ERASMUS UNIVERSITY ROTTERDAM

Erasmus School of Economics

Master Thesis [Urban Port and Transport Economics]

The Impact of Online Retailing on Physical Shopping Travel Patterns

Quantitative analysis at the Netherlands

Name student: Christian Poort

Student ID Number: 360135

Supervisor: Jan-Jelle Witte

Second assessor: Giuliano Mingardo

Date final version: 20-08-2018

Abstract

The shift from offline to online shopping is accompanied with many changes for the retail as well as the transport sector. One of the effects of the rising use of the Internet is that more and more people also shop online and this could imply that also the need to travel to stores decreases. Physical stores are however still in business and trips to these stores are still be made. It implies that people not always as rational human beings that act in a cost minimizing. There could also be such thing where people derive utility from shopping and shopping can be seen as an activity that is fun. This research tries to find if the rising amount of online retail has an effect on transport behaviour. A dataset for the Netherlands is used to investigate this effect quantitatively. The results do show that there is a positive effect of online frequency use on the travel distance for shopping purposes. A different perception of the gain in utility derived from fun shopping could not be found in this investigation. The research as well as literature however did find significant effects for personal characteristics like income, urbanity and gender on travel distance for physical shopping trips.

Table of Contents

1. Introduction

Many trips are made daily all over the world for different purposes. Some people need to travel because of their work; others travel for recreational purposes and others again travel to gather products that they need or want. Back in the days people got their products from their own land, items that they could not get from their own land were acquired via trade. The exchange of goods became much easier by the introduction of money. Together with the innovations in transport was it possible to get almost all the items that are available all over the world. In modern society is the process of gathering products done by visiting stores where these products are available and can be bought with money. This process takes time when consumers need to visit multiple stores at different locations (Jones & Simmons, 1990). The fierce competitive environment by the retail market with its low barriers to entry results in an enormous amount of goods and shops in society, this again contributes to rising searching costs for goods and products by consumers (van Wee, Rietveld, & Meurs, 2002).

1.1 Online retail

The rise of the Internet resulted in lots of benefits for retailers, but also for customers that shop for their goods. By new online retail techniques are customers able to reduce their searching costs drastically. Online shopping gave customers more information about products prices, quality, specifications and the stock in the stores. By this innovation are the costs of transport by customers for shopping decreased a lot (van Wee, Annema, & Banister, 2013). Online retail can be done from anywhere, and travelling to stores by customers could not be needed any longer. Everything that's needed by people can be delivered at home nowadays, and so the need for customers to travel for shopping purposes decreased and could even vanish in the future. This development in transport can also have a huge effect on transport as a whole, when individual trips to shops aren't necessary anymore and transport companies can deliver instead all products to all customers as efficient as possible. Combining the trips of delivering companies also together will decrease of the amount of single trips to the stores enormously. The online innovation together with transport efficiency could therefor also have a positive effects to society when negative externalities

caused by many individual trips will be reduced (Garcia-Sierra, van den Bergh, & Miralles-Guasch, 2015).

1.2 Retail substitution from offline to online

The need to physically search for items and goods decreased because of online innovations for shopping and therefor reduced the searching costs for customers (Mokhtarian, 2004). People are able to get their goods much faster, more convenient and mostly cheaper because of these innovations. Online shopping is taking over a lot of the market share from the physical stores and this will only become more dominant in the future (Eurostat, 2017). The impact on transport for society will be enormous when trips to shops are not needed any longer and distribution companies will deliver all products to all customers in an efficient way (Mokhtarian, 2004). The externalities that result from transport could be reduced by the shift to online shopping when people do not have to travel for consumer goods any longer.

However, not everyone implemented online shopping, only a fraction of people buys often online nowadays, and so physical shopping still has the upper hand (CBS, 2017). Most people still buy their products in physical stores until now and so shopping trips are still be made in order to get the products that the people need or want. This implies that shopping is not only done with the purpose of gathering products that provide utility to individuals. Physical shopping still exists and will exist still in the future because people enjoy shopping and like to compare different kinds of goods. This type of shopping also creates utility for individuals where people gain utility by the shopping trip itself and not only from the cheapest option, this is also called leisure shopping or fun shopping (Close & Kukar-Kinney, 2010).

Transportation to stores isn't directly leading to increased utility for individuals because transport is needed to get the goods; therefor is transport a derived demand of the utility gained by shopping. It is this derived demand of the utility gained by shopping that is creating utility for some individuals. Maximization of the individuals utility could be obtained when the searching costs, in terms of time, of shopping decrease by shopping online versus in store (Simon, 1955).

Different types of shopping trips could be valued differently in terms of utility and so differences in valuation could have their effect on travel behaviour.

1.3 Utility gains from products and shopping trips

There is definitely a shift from offline to online shopping purchases, but there are still drives for the people that still shop offline despite the advantages of online retail (CBS, 2017). Online shopping is particularly explainable by the aspect of time that becomes more important nowadays (Lam & Small, 2001). The time that people account for shopping purposes changes and this has its effect on the retail environment (Mokhtarian & Chen, 2004). The shift to online retail and the importance of time do therefor also have their effects on shopping trips and transport. The utility gained from shopping needs to be examined in order to understand what the effect of online shopping is on travel behaviour. The part of utility maximization of shopping will first be investigated in the theoretical background section by comparing literature about this subject.

Next to utility maximization for different types of shopping trips like grocery shopping or fun shopping, is there also the distribution of free time of individuals that is becoming more important. Online shopping has a large impact on the time distribution for shopping purposes of individuals. People do account a certain amount of time for leisure or shopping (van Wee et al., 2002). Individuals could redistribute their time for other purposes when certain shopping trips won't be needed to travel any longer. Is this time used to travel larger distances to get products that are not available nearby? Or are people using their time for other purposes like leisure, or can this time for leisure be founded in physical shopping trips? The research question is formulated by:

> *Do people who shop online more, travel longer distances when they shop in physical stores?*

The time that people spare by online shopping could be used for extra time for physical shopping. This could be the case especially when shopping is considered as a nice experience and is a type of leisure or fun. People account mentally some amount of time for shopping and some shopping trips could therefor not be needed any longer by online shoppers. This would imply that online shoppers

redistribute their time and are left with more time for leisure activities. This redistribution could also lead to people that will travel further when they do go shopping at physical stores.

1.4 Theoretical relevance

This report will have a large contribution to existing literature about transport behaviour but will also be useful for retailers because of the effects of e-commerce in retail and the willingness to travel to stores. Especially the aspect of the effect of online retail on transport behaviour isn't investigated in detail in other literature and therefor will be very interesting for these studies. Other research had already examined for example the effect of online retail on travel behaviour where geographical factors are important determinants. This research will focus on the Netherlands where infrastructure is good and where shops can be found in relative small distance to its consumers. This contribution will have an influence on the way shops are perceived and in which way goods are transported to its consumers. Different shopping trips could be made online and create a time reduction for people. But is this extra time used for other shopping purposes and is the utility gained by shopping an explanation for larger shopping travel distances?

