer	Oosteind	Oosterhout
Bentveld	Zandvoort	Oosterblokker	Hoorn
Bergen aan Zee	Bergen	Oosterend (Nh.)	Den Burg
Bergschenhoek	Berkel en Rodenrijs	Oosterland	Zierikzee
Berkenwoude	Gouda	Oosterleek	Hoorn
Berkhout	Hoorn	Oosterwijk	Leerdam

Town	New Central location	Town	New Central location
Beusichem	Nieuwegein	Oost-Graftdijk	Alkmaar
Blaricum	Laren	Oosthuizen	Purmerend
Bleskensgraaf	Dordrecht	Oostknollendam	Wormerveer
Bleskensgraaf	Dordrecht	Oostwoud	Hoorn
Bleskensgraaf Ca	Dordrecht	Oostzaan	Zaandam
Blokker	Hoorn	Ophemert	Tiel
Boesingheliede	Haarlem	Opheusden	Rhenen
Bosch en Duin	Bilthoven	Opijnen	Zaltbommel
Brandwijk	Dordrecht	Opmeer	Hoorn
Breezand	Den Helder	Opperdoes	Medemblik
Breukeleveen	Utrecht	Oterleek	Heerhugowaard
Broek in Waterland	Amsterdam	Otterlo	Ede
Broek op Langedijk	Heerhugowaard	Ottoland	Schoonhoven
Brouwershaven	Zierikzee	Oud Ade	Leiden
Bruchem	Zaltbommel	Oud Zuilen	Utrecht
Bruinisse	Zierikzee	Oud-Alblas	Alblasserdam
Buitenkaag	Sassenheim	Oude Meer	Aalsmeer
Buren (Gld.)	Tiel	Oude Niedorp	Heerhugowaard
Burgerbrug	Schagen	Oude Wetering	Nieuw-vennep
Burgerveen	Nieuw-vennep	Oude-Beijerland	Oud-Beijerland
Burgh-Haamstede	Renesse	Oudendijk	Oud-Beijerland
Buurmalsen	Geldermalsem	Oudendijk (Nh.)	Oud-Beijerland
Callantsoog	Schagen	Oudenhoorn	Hellevloetsluis
Cothen	Wijk bij Duurstede	Ouderkerk aan den IJssel	Krimpen aan den IJssel
Cruquius	Hoofddorp	Oudeschild	Den Burg
De Cocksdorp	Den Burg	Oudesluis	Schagen
De Glind	Barneveld	Oude-Tonge	Middelharnis
De Goorn	Hoorn	Oudkarspel	Heerhugowaard
De Heen	Bergen op Zoom	Oudorp (Nh.)	Alkmaar
De Hoef	Barendrecht	Oud-Vossemeer	Bergen op Zoom
De Koog	Den Burg	Overberg	Veenendaal
De Kwakel	Aalsmeer	Overveen	Haarlem
De Lier	Naaldwijk	Papekop	Woerden
De Rijp	Purmerend	Pernis Rotterdam	Vlaardingen
De Waal	Den Burg	Petten	Schagen
De Weere	Hoorn	Piershil	Spijkenisse
De Woude	Purmerend	Poederoijen	Wijk en Aalburg
De Zilk	Hillegom	Poeldijk	Naaldwijk
Deil	Geldermalsem	Polsbroek	Schoonhoven
Delfgauw	Delft	Poortugaal	Hoogvliet Rotterdam
Den Bommel	Roosendaal	Poortvliet	Bergen op Zoom
Den Hoorn	Delft	Purmer	Purmerend
Den Hoorn (Zh.)	Delft	Purmer	Purmerend
Den Hoorn Texel	Den Burg	Purmerland	
Den Ilp	Purmerend	Raamsdonk	Raamsdonkveer

Town	New Central location	Town	New Central location
Den Oever	Schagen	Ravenswaaij	Wijk bij Duurstede
Diemen	Amsterdam	Reeuwijk	Gouda
Dinteloord	Roosendaal	Renswoude	Veenendaal
Dirkshorn	Schagen	Rhenoy	Leerdam
Dirksland	Renesse	Rijnsaterwoude	Nieuw-vennep
Dodewaard	Tiel	Rijnsburg	Oegstgeest
Dorst	Breda	Rijpwetering	Leiden
Dreischor	Zierikzee	Rijsenhout	Nieuw-vennep
Driebruggen	Woerden	Rijswijk (Gld.)	Wijk bij Duurstede
Driehuis	Ijmuiden	Rijswijk (Nb.)	Gorinchem
Driehuis	Ijmuiden	Rockanje	Hellevloetsluis
Driehuis (Nh.)	Ijmuiden	Santpoort-Noord	Ijmuiden
Driehuizen	Alkmaar	Santpoort-Zuid	Ijmuiden
Duivendrecht	Amsterdam	Schagerbrug	Schagen
Dussen	Raamsdonkveer	Schardam	Hoorn
Echteld	Tiel	Scharendijke	Renesse
Eck en Wiel	Wijk bij Duurstede	Scharwoude	Hoorn
Ederveen	Veenendaal	Schellinkhout	Hoorn
Eemdijk	Laren	Schelluinen	Gorinchem
Eemnes	Laren	Schermerhorn	Alkmaar
Egmond aan den Hoef	Egmond aan Zee	Scherpenisse	Tholen
Egmond-Binnen	Egmond aan Zee	Schipluiden	Delft
Ellemeet	Renesse	Schoonrewoerd	Leerdam
Elst (Ut.)	Veenendaal	Schoorl	Bergen
Enspijk	Geldermalsem	's-Graveland	Hilversum
Erichem	Tiel	Sijbekarspel	Hoorn
Everdingen	Nieuwegein	Simonshaven	Spijkenisse
Fijnaart	Zevenbergen	Sint Maarten	Schagen
Gameren	Zaltbommel	Sint Maartensbrug	Schagen
Garderen	Raamsdonkveer	Sint Maartensvlotbrug	Schagen
Geertruidenberg	Spijkenisse	Sint Martensvlotbrug	Schagen
Geervliet	Spijkenisse	Sint Pancras	Heerhugowaard
Gelderswoude	Zoetermeer	Sint Philipsland	Steenbergen
Giessen	Zaltbommel	Sint-Annaland	Steenbergen
Giessenburg	Gorinchem	Sint-Maartensdijk	Tholen
Goedereede	Brielle	Sleeuwijk	Gorinchem
Gouderak	Gouda	Slootdorp	Medemblik
Goudriaan	Gorinchem	Snelrewaard	Oudewater
Goudswaard	Spijkenisse	Soesterberg	Zeist
Graft	Alkmaar	Sommelsdijk	Middelharnis
Groenekan	Utrecht	Spaarndam	Haarlem
Groot-Ammers	Schoonhoven	Spaarndam Gem. Haarlem	Haarlem
Grootebroek	Bovenkarspel	Spanbroek	Hoorn
Grootschermer	Alkmaar	Spierdijk	Hoorn
Haaften	Zaltbommel	Spijk gem. Lingewaal	Gorinchem