1.5 Structure of the report

This report started with the introduction where the problem is mentioned as well as the research question and the relevance of the report. Secondly, a theoretical background will be followed where the existing literature will be discussed and where theoretical concepts are explained in order to understand the research. Thereafter, the data will be examined that is used for the investigation. Followed by the methodology where the concepts of the research are explained as well as discussed in order to give a representation of the investigation together with the research question of this report. The next section contains the results of the research and explains the most important findings. Finally, the conclusion and discussion about the research with the limitations and recommendations for further research for this thesis will be discussed.

2. Theoretical background

The history of research about online retailing and transport behaviour needs to be discussed in order to get a wide view of the problem and to understand why this research differs from others. Literature about time, travel behaviour and substitution of shopping trips by online retail will be reviewed as well as other literature about the effects of online retailing on transport and travel behaviour. Important definitions like the Generalized Transport Costs (GTC) and the Prospect Theory (PT) are explained and discussed in line with the research, these theoretical concepts and literature contribute to the motivation of this research. The implementation of the concepts from the literature is needed to explain the shift from physical shopping to online shopping and is motivated by the Rational Actor Model and the Generalized Transport Costs theory. Important physiological reasons need to be understood in order to motivate the change of travel behaviour by individuals.

2.1 Rational Actor Model

To investigate the shift from offline to online retail we need to understand some basic ideas behind decision making and travel behaviour. In the beginning we assume that in the ideal world all human individuals act rationally. The assumption is used in most economic models and theories in order to find plausible explanations for situations without too many changing variables. Individuals act rational when their behavioural choices in practice are consistent with their individual goals (Simon, 1955). Some choices by individuals can be rational even when the choices are not morally or ethically defendable.

The Rational Actor Model is based on multiple choices that can be ordered by priority. The decision maker analyses the options for the benefits, for the costs and for their possible outcomes. Losses, profits and likelihood of the multiple options are ranked from the best to the worst option. The rational actor is able to choose the best option with the lowest costs. The optimal choice is linked with rational behaviour and maximizes the individuals their utility.

Individuals do make use of this technique in practice mostly unconsciously in their travel behaviour but also in their shopping behaviour (Simon, 1955). People maximize the utility from a product by minimizing the costs to acquire it.

Transport costs are incorporated into this and so there could be significant evidence that people would like to minimize the distance that they need to travel to get certain kind of goods (Lo, Luo, & Siu, 2006). Especially online retail minimizes transport costs drastically in terms of time and money (Annema, 2013). By implying the rational actor model could there be said that better online retail conditions would decrease the costs of acquiring a product in terms of time and money. The opportunity costs of physical retail are therefor higher and according to the model will people choose the option with the lowest costs and the highest benefits. Online retail decreases therefor the opportunity costs of a good compared to physical retail that implies extra travel costs. The higher opportunity costs by the transport costs for physical retail results in a higher loss, but physical retail can also be compared with higher benefits. The physical stores are able to give their clients a better shopping experience via the personal approach or the experience in the shop and this could be valued as an extra benefit in utility from physical shopping. People could value a shopping experience as extra benefits compared to the benefits by online retail in their rational actor model. The question arises where the distinction is made between physical and online retail. For which people is the shopping experience more important and for which people the costs?

This research tries to find an answer on the question if online retail does have a significant effect on travel behaviour. Are people really travelling less by improvements of online retail? Or do people make certain trips, like grocery shopping, not any longer? Or are people right travelling further to see and buy products which where first not available nearby? Is shopping more then a necessity or could it be that people like shopping as a leisure activity? Do people act via the rational actor model and just minimize their transport costs for shopping or is this not the case and do they want to maximize utility?

2.2 Travel behaviour and prospect theory

Next to the rational actor model is the know-how about travel behaviour needed in order to understand the change of the actual transport behaviour caused by e-commerce. Human behaviour is mostly not in line with the expected behaviour of rational humans of the rational actor model, often noted as the ''homo

economicus'' (Annema, 2013). The behaviour in practice mostly differs from the behaviour of rational individuals modelled by the theory because people are bounded rational in their thinking capacity, available information and time (Annema, 2013). This implicates that travel behaviour mostly cannot be derived from the rational actor model. Not all decisions are focussed on minimizing the costs or maximizing the individual utility. People could value physical shopping for example, as a leisure activity that gives them extra utility.

Decisions that are made in practice are bounded to the limitations and therefor mostly not in line with the expected behaviour modelled by the theory of expected utility. The expected utility theory models the expected outcomes of different decisions with uncertainty. Rational persons should choose the option that is in their best, and so they maximize their expected utility (Rabin et al., 2000). The theory suggests that rational decision makers, with a certain kind of risk preferences, maximize their expected utility without other factors influencing their decision (Rabin et al., 2000).

In multiple transport studies is shown that these assumptions are not met, and that individuals are not acting in a rational way (P. Mokhtarian, 2004) (Lo et al., 2006) (Van De Kaa, 2010) (Rotem-Mindali & Weltevreden, 2013) (Garcia-Sierra et al., 2015). A theory that does account for these irrationalities of human behaviour in travel behaviour is captured by the prospect theory. This theory takes irrational behaviour into account that occurs by the risk of the prospects (Kahneman & Tversky, 1979). Compared to the rational actor model takes this theory also risk into account for the choice options and does it not only accounts the costs and benefits. The rational actor model falls short when the risks of multiple outcomes need to be considered and so the prospect theory could give better forecasts of human decisions with different probabilities.

According to the prospect theory do people tend to underweight outcomes with risk in comparison to outcomes they could get with certainty and results in risk aversion of people (Kahneman & Tversky, 1979). Risk aversion is also visible when the same choice is presented in different ways to people they tend to choose the option where the loss is minimized (Tversky & Kahneman, 1991). People are more likely to choose options that minimize their loss. A loss of the same magnitude as a profit results in decreased utility for the individual because losses are relatively overvalued as profits (Van De Kaa, 2010). The division of gains and losses with different values can be seen in figure 1, where losses are overvalued compared to gains equivalent in size.