Town	New Central location	Town	New Central location
Haarzuilens	Utrecht	Spijkerboor (Nh.)	Purmerend
Haastrecht	Gouda	Sprang-Capelle	Waalwijk
Hagestein	Nieuwegein	Stad aan 't Haringvliet	Middelharnis
Halfweg	Haarlem	Standdaarbuiten	Oudenbosch
Halfweg (Nh.)	Haarlem	Starnmeer	Alkmaar
Hank	Oosterhout	Stavenisse	Tholen
Haringhuizen	Schagen	Stellendam	Hellevloetsluis
Harmelen	Woerden	Stolwijk	Gouda
Harskamp	Barneveld	Stompetoren	Alkmaar
Hauwert	Hoorn	Stoutenburg	Amersfoort
Hazerswoude	Alphen aan den Rijn	Streefkerk	Alblasserdam
Hazerswoude-Dorp	Alphen aan den Rijn	Strijensas	s-gravendeel
Hazerswoude-Rijndijk	Alphen aan den Rijn	Stroe	Barneveld
Heenvliet	Spijkenisse	't Goy	Houten
Heerjansdam	Barendrecht	't Veld	Schagen
Heesselt	Zaltbommel	t Zand	Anna Paulowna
Hei- en Boeicop	Nieuwegein	't Zand	Anna Paulowna
Heijningen	Roosendaal	't Zand (Nh.)	Anna Paulowna
Heinenoord	Barendrecht	Ter Aar	Alphen aan den Rijn
Hekelingen	Spijkenisse	Ter Heijde	Monster
Hekendorp	Gouda	Terheijden	Breda
Hellouw	Zaltbommel	Terschuur	Barneveld
Hensbroek	Heerhugowaard	Teteringen	Breda
Herkingen	Renesse	Tienhoven (Zh.)	Nieuwegein
Herwijnen	Zaltbommel	Tinte	Brielle
Heukelum	Gorinchem	Tricht	Geldermalsem
Hippolytushoef	Schagen	Tuil	Zaltbommel
Hollandsche Rading	Hilversum	Tuitjenhorn	Schagen
Honselersdijk	Naaldwijk	Tull en 't Waal	Nieuwegein
Hoogblokland	Gorinchem	Twisk	Medemblik
Hooge Zwaluwe	Made	Uitdam	Monnickendam
Hoogkarspel	Enkhuizen	Uitgeest	Heemskerk
Hoogland	Amersfoort	Uithoorn	Aalsmeer
Hooglanderveen	Amersfoort	Ursem	Heerhugowaard
Hoogmade	Leiden	Valkenburg (Zh.)	Leiden
Hoogwoud	Hoorn	Varik	Tiel
Hoornaar	Gorinchem	Velsen-Noord	Beverwijk
Huis ter Heide (Ut.)	Zeist	Velsen-Zuid	Beverwijk
Huisduinen	Den Helder	Velserbroek	Santpoort-Noord
IJzendoorn	Tiel	Venhuizen	Hoorn
Ilpendam	Purmerend	Vianen	Nieuwegein
Ingen	Wijk bij Duurstede	Vianen (Ut.)	Nieuwegein
Jaarsveld	Nieuwegein	Vierpolders	Brielle
Jisp	Purmerend	Vijfhuizen	Haarlem
Julianadorp	Den Helder	Vinkeveen	Mijdrecht

Town	New Central location	Town	New Central location
Kaag	Sassenheim	Vleuten	
Kamerik	Woerden	Vlist	Gouda
Katwoude	Monnickendam	Vogelenzang	Hoofddorp
Kedichem	Gorinchem	Voorburg	:eidschendam
Kerk-Avezaath	Tiel	Voorhout	Oegstgeest
Kerkwerve	Zierikzee	Vreeland	Hilversum
Kesteren	Rhenen	Vuren	Gorinchem
Kinderdijk	Ridderkerk	Waardenburg	Zaltbommel
Klaaswaal	Oud-Beijerland	Waarder	Woerden
Klundert	Zevenbergen	Waarland	Schagen
Kockengen	Utrecht	Waddinxveen	Gouda
Koedijk	Alkmaar	Wadenoijen	Tiel
Koedijk	Alkmaar	Wagenberg	Breda
Koedijk gem. Alkmaar	Alkmaar	Warder	Edam
Kolhorn	Schagen	Warmenhuizen	Bergen
Koog	Den Burg	Warmond	Oegstgeest
Koog aan de Zaan	Zaandam	Waspik	Raamsdonkveer
Kootwijk	Apeldoorn	Watergang	Haarlem
Kootwijkerbroek	Barneveld	Waverveen	Mijdrecht
Kortenhoef	Hilversum	Wekerom	Barneveld
Koudekerk aan den Rijn	Alphen aan den Rijn	Werkhoven	Houten
Kreileroord	Enkhuizen	Wervershoof	Medemblik
Krimpen aan de Lek	Krimpen aan den IJssel	Westbeemster	Purmerend
Kruisland	Roosendaal	Westbroek	Utrecht
Kudelstaart	Amstelveen	West-Graftdijk	Alkmaar
Kwadijk	Purmerend	Westknollendam	Zaandam
Kwintsheul	s-Gravenhage	Westmaas	Oud-Beijerland
Lage Vuursche	Soest	Westwoud	Hoorn
Lage Zwaluwe	Made	Westzaan	Zaandam
Langbroek	Wijk bij Duurstede	Weteringbrug	Nieuw-vennep
Langerak (Zh.)	Schoonhoven	Wieringerwaard	Anna Paulowna
Langeweg	Zevenbergen	Wieringerwerf	Medemblik
Leerbroek	Nieuwegein	Wijdenes	Hoorn
Leiderdorp	Leiden	Wijdewormer	Zaandam
Leimuiden	Nieuw-vennep	Wijk aan Zee	Beverwijk
Leimuiderbrug	Nieuw-vennep	Wijngaarden (Zh.)	Sliedrecht
Lexmond	Nieuwegein	Willemstad	Oud-Beijerland
Lienden	Rhenen	Willemstad (Nb.)	Oud-Beijerland
Lijnden	Amsterdam	Wilnis	Mijdrecht
Limmen	Heiloo	Winkel	Schagen
Linschoten	Woerden	Woerdense Verlaat	Woerden
Loenen aan de Vecht	Breukelen	Wormer	Zaandam
Loenen aan den Vecht	Breukelen	Woubrugge	Alphen aan den Rijn
Loenersloot	Breukelen	Woudrichem	Gorinchem
Loosdrecht	Hilversum	Zaandijk	Zaandam

Town	New Central location	Town	New Central location
Lopik	Schoonhoven	Zegveld	Woerden
Lopikerkapel	Schoonhoven	Zevenbergschen Hoek	Zevenbergen
Lutjebroek	Enkhuizen	Zevenhoven	Mijdrecht
Lutjewinkel	Schagen	Zevenhuizen	Zoetermeer
Maarn	Woudenberg	Zevenhuizen (Zh.)	Zoetermeer
Maarsbergen	Woudenberg	Zijderveld	Nieuwegein
Maartensdijk	Utrecht	Zijdewind	Schagen
Maasdam	Dordrecht	Zoelen	Tiel
Nederhorst	Hilversum	Zoelmond	Wijk bij Duurstede
Nederhorst den Berg	Hilversum	Zoeterwoude	Leiden
Neerijnen	Barneveld	Zonnemaire	Zierikzee
Nibbixwoud	Hoorn	Zuid-Beijerland	Oud-Beijerland
Nieuwaal	Zaltbommel	Zuidermeer	Hoorn
Nieuw-Beijerland	Oud-Beijerland	Zuiderwoude	Monnickendam
Nieuwe Niedorp	Schagen	Zuidland	Spijkenisse
Nieuwe Wetering	Nieuw-vennep	Zuidoostbeemster	Purmerend
Nieuwekerk aan den Ijssel	Capelle aan den Ijssel	Zuid-Scharwoude	Heerhugowaard
Nieuwendijk (Nb.)	Raamsdonkveer	Zuidschermer	Zaltbommel
Nieuwer ter Aa	Breukelen	Zuilichem	Hoorn
Nieuwerbrug aan den Rijn	Woerden	Zwaag	Hoorn
Nieuwerkerk	Capelle aan den Ijssel	Zwaagdijk-Oost	Hoorn
Nieuwerkerk aan den Ijssel	Capelle aan den Ijssel	Zwaagdijk-West	Hoorn
Nieuwerkerk Ad IJssel	Capelle aan den Ijssel	Zwaanshoek	Hoofddorp
Nieuwersluis	Utrecht	Zwammerdam	Bodegraven
Nieuwe-Tonge	Middelharnis	Zwanenburg	Amsterdam
Nieuwland	Gorinchem	Zwartebroek	Hovelaken
Nieuw-Lekkerland	Alblasserdam	Zwartewaal	Brielle
Nieuwpoort	Schoonhoven		