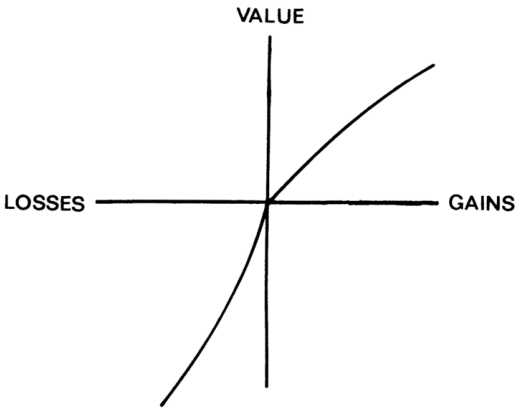

Figure 1: Prospect Theory (Kahneman & Tversky, 1979)

The Prospect Theory is important to understand for the behavioural change in travel patterns of consumers. Changes in travel behaviour could therefor be best explained according to the Prospect Theory. According to the ration actor model are people in search for the highest possible gains in utility but the Prospect Theory also adds that people dislike losses more and that they choose probably the option with the lowest risk for losses. Also travel behaviour is influenced by decisions that can be forecasted by the rational actor model and the prospect theory together. E-commerce could result in changes in travel behaviour because people decrease their travel time and their costs and therefor increase their utility. This could enforce the expectation that people tend to travel less when they shop online more often.

2.3 Generalized Transport Costs

The decision for which way of shopping is used is according to the Rational Actor Model and the Prospect Theory based on costs and benefits. Benefits from shopping are based on the utility that people gain from the goods purchased or the shopping experience. This utility is a nice way to give a value to the gains of shopping but this utility is different for every individual. Utility cannot be

measured and only be ranked between the decisions from best to worst possibility and therefor hard to investigate. However, the costs are made by the product itself but also by all other costs that are needed to make to get the product. These costs are mainly determined by the costs of transportation to the store for physical shopping trips. These costs and distances can be measured and with these numbers it is possible to examine the costs of shopping.

The costs for shopping that can be examined, are modelled by the Generalized Transport Costs model (GTC). This model assumes that people act according to the rational actor model where they want to minimize their costs, so also their costs for transport. Assumptions of the prospect theory are also taken into account but these are difficult to model and so the rational actor model is used for this case. The costs for travel are not only the result of the out of the pocket costs for the transport service but also the out of the pocket costs of travel time (Hanssen, Mathisen, & Jørgensen, 2012). The Generalized Transport Costs (GTC) can therefor be defined by:

$$GTC(D) = P(D) + HT(D)$$

Where P is the price for the transport service dependent from the distance D travelled in kilometres. For shopping could this be the costs of fuel and the use of the car. Next to that are the time costs per hour, H and the trip time, T.

The time and distance that people travel for a trip can also be expressed in a monetary value as seen by the Generalized Transport Costs. Many other researchers tried to find an answer on the question of what the monetary value of only time is. Literature about this subject showed multiple ways to research this question. This question could be answered by the theory about the Travel Time Budget. The uncertain aspect of transportation that can be deducted from the Prospect Theory does contribute to this theory and therefor important to understand for the value of time.

Studies showed that the reliability of travel time is very important in the route choice behaviour because of the risk averse behaviour of travellers (Lo et al., 2006). Earlier experiences of trips are also having a significant effect on the variability of travel times (Lo et al., 2006). This effect also refers to the Prospect Theory (Tversky & Kahneman, 1991). The travel time budget of individuals is

heavily influenced by the uncertainty by external factors within transportation. The time that an individual account for the trip is mentioned as the travel time budget, and can be defined as:

[Travel time budget] = [Expected travel time] + [Travel time margin]

The travel time budget can be seen as the time that people take into account for a trip with their personal travel time margin included. For every individual is this travel time budget different because all trips have different purposes. Some trips could be made earlier in time and so people know from their reference point how long the trip takes. The punctuality of individuals is having a large impact on the travel time margin and this also differs for all individuals (van Wee et al., 2013).

The value of time was in one case investigated by measuring the reliability of roads with different levels of congestion or toll routes (Lam & Small, 2001). The researchers succeeded to give a value of time and reliability of travelling with their pricing experiment. Others argued that the travel time budget is not constant and that it is strongly related to individual and household characteristics. The characteristics about the destination and area are important factors that influence the travel time budget (Mokhtarian & Chen, 2004). This research showed that travel time cannot be given one value because it's not constant over all individuals (van Wee, Rietveld, & Meurs, 2006). The Travel Time Budget is a good interpretation of the monetary value of time but isn't a good predictor for the costs of shopping. The rational actor model can therefor best be used, where all individuals value their utility and costs different.

It seems reasonable that in most cases the Generalized Transport Costs are higher for shopping trips to physical stores compared to people that shop online. Online shoppers do not encounter the costs of the transport service. Also the time needed for the trip to the shopping centre makes online retail much more attractive from this point view. However, people still shop offline and so some people do gain utility from physical shopping and value the costs of transport lower. Different individuals do therefor value the physical shopping different as shopping online. The shopping experience, or maybe the aspect of leisure or fun

obtained by shopping physically could be an important explanation for other utility obtained from shopping.

The question arises when the utility obtained from shopping is higher as the generalized transport costs. This function could be expressed by:

$$U(Shopping) > GTC(D)$$

This interesting fact is not only dependent from the distance for shopping but also the utility obtained from shopping. Utility is not measurable by itself and so difficult to give a number to and therefor can it also not given a monetary value. However, utility can be ranked by the revealed preference of individuals. The different utilities for certain products or activities can be revealed in the different preferences that every individual has. Individuals can rank different options and this shows which option they prefer compared to other options. This can also be used for the valuation of shopping. Where some people value the utility for shopping higher as others we can measure this by the revealed preference of these individuals. This shows that some people have a higher willingness to travel to stores as others.

Travelling with shopping purposes is one of the characteristics of the difference in generalized transport costs over individuals. Some people value shopping more than others and therefor account less value to the travelling time needed for shopping purposes. Not only the value of time to go to shops can be investigated, also the value of time can be investigated of travelling between different shops. The value of time used for shopping between different shops to compare prices, quality and other factors of different goods can be used to value time for shopping purposes (Marmorstein, Grewal, & Fishe, 1992). The researchers concluded that some people receive enjoyment of some types shopping. Shopping for leisure or a fun experience like comparing different stores could give utility for some people. Where others would dislike grocery shopping and so would encounter a negative utility from grocery shopping. The investigation argued that fun shopping like visiting a shopping mall could give positive utility to some individuals that outweigh the costs of travelling. The experiment that investigated this also confirmed this expectation and concluded that travel time can also be subjective (Marmorstein et al., 1992). This would

mean that personal characteristics are very important for the valuation of shopping utility. Different types of shopping could have different valuations of utility and so this should be further investigated.

2.4 GTC for online and physical retail

The valuation of shopping utility and the spatial geography or accessibility of retail stores are the most important factors that influence people to buy online or in store (Farag, Schwanen, Dijst, & Faber, 2007). Literature about the effect of e-commerce on transport is very wide and the findings are very diverse. The results of these investigations are fundamental to understand for this research about the impact of e-commerce on travel behaviour by customers. The findings of the literature will be discussed therefor in more detail in this section.

Shopping mobility was in the early years mainly affected by personal characteristics like income level. People with higher incomes could incur the extra costs of travelling to other regions in order to get a more varied selection of products (Herrmann & Beik, 1968). Those rich people did mainly live out of the large urban city centres but travelled to these city centres and so travel distance for shopping was mainly determined by personal characteristics (Herrmann & Beik, 1968).