Appendix 4: Correlation Matrix

	Size 5k sqm	Size 10k sqm	Size 20k sqm	Size 40k sqm	Size 60k sqm	Size Large	Male
Size 5k sqm	1						
Size 10k sqm	-0.1033	1					
Size 20k sqm	0	-0.2559	1				
Size 40k sqm	-0.1029	-0.1647	-0.255	1			
Size 60k sqm	-0.1002	-0.1605	-0.2484	-0.1599	1		
Size Large	-0.1419	-0.2272	-0.3518	-0.2264	-0.2206	1	
Male	-0.0122	-0.0162	0.0072	-0.0004	0.0122	0.0032	1
Female	0.0122	0.0163	-0.0072	0.0005	-0.0122	-0.0032	-1
Education Low	0.013	0.0159	0.0086	0.0237	-0.0005	-0.0484	-0.062
Education Middle	0.0018	-0.0104	0.0011	0.0227	0.0163	-0.0252	-0.0289
Bachelor	-0.0035	-0.0048	0.0109	-0.0094	-0.0074	0.008	0.0361
Master/Doctoral	-0.015	-0.0015	-0.0276	-0.0439	-0.0105	0.0832	0.0648
One-person Household	-0.0311	-0.0347	-0.0375	-0.0271	0.0169	0.0942	-0.0297
Single Parent	-0.0105	-0.0133	-0.0069	-0.0046	0.0137	0.0168	-0.0794
Two-person Household	0.0132	0.0216	0.0345	0.0152	-0.0009	-0.0734	0.1317
Family	0.0247	0.023	0.0045	0.013	-0.0226	-0.0299	-0.0728
With Legal guardian/parents	-0.0071	-0.0194	-0.0049	0.0079	-0.002	0.0203	-0.0151
Household other	-0.0074	-0.0056	-0.0085	-0.0136	0.004	0.0257	-0.0242
Income U950	-0.0196	-0.0237	-0.0245	-0.0124	0.0176	0.0524	-0.0429
Income U1300	-0.0112	-0.0179	-0.0252	-0.0097	0.0166	0.0423	-0.0674
Income U1900	-0.0014	-0.009	-0.0056	-0.0028	0.0054	0.012	-0.0413
Income U3150	-0.0014	0.0044	0.0019	0.0122	-0.0023	-0.0129	0.0581
Income O3150	0.0191	0.0139	0.0171	-0.0077	-0.0189	-0.0187	0.1331
Income Unknown	-0.0038	0.0076	0.0088	0.0079	0.002	-0.0215	-0.1237
Age	0.021	0.0402	0.058	0.007	0.0003	-0.1121	0.1922
Rating Atmosphere	-0.2915	-0.2655	-0.0342	0.0484	0.1082	0.2907	0.0084
Rating Feeling	0.1912	0.1702	0.1876	0.0746	-0.1217	-0.4086	-0.0059
Rating Access	0.2121	0.2501	0.2862	0.0935	-0.0985	-0.6241	-0.0073
Rating Other	-0.4128	-0.3	0.0008	0.0897	0.1459	0.2859	0.0108
Rating F&B	-0.0164	-0.0819	0.0016	0.0511	0.0118	0.0233	0
Rating Supply	-0.2836	-0.2645	-0.1123	0.0386	0.0562	0.4179	0.0135
Traveldistance	-0.0049	-0.0007	0.006	-0.0335	0.0158	0.0117	0.0022
Deviation in shop allocation	0.1409	0.3305	0.0254	-0.1799	-0.0555	-0.1851	0.0004
F&B amenities	0.0956	-0.0449	-0.2419	-0.2239	-0.0016	0.4237	0.0007

Female	Education Low	Education Middle	Bachelor	Master/Doctoral
1				
0.062	1			
0.0289	-0.3729	1		
-0.0361	-0.3429	-0.429	1	
-0.0649	-0.2376	-0.2973	-0.2734	1
0.0298	0.0273	-0.0078	-0.0081	-0.0123
0.0794	-0.013	0.0245	-0.0055	-0.0078
-0.1317	0.0823	-0.0477	-0.006	-0.0298
0.0728	-0.1103	0.0293	0.0276	0.0632
0.0152	-0.0152	0.0663	-0.0278	-0.0323
0.0242	-0.0007	0.007	-0.0067	-0.0132
0.0429	0.0462	0.0269	-0.0417	-0.0358
0.0674	0.1267	0.0183	-0.0774	-0.0776
0.0413	0.1414	0.0413	-0.08	-0.1134
-0.0581	-0.0116	0.0507	0.0213	-0.0646
-0.1331	-0.2179	-0.1196	0.1277	0.2595
0.1237	0.0412	0.0181	-0.0357	-0.0578
-0.1922	0.2145	-0.0819	-0.0561	-0.0856
-0.0084	-0.0105	-0.0067	0.0031	0.0181
0.006	0.0305	0.0013	-0.0044	-0.0364
0.0074	0.0673	0.0374	-0.0144	-0.1123
-0.0108	-0.0149	-0.0063	0.0034	0.0224
0	-0.0032	-0.0171	0.0042	0.0186
-0.0135	-0.0276	-0.0118	0.0035	0.0458
-0.0022	0.0005	-0.001	0.0014	0.0007
-0.0004	0.0051	-0.0057	-0.006	0.0077
-0.0007	-0.0195	-0.0174	-0.0003	0.0471

|	One-person H	Single Parent	Two-person Hous	Family	With Legal gua	Household ot	Income U950	Income U1300	Income U1900	Income U3150	Income O3150	Income Unknown
One-person Household	1											
Single Parent	-0.1194	1										
Two-person Household	0	-0.1871	1									
Family	-0.3238	-0.1231	-0.5074	1								
With Legal guardian/parents	-0.0726	-0.0276	-0.1138	-0.0748	1							
Household other	-0.07	-0.0266	-0.1096	-0.0721	-0.0162	1						
Income U950	0.1495	0.0531	-0.111	-0.0776	0.06	0.0522	1					
Income U1300	0.2103	0.0849	-0.1273	-0.1075	0.0064	0.0152	-0.0429	1				
Income U1900	0.1568	0.0523	-0.0529	-0.1166	-0.0075	0.0013	-0.0648	-0.1019	1			
Income U3150	0.009	-0.011	0.0338	-0.03	-0.0288	-0.0138	-0.1101	-0.1731	-0.2619	1		
Income O3150	-0.2455	-0.0758	0.0875	0.2002	-0.0427	-0.0364	-0.1017	-0.1598	-0.2418	-0.4107	1	
Income Unknown	-0.0583	-0.0204	0.0308	0.0069	0.06	0.0255	-0.081	-0.1273	-0.1926	-0.3271	-0.3021	1
Age	0.0482	-0.0685	0.3467	-0.3214	-0.2538	-0.0564	-0.1175	0.0024	0.0567	0.0376	-0.0617	0.0233
Rating Atmosphere	0.0357	0.0037	-0.0183	-0.0214	0.0076	0.0113	0.0239	0.0197	0.0078	0.0041	-0.0167	-0.0143
Rating Feeling	-0.0663	-0.0198	0.0505	0.0248	-0.0138	-0.0143	-0.0372	-0.0314	-0.01	0.01	0.0197	0.0091
Rating Access	-0.1088	-0.0147	0.0707	0.0437	-0.009	-0.028	-0.0601	-0.0412	-0.0072	0.0171	0.0054	0.03
Rating Other	0.0433	0.0067	-0.0222	-0.0254	0.007	0.0107	0.0249	0.0227	0.0074	0.0026	-0.0164	-0.0149
Rating F&B	0.0107	-0.009	0.0034	-0.0088	-0.0076	0.0029	0.0032	0.0005	0.0067	0.004	0.0038	-0.0163
Rating Supply	0.0478	0.0085	-0.0295	-0.0263	0.0144	0.0168	0.0306	0.018	0.0084	-0.0019	-0.012	-0.015
Traveldistance	0.0024	-0.0012	-0.0039	0.0013	0.0075	-0.0031	0.0003	-0.0009	0.0028	-0.0015	0.0028	-0.0035
Deviation in shop allocation	-0.0094	-0.0137	0.0153	-0.0014	-0.0054	0.0042	-0.0038	-0.0089	-0.001	0.0011	0.0019	0.0045
F&B amenities	0.0867	0.0123	-0.0489	-0.0446	0.0067	0.0266	0.049	0.0404	0.0205	-0.0081	-0.0218	-0.0284

	Age	Rating Atmosphere	Rating Feeling	Rating Access	Rating Other	Rating F&B	Rating Supply	Traveldistance	Deviation in shop allocation	F&B amenities
Age	1									
Rating Atmosphere	-0.0303	1								
Rating Feeling	0.0761	0.3155	1							
Rating Access	0.094	-0.1472	0.6037	1						
Rating Other	-0.0336	0.9025	0.1023	-0.2574	1					
Rating F&B	0.0168	0.5378	0.3922	0.021	0.534	1				
Rating Supply	-0.0473	0.688	0.2187	-0.2279	0.5466	0.3491	1			
Traveldistance	0.0004	0.0184	-0.0137	-0.0148	0.033	0.0308	0.0274	1		
Deviation in shop allocation	0.0309	-0.2246	0.1555	0.1385	-0.2935	-0.1567	-0.2437	-0.0082	1	
F&B amenities	-0.0577	0.0273	-0.3841	-0.5168	0.1008	0.1724	-0.019	0.021	-0.2234	1

Appendix 5: Output PCA

Principal components/correlation

Number of observations	56,523
Number of components	13
Trace	13
Rho	1

Variable	Compound 1	Compound 2
Supply & diversity of shops		0.3915
Atmosphere		0.4552
Supply F&B		0.3823
Events		0.402
Facilities		0.3601
Cleanliness	0.3625	
Safety	0.3198	
Parking facilities (car)	0.3937	
Parking tariffs	0.3973	
Parking facilities (bicycle)	0.3643	
Accessibility public transportation	0.4098	
Accessibility car		
Accessibility bicycle		

Scree plot of eigenvalues after pca
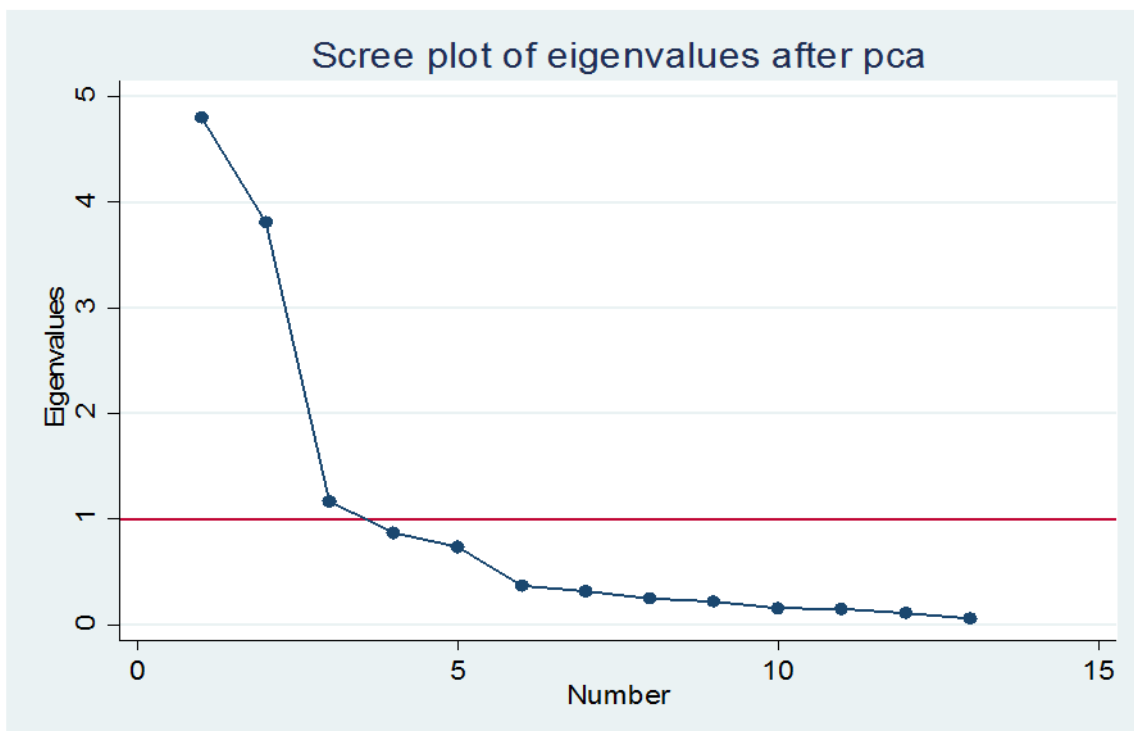
Appendix 6: VIF scores per product group