Accessibility of retail stores is one of the most important factors that influence peoples buying behaviour via online or physical stores. When the physical retail accessibility is bad would this result in longer travel distances by consumers and more online retail in these areas (Visser & Lanzendorf, 2004). The increasing amount of e-commerce results in a shift of the distribution centres from the large cities to more remote areas (Visser & Lanzendorf, 2004). This would mean that sales via e-commerce can be delivered more convenient and quick and therefor would reinforce the closure of physical stores in these remote areas. The accessibility of retail shops will decrease and people would need to travel further if they still want to shop at physical retail stores (Visser & Lanzendorf, 2004).

Other research about this subject concludes controversially that increasing e-commerce would not affect travel distance. Online stores will compete with physical retail stores but will share the market for retail products. Customers

will buy from online stores as well as the physical stores. The travel distance for shopping purposes would therefor not reduce (Mokhtarian, 2004).

The relationship between online and in-store shopping is interesting for the amount of travelled kilometres. Travel to retail stores will definitely decrease when online stores are dominating the market and retail stores will hardly exist. However, Research showed that online retail shopping has a positive effect on the frequency of shopping trips to physical retail stores (Farag et al., 2007). But the risen frequency of shopping trips also has a positive effect on the amount of online purchases (Farag et al., 2007). The question arises which kind of shopping (grocery or fun) has a positive effect on the other. The research also shows that urban residents shop more online as rural residents despite the high accessibility of stores (Farag et al., 2007). However, inhabitants of places with a lot of stores in reach seemed to search online less often as people with low accessibility to stores (Farag et al., 2007).

Others conclude that the relationship between online and in-store shopping is difficult to forecast and that transition is not only determined by cost minimization (Rotem-Mindali & Salomon, 2007). The change of shopping behaviour is difficult to examine because there are lots of factors like the level of technology that could influence this (Rotem-Mindali & Salomon, 2007). Research concluded that online retailers and physical retailers are living in a symbiosis where they strengthen each other (Lee, Sener, Mokhtarian, & Handy, 2017). This could answer the question not only more retail online has a positive effect on physical retail but also the other way around. People can buy and watch at the Internet as well as in the physical store and this results in higher shopping frequencies, and so more shopping trips (Lee et al., 2017). The results of the researchers is still in line with recent numbers about retail, where in-store shopping still has the upper hand (CBS, 2017).

Most literature that is discussed until now was mostly about accessibility of retail stores in urban areas. Research about this subject is also executed in a situation where accessibility of retail stores is bad. The researchers examined the impact of the e-commerce on consumer mobility in the Scottish islands. The circumstances with limited transport infrastructure, inconsistent supply and restricted availability of the supply of goods is clearly visible on these islands.

This research showed that the impact of e-commerce on consumer travel is only small (Calderwood & Freathy, 2014).

Internet usage and online retailing is rising rapidly over the last decades, but is used different over the world. Personal characteristics like age and income seem to have a significant effect on the frequency usage of online retail. This research also showed evidence for the case of England that people in more rural areas adapted online retail more often (Clarke, Thompson, & Birkin, 2015). This report is in contrast with most others where it concludes that people use online retail more when accessibility of retail stores is bad (Clarke et al., 2015).

2.5 Online impact on travel behaviour

Shopping became much more convenient by the innovations of online shopping at the Internet stores. Consumer reduce their time and effort by the risen convenience levels due to the online shopping services (Jiang, Yang, & Jun, 2013). The value of time is more important as ever nowadays in a very busy world. Time saving options are therefor not only more convenient but also increase utility from shopping for customers (Jiang et al., 2013). The online stores decreased the effort that is needed to collect products for consumers because products can be delivered at home. People do not have to travel to stores any longer in order to get the products they want. Online shopping is in these two aspects more convenient as physical stores. However, there are still many shops and lots of people travel to go to stores that like. Shopping in store therefore seems not yet to be overruled by online shopping but competition for retail shops is getting stronger (CBS, 2017).

People are able to compare quality, prices and availability of products very easy online and it saves them a lot of effort (Park & Kim, 2003). Instead of searching for the right product at the right store and travel between or to these shops costs them time, money and effort. This trend is only becoming stronger by innovations of web stores. Computers will know your personal taste and will advise customers other products that are similar to the products you already picked out.

The shopping mode choice between physical stores and online stores is mainly set by the valuation of time by different people (Hsiao, 2009). Some do not mind

to travel one hour to visit some stores while others think they could do much better things in the time that they do not need to travel. This valuation of travel time also differs for different kind of shopping trips (Hsiao, 2009). Travelling to the stores as well as travelling between stores is value as very costly in terms of time by lots of individuals (Overby & Lee, 2006). Online stores do have a competitive advantage in the value of the time for their customers and physical stores have an advantage in the shopping experience that they can offer (Childers, Carr, Peck, & Carson, 2001). Especially the factor time is the main driver for customers to choose for online shopping. The reduction of time needed to get products and get them delivered at home results in positive utility for the individual (Hsiao, 2009). Travelling to physical stores therefor does encounter costs of time but also of monetary value. Customers that bought products online value their products higher in comparison with the same product only bought in a physical store for the same price (Close & Kukar-Kinney, 2010). The costs that they encounter from travelling do have an effect on perceived customer value.

3. Data and methodology

3.1 Dataset

This thesis makes use of a combined dataset with data retrieved from two knowledge institutes in the Netherlands. The total dataset has 116,298 observations about trips from people in the over the years 2013 and 2014.

3.1.1 MPN

Data from the ministry of infrastructure and water is used to get insight about the movements of people with different characteristics in the Netherlands. The knowledge institute ''Mobiliteitspanel Nederland'' (MPN) retrieves this data from many households, which results in 116,298 observations of trips over two years. The institute sends surveys to their respondents with multiple questions about their travel behaviour and personal characteristics. The institute is able to investigate behavioural changes in transport with the data that is retrieved from the surveys. This report makes use of the dataset of MPN for the years 2013 and 2014. My thesis supervisor J.J. Witte shared the data for these years, which could only be used for educational purposes like this investigation.

3.1.2 CBS

Data from the Central Bureau of Statistics (CBS) in the Netherlands is also needed for this thesis to investigate the travel behaviour in more detail. Data from this institute about postal codes and neighbourhoods is used to add to the dataset of MPN. The combination of these two datasets is giving more precise information about the origin destination as well as the arrival destination of the people survived by MPN. This level of detail is very important for the implications of this thesis, especially for the data of the Netherlands where travel distances are relatively small.

3.1.3 Household clusters

The data retrieved from the dataset is clustered by households and are separated via household identity numbers. The dataset thus consists of all trips made by individuals but some individuals or households recorded more then 1 trip. Trips are therefor clustered by household identities and the models need to be

corrected for this in order to prevent biased standard errors. The models that predict travel distance for shopping trips are corrected with the ''cluster'' expression in STATA.