Product group	Electronics	Electronics	Apparel	Apparel	Media	Media	Sport	Sport
Model #	(1)	(2)	(1)	(2)	(1)	(2)	(1)	(2)
Central area size								
<5000 sqm	2.54	1.65	2.54	1.49	2.5	1.64	2.55	1.68
10,000sqm - 20,000 sqm	1.66	1.69	1.65	1.74	1.67	1.7	1.67	1.7
20,000 sqm - 40,000 sqm	1.64	1.73	1.63	1.78	1.62	1.72	1.64	1.74
40,000 sqm - 60,000 sqm	1.52	1.76	1.53	1.97	1.5	1.73	1.49	1.72
>60,000 sqm	3.51	6.02	3.54	6.7	3.54	6.12	3.52	6.08
Rating Vitality		4.46		5.82		4.46		4.49
Rating Accessibility		1.5		1.12		1.5		1.48
Amsterdam	1.18	1.4	1.18	1.47	1.18	1.41	1.17	1.38
Utrecht	1.37	1.55	1.38	1.63	1.39	1.57	1.39	1.58
Rotterdam	1.34	1.47	1.35	1.53	1.33	1.46	1.33	1.45
The Hague	1.58	1.87	1.6	1.92	1.58	1.87	1.57	1.85
Deviation from mean allocation	6.41	7.93	6.42	7.91	6.44	7.94	6.47	7.98
Distance	42.73	49.69	42.59	49.97	42.97	49.57	42.72	49.58
% F&B								
40% - 50%	3.13	3.21	3.12	3.27	3.12	3.21	3.11	3.18
>50%	3.04	3.45	3.07	3.6	3.06	3.48	3.04	3.47
Age group								
< 25	1.45	1.44	1.52	1.51	1.49	1.48	1.54	1.53
25 - 34	1.51	1.52	1.54	1.55	1.52	1.54	1.61	1.62
35 - 44	1.74	1.74	1.75	1.75	1.76	1.77	1.89	1.9
45 - 54	2.01	2.01	2.02	2.01	2.03	2.03	2.17	2.17
65 -74	2.07	2.07	2.06	2.07	2.04	2.04	1.97	1.97
>75	1.33	1.33	1.33	1.34	1.31	1.32	1.22	1.22
Wage level								
< €950	1.21	1.22	1.23	1.24	1.21	1.22	1.21	1.22
€950 - €1,300	1.34	1.35	1.37	1.37	1.32	1.32	1.29	1.3
€1,300 - €1,900	1.53	1.53	1.55	1.55	1.5	1.5	1.47	1.47
> €3,150	2.18	2.2	2.16	2.18	2.24	2.25	2.3	2.32
Household composition								
Single-person	1.91	1.94	1.94	1.97	1.92	1.95	1.85	1.89
Single-parent	1.24	1.25	1.26	1.26	1.24	1.25	1.29	1.29
Family with kids	2.17	2.15	2.17	2.15	2.19	2.18	2.4	2.38
Living with Parents/Legal guardian	1.3	1.29	1.35	1.33	1.32	1.3	1.35	1.32
Level of Education								
Low	1.81	1.82	1.84	1.84	1.72	1.72	1.69	1.7
Bachelor	1.99	2.01	1.97	1.99	2.06	2.08	2.05	2.07
Doctoral	1.7	1.74	1.69	1.73	1.78	1.82	1.77	1.82
Gender								
Male	2.19	2.18	2.1	2.09	2.13	2.13	2.12	2.12
Additional Supply	44.97	62.33	45.07	62.11	45.46	62.48	45.12	62.6
Constant	4.49	5.24	4.5	5.32	4.52	5.25	4.52	5.27