3.1.4 Sorting of the dataset

Not all observations of the dataset can be used for the investigation of shopping trips. The dataset consists of all trips that are monitored by households. This also results in multiple trips for the same person in one year. The dataset should be corrected for this bias where some people recorded lots of small trips because they live in an urban area, where others only filled in one shopping trip. The dataset should therefor be corrected where only one shopping trip per year is used for the investigation. The dependent variable that is investigated is shopping distance and so this research will make use of the longest trip of every individual per year and not all of its recorded shopping trips. All other shopping trips of an individual that recorded multiple shopping trips are rejected. This has a large impact on the dataset but there are still many observations left and so the investigation can go on. This investigation will be based on this dependent variable and is called ''SHOP_LANGSTE''

3.2 Methodology

Investigation can start now the whole dataset is complete and the data is sorted and corrected. This research is executed by the use of the computer program STATA. The following commands and results are executed by the use of this program. The quantitative data is investigated using Ordinary Least Squares models. The models are used to forecast the travel distance of online and offline shoppers.

3.2.1 Dependent variable: Shopping distance

A variable should be chosen or created that can investigate empirically what the effect of online shopping is on travel distance. This variable is then set up in an Ordinary Least Squares model that can test this effect. Travel distance for shopping should therefor be modelled as a function of online shopping behaviour. The travel distance for shopping purposes has to be taken out of the data by making a new variable called ''SHOPDISTANCE''. This variable is made by

the variable, where an individual says the trip recorded was used for shopping purposes, multiplied with the distance travelled for the trip that day. This dummy variable multiplied with a continuous variable of distance results in all distances that where made for shopping purposes. SHOPDISTANCE is the distance that is travelled for shopping and covers the recorded shopping trips. This data is also corrected for the households, only the longest shopping trips per year were used when households filled in the distance for multiple shopping trips. It still results in a large dataset of 36,556 observations from the original dataset of 116,298 respondents to investigate the travel distance for shopping purposes. The single shopping trips that are made vary from 0 to 400 kilometres. The mean is around 2.12 kilometres and implies that many trips are recorded for short distances. These descriptive statistics of the dependent variable are shown below.

Variable	Obs	Mean	Std. Dev.	Min	Max
SHOPDISTANCE	36,556	2.119611	11.87101	0	400

3.2.2 Online shopping frequency

Next, the variable for the online shopping behaviour needs to be taken into account. This variable is deducted from the question of how many times an individual uses the Internet for online shopping purposes. This question is divided in eight answer possibilities respectively:

- 4 days or more a week (option 1)
- 1 to 3 days a week (option 2)
- 1 to 3 days a month (option 3)
- 1 to 2 days per quarter a year (option 4)
- Less then 1 day per quarter a year (option 5)
- Not asked, person (almost) does not use
- Not asked, person is younger then 12 years old
- Person did not give an answer to the question

The observations deducted from this question with their frequencies are shown below. Other answer possibilities are excluded from the dataset and are interpreted as missing values. Still 42,213 observations where left in the dataset.

Use of the Internet for buying products, Services or online shopping	Freq.	Percent	Cum.
4 days or more a week	832	1.97	1.97
1 to 3 days a week	4,383	10.38	12.35
1 to 3 days a month	15,956	37.80	50.15
1 to 2 days per quarter	11,313	26.80	76.95
Less then 1 day per quarter or not	9,729	23.05	100.00
Total	42,213	100.00	

The question could be answered by five different answer possibilities and so this variable needs to be changed in order to add this variable into the ordinary least squares model. The variable of online frequency isn't binary because the answer possibilities are not zero or one and the variable is also not continuous because of the limited choice possibilities. This variable is called a limited dependent variable where there are only five answer possibilities. The choice possibilities of the limited dependent variable are not equal in weight and so the ranks of the choice possibilities are not equal to each other. I created five different binary dummy variables to tackle this problem of limited dependent variables. Where the first dummy variable called ''ONLINEFREQ1'' is the answer possibility where people use Internet more than four times a week to shop online, the answer to this question can only be filled in with a one for yes and a zero for no. ''ONLINEFREQ2'' is the second dummy variable where a one stand for Internet usage of one to three days a week for online shopping. The other three answer possibilities are also transformed into dummy variables respectively.

3.2.3 Urbanity

Next to the online shopping frequency are personal characteristics very important according to multiple researchers and described in the theoretical background section (Farag et al., 2007). The literature showed that especially urbanity has a significant effect on travel distance for shopping and so this variable should be included in the model as a control variable (Herrmann & Beik, 1968). The urbanity of a person can be found in the large dataset where people could choose between 5 types of urbanity. These types differ from very urban cities to very rural areas in the Netherlands. Where very urban municipalities are classified to have more than 2,500 addresses per square kilometre and very rural municipalities would have less then 500 addresses per square kilometre. The frequencies of urbanity for the dataset are shown in the figure below.

Urbanity (municpality level).	Freq.	Percent	Cum.
Very strong urban (2500 or more inhabita	7,635	18.09	18.09
Very urban (1500 to 2500 inhabitants)	12,110	28.69	46.77
Moderately urban (1000 to 1500 inhabitan	9,540	22.60	69.37
Low urban (500 to 1000 inhabitants)	8,671	20.54	89.92
Not urban(less then 500 inhabitants)	4,257	10.08	100.00
Total	42,213	100.00	

3.2.4 Gender

The gender variable is another important personal characteristic that is spoken of in the literature (Herrmann & Beik, 1968). Maybe the most visible differences in shopping behaviour can be found between men and woman. From a subjective point of view I assume that women attach mostly more value to shopping as men. Travelling for shopping purposes could therefor create more utility for woman as for men. The increased utility could be a reason for larger travel distances for woman as for men. The variable for gender therefor could have a significant effect on travel distance and therefor should be included in the model. The frequencies of the gender are listed below, where the total number of respondents is 42,213 corrected for cluster and multiple trips per person.

Gender	Freq.	Percent	Cum.
Male	21,864	51.79	51.79
Female	20,349	48.21	100.00
Total	42,213	100.00	

3.2.5 Income

The final factor that needs to be fitted into the model is another variable that is frequently spoken of in the literature (Visser & Lanzendorf, 2004). An important personal characteristic, especially for shopping, is the variable for income. The dataset also provides information from the respondents about income. Respondents could answer a multiple-choice question with six different answer possibilities. A seventh possibility was where people did not respond to the question. Dummy variables need to be created in order to use the instrumental variable of income in the model. The question could be answered by the following answer possibilities of income in euros per year:

- <12.500 euro per year (option 1)
- 12.500-<26.200 euro per year (option 2)
- 26.200-<38.800 euro per year (option 3)

- 38.800-<65.000 euro per year (option 4)
- 65.000-<77.500 euro per year (option 5)
- >=77.500 euro per year (option 6)
- Unknown (option 7)

The frequencies of the respondent answers are shown below. Every possibility is transformed in a dummy variable where the first possibility of an income below 12.500 euro a year is named ''INC1''. The second possibility of an income between 12.500 and 26.200 euros per year is named ''INC2'', the other answer possibilities are also transformed in the same way to dummy variables. The other possibility needs to be dropped from the dataset so there are six different dummy variables created with a total of 42,213 observations.