Appendix 7: Do-file of the analysis

Do-File for the Electronics product group. Do-Files are identical among the product groups

```
set mem 5000
set matsize 11000
set maxvar 32767
import excel "X:\My Documents\Stata\Deze Justin\Deze ja\Je bent er bijna\Excel files per sector\H&L
gesplitst\Electronics Stata.xlsx", sheet("Blad1") firstrow
encode reistijd, generate(Traveltime)
encode straal, generate (Radius)
encode reisafstand, generate(Traveldistance)
generate lnTraveltime = ln(Traveltime)
generate lnRadius = ln(Radius)
generate lnAdditionalSupply = ln(Extrametrages)
generate DummySize5k = 0
replace DummySize5k = 1 if var45 < 5000
generate DummySize10k = 0
replace DummySize10k = 1 if var45 >= 5000 & var45 < 10000
generate DummySize20k = 0
replace DummySize20k = 1 if var45 >= 10000 & var45 < 20000
generate DummySize40k = 0
replace DummySize40k = 1 if var45 >= 20000 & var45 < 40000
generate DummySize60k = 0
replace DummySize60k = 1 if var45 >= 40000 & var45 < 60000
generate DummySizeLarge = 0
replace DummySizeLarge = 1 if var45 >= 60000
generate RatingAtmosphere = (faciliteiten + sfeerenuitstraling)/2
generate RatingFeeling = (veiligheid + netheid)/2
generate RatingAccess = (parkeermogelijkhedenauto + parkeertarievenauto + stallingsmogelijkhedenfiets +
bereikbaarheidauto + bereikbaarheidOV + bereikbaarheidfiets)/6
generate RatingOther = (evenementen + faciliteiten)/2
generate RatingFandB = aanbodhoreca
generate RatingSupply = winkelaanbod
drop if V34_aankoopnr == 1
drop if V34_aankoopnr == 2
drop if V34_aankoopnr == 4
drop if V34_aankoopnr == 5
drop if V34_aankoopnr == 6
drop if V34_aankoopnr == 7
drop if V34_aankoopnr == 11
drop  reistijd reisafstand straal Provincie
drop kern_classificatie
drop V34_kern_classificatie V34_gemeente_classificatie
generate Online = 0
replace Online = 1 if V34_aankoopnr == 3
generate Male = 0
replace Male = 1 if Geslacht == 1
generate Female = 0
replace Female = 1 if Geslacht == 2
generate EduLow = 0
generate EduMid = 0
generate Bachelor = 0
generate MasterDoctoral = 0
replace EduLow = 1 if Opleiding == 1
replace EduLow = 1 if Opleiding == 2
replace EduLow = 1 if Opleiding == 3
replace EduMid = 1 if Opleiding == 4
replace EduMid = 1 if Opleiding == 5
replace Bachelor = 1 if Opleiding == 6
replace MasterDoctoral = 1 if Opleiding == 7
generate HH1 = 0
generate HHSP = 0
generate HH2 = 0
generate HHFam = 0
generate HHLG = 0
generate HHother = 0
replace HH1 = 1 if samhshdn == 1
replace HHSP = 1 if samhshdn == 2
```

```
replace HH2 =1 if samhshdn == 3
replace HHFam = 1 if samhshdn == 4
replace HHLG = 1 if samhshdn == 5
replace HHother = 1 if samhshdn == 6
replace HHother = 1 if samhshdn == 7
replace HHother = 1 if samhshdn == 8
replace HHother = 1 if samhshdn == 9
generate IncomeU950 = 0
generate IncomeU1300 = 0
generate IncomeU1900 = 0
generate IncomeU3150 = 0
generate IncomeO3150 = 0
generate IncomeUnknown = 0
replace IncomeU950 = 1 if Inkomencat == 1
replace IncomeU1300 = 1 if Inkomencat == 2
replace IncomeU1900 = 1 if Inkomencat == 3
replace IncomeU3150 = 1 if Inkomencat == 4
replace IncomeO3150 = 1 if Inkomencat == 5
replace IncomeUnknown = 1 if Inkomencat == 6
replace IncomeUnknown = 1 if Inkomencat == 7
replace IncomeUnknown = 1 if Inkomencat == 8
replace IncomeUnknown = 1 if Inkomencat == 9
generate Under25 = 0
generate Under35 = 0
generate Under45 = 0
generate Under55 = 0
generate Under65 = 0
generate Under75 = 0
generate Over75 = 0
generate AgeOther = 0
replace Under25 = 1 if Leeftijd <= 24
replace Under35 = 1 if Leeftijd >= 25 & Leeftijd < 35
replace Under45 = 1 if Leeftijd >= 35 & Leeftijd < 45
replace Under55 = 1 if Leeftijd >= 45 & Leeftijd < 55
replace Under65 = 1 if Leeftijd >= 55 & Leeftijd < 65
replace Under75 = 1 if Leeftijd >= 65 & Leeftijd < 75
replace Over75 = 1 if Leeftijd >= 75
replace AgeOther = 1 if Leeftijd == 997
replace AgeOther =1 if Leeftijd == 998
replace AgeOther =1 if Leeftijd == 999
mvdecode Leeftijd, mv(998)
tabulate Online, missing
tabulate DummySize5k, missing
tabulate DummySize10k, missing
tabulate DummySize20k, missing
tabulate DummySize40k, missing
tabulate DummySize60k, missing
tabulate DummySizeLarge, missing
tabulate Male, missing
tabulate Female, missing
tabulate EduLow, missing
tabulate EduMid, missing
tabulate Bachelor, missing
tabulate MasterDoctoral, missing
tabulate HH1, missing
tabulate HHSP, missing
tabulate HH2, missing
tabulate HHFam, missing
tabulate HHLG, missing
tabulate HHother, missing
tabulate IncomeU950, missing
tabulate IncomeU1300, missing
tabulate IncomeU1900, missing
tabulate IncomeU3150, missing
tabulate IncomeO3150, missing
tabulate IncomeUnknown, missing
tabulate Under25, missing
```

```
tabulate Under35, missing