Gross annual income	Freq.	Percent	Cum.
minimum (<12.500)	1,971	4.67	4.67
below average (12.500-<26.200)	5,999	14.21	18.88
average (26.200-<38.800)	9,387	22.24	41.12
1-2x average (38.800-<65.000)	15,094	35.76	76.87
2x average (65.000-<77.500	4,276	10.13	87.00
more then 2x average(>=77.500)	5,486	13.00	100.00
Total	42,213	100.00	

3.3 OLS Models

There need to be made some changes to the dataset to be able to investigate the research question. Estimated Ordinary Least Squares models are created by the use of the dataset in order to give an answer to the research question.

3.3.1 Model 1

The first model is created by the use of multiple important personal characteristics. These characteristics do have an effect on travel distance according to the literature and therefor need to be modelled. The model also takes into account online retail frequency and is therefor used as the base model of this research for model 2 and three.

3.3.2 Model 2

This research makes a distinction between shopping trips that are made because of necessity or shopping trips that are made with the purpose of a shopping trip as a leisure activity.

The distinction between fun shopping is difficult to determine exactly and heavily depends on the personal characteristics of the individual. However, some researchers tried to make the distinction based on multiple reports and investigations. The differences between the types of shopping may result in higher willingness to pay for travel for leisure/fun shopping trips. Research about higher-level shopping centres was able to investigate from where people came. There was found that the local convenience centres mainly attracted people that travelled up to 30 minutes (Jones & Simmons, 1990). The higher-level shopping centres that where compared with fun shopping attracted significantly more people that travelled more than 30 minutes (Jones & Simmons, 1990).

The variable that is created to make the distinction between fun shopping and shopping for necessities is named ''FUN''. The distinction is made by making use of the variable for travel time according to the research about convenience and higher-level shopping centres (Jones & Simmons, 1990). The variable that is created filters out all trips that are less then 30 minutes to travel and gives them the value 0. Trips longer then 30 minutes are given the value of 1. This dummy variable is strongly correlated with the independent variable of travel distance but is created to make the distinction between shopping trips with different purposes.

The distinction of different kind of shopping trips can be investigated by including this variable to a standard model with personal characteristics of shoppers. By comparing this model with the standard model with personal characteristics we could investigate the research question.

3.3.3 Model 3

The difference of shopping distance between online and offline shoppers can best be investigated by the use of the fun-shopping variable. This difference can better be distinguished by making use of an interaction term that consists of the online frequency and the element of different shopping trips. Model three has the interaction terms of these variables therefor incorporated into the model and is so able to give a better forecast about the shopping distance.

4. Results

The different variables discussed in the last section are merged together into three ordinary least squares models. The dependent variable of the longest distance that people travel for physical shopping purposes is regressed to the multiple independent variables. Three models are created and compared to each other in order to investigate the effect of online shopping to shopping travel distance. The models are shown below and discussed in more detail in the following sections.

4.1 Quantitative Analysis

	(1)	(2)	(3)
	SHOP_LANGSTE	SHOP_LANGSTE	SHOP_LANGSTE
MALE	-0.884***	-1.016***	-1.010***
	(-3.80)	(-4.23)	(-4.22)
ONLINEFREQ1	0.686	0.542	0.267
	(1.03)	(0.81)	(0.84)
ONLINEFREQ2	0.340	0.119	0.518**
	(1.11)	(0.36)	(3.25)
ONLINEFREQ3	0.581	0.443	0.284**
	(1.96)	(1.45)	(2.58)
ONLINEFREQ4	0.682	0.634	0.192*
	(1.79)	(1.67)	(2.04)
URB1	-0.655	-0.836	-0.856
	(-0.52)	(-0.66)	(-0.68)
URB2	-1.092	-1.072	-1.079
	(-0.90)	(-0.89)	(-0.90)
URB3	-0.620	-0.479	-0.482
	(-0.49)	(-0.39)	(-0.39)
URB4	-1.586	-1.491	-1.496
	(-1.31)	(-1.25)	(-1.26)
INC2	0.877	1.007	1.034
	(1.67)	(1.87)	(1.91)
INC3	0.868	1.016*	1.049*
	(1.76)	(1.98)	(2.04)
INC4	0.669	0.836	0.856
	(1.35)	(1.59)	(1.62)
INC5	0.480	0.593	0.602
	(0.97)	(1.16)	(1.18)
INC6	1.198	1.242	1.256
	(1.86)	(1.90)	(1.92)
FUN		2.643***	2.190**
		(5.61)	(3.13)
FUNONLINE1			0.778
			(0.47)
FUNONLINE2			-0.855

			(-1.02)
FUNONLINE3			0.482
			(0.59)
FUNONLINE4			1.285
			(1.22)
_cons	2.264	1.296	1.428
	(1.88)	(1.17)	(1.33)
N	36556	36556	36556

t statistics in parentheses
* p<0.05, ** p<0.01, *** p<0.001

Some dependent variables are having clearly a significant effect where others do not have as significant effect at all. Nevertheless, we can have reasonable expectations and forecasts about the travel behaviour that is affected by online retail. The models are estimated with the data of 36,556 individual observations. The forecasted variables are shown for each model and can so be compared to each other. The interaction term in model 3 is an important addition of the other two models and the differences for fun shopping and online frequency are therefor shown in table 1 below. The distances for non-fun shopping are the distances found by model 3 for the different online frequency dummy variables, respectively. The values are calculated for a woman that lives in a very rural area with a very low income, all these values take the value zero. This situation is a base situation for the model where other formats result in other values for the travel distance. Together with the constant term we make the distinction between fun-shopping and non-fun shopping. The differences for fun shopping and non-fun shopping can be related to the interaction term values. Where the fun shopping values are calculated by adding up the constant factor and the fun factor with the different online frequency dummy variables and fun online dummy variables. The differences are very different; this can possibly be explained by the insignificant values of the interaction term dummy variables that are used for this calculation. Also the fun variable is correlated with shopping distance because of the interpretation and could therefor give a high value to the distance travelled for shopping. However, the numbers show clearly the difference between fun shopping and non-fun shopping where the distance travelled for fun shopping is larger.