tabulate Under45, missing
tabulate Under55, missing
tabulate Under65, missing
tabulate Under75, missing
tabulate Over75, missing
tabulate AgeOther, missing
summarize Radius lnRadius Extrametrages lnAdditionalSupply Traveltime lnTraveltime RatingOther
RatingAccess RatingAtmosphere Leeftijd STD horeca
replace lnAdditionalSupply = 0 if (lnAdditionalSupply == .)
summarize lnAdditionalSupply
replace lnAdditionalSupply = 0 if (lnAdditionalSupply == .)
summarize lnAdditionalSupply
corr horeca hor aanbodhoreca
corr STD winkelaanbod
gen Testsize = 0
replace Testsize = 1 if DummySize5k == 1
replace Testsize = 2 if DummySize10k == 1
replace Testsize = 3 if DummySize20k == 1
replace Testsize = 4 if DummySize40k == 1
replace Testsize = 5 if DummySize60k == 1
replace Testsize = 7 if DummySizeLarge == 1
gen Testage = 0
replace Testage = 1 if Under25 == 1
replace Testage = 2 if Under35 == 1
replace Testage = 3 if Under45 == 1
replace Testage = 4 if Under55 == 1
replace Testage = 5 if Under65 == 1
replace Testage = 6 if Under75 == 1
replace Testage = 7 if Over75 == 1
replace Testage = 8 if AgeOther == 1
gen Testwage = 0
replace Testwage = 1 if IncomeU950 == 1
replace Testwage = 2 if IncomeU1300 == 1
replace Testwage = 3 if IncomeU1900 == 1
replace Testwage = 4 if IncomeU3150 == 1
replace Testwage = 5 if IncomeO3150 == 1
replace Testwage = 6 if IncomeUnknown == 1
gen TestHH = 0
replace TestHH = 1 if HH1 == 1
replace TestHH = 2 if HHSP == 1
replace TestHH = 3 if HH2 == 1
replace TestHH = 4 if HHFam == 1
replace TestHH = 5 if HHLG == 1
replace TestHH = 6 if HHother == 1
gen TestEdu = 0
replace TestEdu = 1 if EduLow == 1
replace TestEdu = 2 if EduMid == 1
replace TestEdu = 3 if Bachelor == 1
replace TestEdu = 4 if MasterDoctoral == 1
gen TestGender = 0
replace TestGender = 1 if Male == 1
replace TestGender = 2 if Female == 1
corr Male Female EduLow EduMid Bachelor MasterDoctoral HH1 HHSP HH2 HHFam HHLG HHother
IncomeU950 IncomeU1300 IncomeU1900 IncomeU3150 IncomeO3150 IncomeUnknown Leeftijd
RatingAtmosphere RatingFeeling RatingAccess RatingOther RatingFandB RatingSupply Traveltime STD horeca
generate RatingPractical = (veiligheid + netheid + parkeermogelijkhedenauto + parkeertarievenauto +
stallingsmogelijkhedenfiets + bereikbaarheidauto + bereikbaarheidOV + bereikbaarheidfiets)/8
generate RatingVitality = (evenementen + faciliteiten + aanbodhoreca + winkelaanbod)/4
corr DummySize5k DummySize10k DummySize20k DummySize40k DummySize60k DummySizeLarge Male
Female EduLow EduMid Bachelor MasterDoctoral HH1 HHSP HH2 HHFam HHLG HHother IncomeU950
IncomeU1300 IncomeU1900 IncomeU3150 IncomeO3150 IncomeUnknown Leeftijd RatingAtmosphere
RatingFeeling RatingAccess RatingOther RatingFandB RatingSupply Traveltime STD horeca
fvset base 5 Testage
tabulate Testwage
fvset base 4 Testwage
tabulate Testsize
```

```
fvset base 3 Testsize
tabulate TestHH
fvset base 3 TestHH
tabulate TestEdu
fvset base 2 TestEdu
tabulate TestGender
fvset base 2 TestGender
logistic Online ib(2).Testsize ib(5).Testage ib(4).Testwage ib(3).TestHH ib(1).TestEdu ib(2).TestGender
RatingVitality lnTraveltime lnAdditionalSupply horeca STD
vif, uncentered
generate FB1 = 0
generate FB2 = 0
generate FB3 = 0
replace FB1 = 1 if horeca <0.4
replace FB2 = 1 if horeca >= 0.4 & horeca <0.5
replace FB3 = 1 if horeca >= 0.5
gen TestFB = 0
replace TestFB = 1 if FB1 == 1
replace TestFB = 2 if FB2 == 1
replace TestFB = 3 if FB3 == 1
replace Testwage = 6 if Testwage == 0
replace TestHH = 6 if TestHH == 0
drop if IncomeUnknown == 1
drop if AgeOther == 1
drop if Geslacht == 7
drop if Geslacht == 8
drop if Geslacht == 9
drop if Opleiding == 8
drop if Opleiding == 18
drop if Opleiding == 19
drop if samhshdn == 7
drop if samhshdn == 8
drop if samhshdn == 9
drop if Geslacht == .
encode Plaatsnaamhuis, generate (Plaats)
gen Adam = 0
gen Rdam = 0
gen Utr = 0
gen TH = 0
replace Adam = 1 if Plaatsnaamhuis == "Amsterdam"
replace Utr = 1 if Plaatsnaamhuis == "Utrecht"
replace Rdam = 1 if Plaatsnaamhuis == "Rotterdam"
replace TH = 1 if Plaatsnaamhuis == "'s-Gravenhage"
logistic Online i.Testsize lnRadius i.TestFB STD i.Testage i.Testwage i.TestHH i.TestEdu i.TestGender
lnAdditionalSupply Adam Utr Rdam TH
vif, uncentered
drop if RatingVitality ==.
pca winkelaanbod sfeerenuitstraling  aanbodhoreca evenementen faciliteiten netheid veiligheid
parkeermogelijkhedenauto parkeertarievenauto stallingsmogelijkhedenfiets bereikbaarheidauto bereikbaarheidOV
bereikbaarheidfiets
screeplot, yline(1)
predict PCRatingVit PCRatingAcc
logistic Online i.Testsize PCRatingVit PCRatingAcc Adam Utr Rdam TH lnRadius i.TestFB STD i.Testage
i.Testwage i.TestHH i.TestEdu i.TestGender lnAdditionalSupply
vif, uncentered
```