Table 1: Estimation of total travel distance for shopping purposes

Online frequency	Fun shopping	Non-fun shopping
1 (more than 4 times a week)	4.663 km	1.695 km
2 (1 to 3 times a week online)	3.281 km	1.946 km
3 (once to 3 times a month)	4.384 km	1.712 km
4 (once or twice per three months)	5.093 km	1.620 km
5 (never online)	3.618 km	1.428 km

4.2 Gender

The control variable for gender has a significant effect on the travel distance for shopping for all of the models. The results of -0.884, -1.016 and -1.010 kilometres show that men are significantly travelling on average about 1 kilometre less for their longest shopping trip compared to woman, where all other variables are held constant. When all other variables are held constant is also referenced as the ceteris paribus condition, this term will be used for the results of the other variables. The gender effect is significant for t-statistic probabilities of even 0.1% and so this conclusion can be given almost every time. This model therefor implicates that men travel on average less in terms of distance for their shopping trips as women.

4.3 Online frequency

The online shopping frequency is the dependent variable that motivated this research. Unfortunately, not all dummy variables for online shopping frequency are significant at a significance level of 5% and the values for some variables can therefor not be interpreted from the model. An F-test of all dummy variables used for online shopping frequency implicates that the overall significance of the dummy variables is significant for a value of 1%. Some values may not correspondent with the dataset but the model is able to say something about the relationship between online shopping frequency and travel distance for shopping purposes. The variable of ''ONLINEFREQ5'' is not taken into the model because of the dummy variable trap. The variable suggests that people do (almost) never shop online. The other variables for online shopping frequency are compared to this situation of the variable where people do not shop online.

```
. test ONLINEFREQ1 ONLINEFREQ2 ONLINEFREQ3 ONLINEFREQ4

 ( 1)  ONLINEFREQ1 = 0
 ( 2)  ONLINEFREQ2 = 0
 ( 3)  ONLINEFREQ3 = 0
 ( 4)  ONLINEFREQ4 = 0

       F(  4,  3156) =    3.32
            Prob > F =    0.0101
```

The ''ONLINEFREQ1'' variable implicates that people that shop more than four times a week online are travelling 0.686 (model 1), 0.524 (model 2) and 0.267 (model 3) kilometres further as people that do not shop online, ceteris paribus. This result is as said not significant for all models at a significance level of 5% and therefor can this value not be interpreted from the model. The variable ''ONLINEFREQ2'' where people shop online 1 to 3 times a week would suggest that these people will travel 0.340(1), 0.119(2) and 0.518**(3) kilometres further, ceteris paribus, when they shop in physical shops. This variable is only significant in model 3 is the variable significant at a level of 1%, this implicates that people that shop 1 to 3 times a week online travel 0.518 kilometres further as people that do not shop online, ceteris paribus. ''ONLINEFREQ3'' is also not significant in all models but suggests that people that shop online once to three times a month would travel 0.581(1), 0.443(2) and 0.284**(3) kilometres further when they shop in real life compared to people that do not shop online, ceteris paribus. For model 3 is this result again the only one that is significant at a 1% significance level and so there can be concluded that these people travel on average 0.284 kilometres further as people that do not shop, online ceteris paribus. ''ONLINEFREQ4'' is also not significant for model 1 and 2 but again for model 3 it is. Model 3 concludes that people that shop online once or twice per three months travel 0.192*(3) kilometres further as people that do not shop online, ceteris paribus.

Most of the results are not significant of the models 1 and 2. However, most results of model 3 are significant and the results of these dummy variables can be used to give a reasonable forecast about the relationship between online usage and travel distance to stores. Where only ''ONLINEFREQ1'' is not in line with the results and also not significant there could be concluded that there is a positive relation between online shopping frequency and travel distance. Where

people with a low frequency travel 0.192* kilometres further as people that do not shop online at all. People that shop online more travel 0.284** kilometres more and again people that even shop more frequent online travel 0.518 ** kilometres more. From this rising number of distances related to the online frequency can be concluded that people that shop more online also travel longer for physical shopping. However, the result is only a small distance where people that shop online often travel on average 518 metres more as people that do not shop online.

4.4 Urbanity

The model shows not one significant outcome for the explanatory variable of urbanity on shopping travel distance and. An F-test however confirmed the joint significance of the dummy variables made for the variable of urbanity. The results of this variable are hard to interpret because of the insignificant values but the conclusions can carefully by interpreted. The values are not varying a lot for their effect on travel distance and big differences won't be the result of this variable.

```
. test URB1 URB2 URB3 URB4

 ( 1)  URB1 = 0
 ( 2)  URB2 = 0
 ( 3)  URB3 = 0
 ( 4)  URB4 = 0

       F(  4,  3156) =    2.99
            Prob > F =    0.0177
```

The dummy variables that are made are compared to the situation where less then 500 inhabitants per square kilometre live (URB5); this very rural situation is used as the base level.

The variable ''URB1'' is used for strong urban areas with more then 2,500 inhabitant per square kilometre live. People that live in these very dense areas are travelling -0.655(1), -0.836(2) and -0.856(3) kilometres on average, ceteris paribus, compared to people that live in very rural areas. People that live in urban areas with 1500 to 2500 inhabitants per square kilometre, mentioned by ''URB2'', are likely to travel even -1.092(1), -1.072(2) and -1.079(3) kilometres on average as people in very rural areas, ceteris paribus. People that live in areas with 1500 to 1000 inhabitants per square kilometre, variable ''URB3'', are likely to travel -0.620(1), -0.479(2) and -0.482 kilometres on average as people in very

rural areas, ceteris paribus. The last dummy variable created is ''URB4'', for areas that have inhabitants with 1000 to 500 people per square kilometre. These are also likely to travel less kilometres compared to people in very rural areas, ceteris paribus. The results for this respectively: -1.586(1), -1.491(2) and -1.496(3) kilometres on average.

A reasonable conclusion can be given about the relationship between urbanity and travel distance for shopping because of the high number of observations together with the F-test significance. The results show that there is clearly a negative effect on the distance travelled for shopping. Especially people that live in rural areas seem to travel less on average compared to people that live in very rural areas. This seems contradictory to the literature where people in the city have better accessibility to facilities (Jones & Simmons, 1990)(Annema, 2013) and so the need to travel is lower. However, others just mention that people in the city could have other preferences and so they like to travel more in order to have items that are not sold in the city. A negative relationship can be seen but the differences are still very small. Concluding remarks are therefor hard to make and so real differences cannot be deducted from these results.

4.5 Income

Other results are based on the factor income of the participants of the surveys. This personal characteristic seemed from the literature to be very impressionable for the travel distance. However, this model only shows small (mostly not significant) differences for the travel distances influenced by the factor income. The F-test shows that the combined variables are not significant and so we have to be careful with conclusions about the models .Not all results about income can be interpreted, but the overall relationship between income and travel distance could be distracted from the results.

```
. test INC2 INC3 INC4 INC5 INC6

 ( 1)  INC2 = 0
 ( 2)  INC3 = 0
 ( 3)  INC4 = 0
 ( 4)  INC5 = 0
 ( 5)  INC6 = 0

       F(  5,  3156) =    1.22
            Prob > F =    0.2947
```

The variable of ''INC1'' is left out of the model, all other variables are compared to this situation where income is below 12,500 euro per year. The variable of ''INC2'' for people with an income between 12.500 en 26.200 euro per year, resulted in an outcome where these people travel 0.877(1), 1.007(2) and 1.034(3) kilometres more on average for shopping purposes, ceteris paribus. These values are not significant at as significance level of 5%. ''INC3'', for people with an income between 26.200 and 38.800, shows an positive effect of 0.868(1), 1.016*(2) and 1.049*(3) kilometres on average travelled more for shopping purposes compared to people with an income below 12.500 euro per year, ceteris paribus. This effect can be interpreted at a 5% significance level for the models 2 and 3 and so these values can be used in this investigation. For the ''INC4'' variable of people with an income between 38.800 and 65.000 euro per year is found that they travel 0.669(1), 0.836(2) and 0.856(3) kilometres more on average as people with a low income, ceteris paribus. These results are again insignificant for all models at a significance level of 5%. Furthermore, the ''INC5'' group with people with an income between 65.000 and 77.500 euro per year was found that they travel 0.480(1), 0.593(2) and 0.602(3) kilometres on average more as people with an income below 12.500 euro per year, ceteris paribus. This variable does not show significant values. Finally the highest incomes are classified by the variable ''INC6''. The results are also not significant at a significance level of 0.05. The variables shows that these people travel 1.198(1), 1.242(2) and 1.256(3) kilometres on average more for shopping as people with a low income, ceteris paribus.

The effect of income on travel distance can be seen as positive, where people with more income as the base level travel significantly more on average when they travel for shopping. The differences are however small and this variable cannot be interpreted as a big game changer for the distance travelled for shopping purposes

4.6 Fun shopping

The variable for leisure shopping is noted as ''FUN'' and is measured as the distance travelled for shopping by people that travelled more then 30 minutes. This variable is very significant with a significance level of 0.001 and a value of

2.643 kilometres more on average for model 2 and 2.190 kilometres more on average for model 3, ceteris paribus. Distance and travel time are highly dependent from each other and therefor highly correlated. This results in very significant values, purposes of the shopping trip where not recorded in the dataset and so the variable of time was used to classify people in the fun shopping category or not.

4.7 Fun shopping and online frequency interaction

The interaction terms are modelled in the last model to investigate the relationship of online frequency and fun shopping on travel distance better. However, there are no significant values found by the model. Also an F-test for overall significance for the interaction variables shows no significant result. The values found in model 3 for the interactions terms can therefor not be interpreted any further.

```
. test FUNONLINE1 FUNONLINE2 FUNONLINE3 FUNONLINE4

 ( 1)  FUNONLINE1 = 0
 ( 2)  FUNONLINE2 = 0
 ( 3)  FUNONLINE3 = 0
 ( 4)  FUNONLINE4 = 0

       F(  4,  3156) =    1.44
            Prob > F =    0.2196
```

The results of the coefficients show positive values except for the variable FUNONLINE2, which seems to be strange. The interaction terms together with the online frequency dummy variables results in the table that is created at section 4.1. With some precaution can be said that people that are online a lot (more then 4 times a week) and people that are very often online (once to twice per three months) are travelling further for fun shopping activities on average as others, ceteris paribus.

5. Conclusion, Discussion and limitations

5.1 Conclusion

Online retail has changed the way of shopping forever. The effects for retail are clearly visible. Customers are able to compare their products for quality and price very easily and shopping trips to the stores and city centres could probably vanish. However, online retail did not take over physical retail yet and so the question arises if shopping trips to the stores will still be there or that these trips will change and people will travel less for shopping. Online retail is cheaper and possibly more convenient and could therefor decrease the travel to the stores. However, some people do gain utility from shopping because they like it to compare and visit the stores. This utility function could be different for different kind of products and utility from shopping trips should be investigated. This thesis investigates this topic by setting up the research question:

> *Do people who shop online more, travel longer distances when they shop in physical stores?*

First can be concluded according to the theoretical background of this study that all people value time differently. The Generalized Transport Costs are mostly equal where the utility gained by a person can be very different for each individual. For some is the gain in utility by physical shopping higher as for the other and so here could investigation show the revealed preference for physical shopping for individuals.

Furthermore is the quantitative research about travel distance and online frequency use in the Netherlands investigated. The investigation showed that people that shop online more often also travel longer distances when they do go shopping at physical stores. Compared to the situation where people do not shop online at all is there a difference of approximately 1 kilometre. It is possible to conclude that online frequency does have an impact on the travel distance based on the large dataset and the significant values found in the analysis. This conclusion is contradictory for some literature found about the subject, where was concluded that people would travel less kilometres when they shop online more often (Clarke et al., 2015).

Other characteristic findings in the investigation showed similar results as found in the literature about the effect of personal characteristics on travel distance (Herrmann & Beik, 1968). The results showed that men do significantly travel about 1 kilometre less then women when they visit shops. Also the online frequency of people has a significant effect on the travel distance where people that use the Internet more are travelling about 0.5 kilometres more then people that do not use the Internet at all. The urbanity of the area where people live was another variable that was investigated but showed negative outcomes. However, the results of this variable where insignificant and could therefor not be interpreted. Income was proven to be a very important determinant for travel distance. Not all results showed a significant effect but there can be concluded that rich people travel further than people with very low incomes. Other variables that where tested showed no important significant results.

5.2 Discussion and limitations

Online retail has an effect on the amount of purchases made in physical stores and so possibly also in the amount of trips and the distance travelled to these stores. The omitted variable bias is a large limitation of this investigation. The given dataset that was used did not provided enough detailed information for this investigation. Information about the type of shopping was for example not included. The dummy variable for fun shopping is based on travel time and so highly correlated with the dependent variable of travel distance. Further investigation should therefor focus on more detailed information from the surveys. Another limitation for the investigation is that only data of the years 2013 and 2014 is used. More information about the effect of online retail and the shift from offline to online could be explained when data from many years was available.

Another large limitation can be found in the dataset itself. Data about the shopping purpose is not included and could have a large impact for the answer on research question 1. Next to that is the data based on single trips recorded by households. Only the longest trip with a shopping purpose is used for the investigation. Further research could research the pattern of shopping trips for an individual in more detail over a longer period of time.

This research could also focus more on the effect of personal characteristics to travel distance to the stores. The literature also discussed the effect of personal characteristics in more detail. This research did found significant effects for personal characteristics and so these could be investigated in more detail. Furthermore, could this research also be done with data from other countries where other physical barriers are more the case where for example in some countries not everyone has access to the Internet?

References

Annema, J. A. (2013). Transport resitance factors: time, money and effort. In B. Van Wee, J. A. Annema, & D. Banister (Eds.), *The transport system and transport policy: an introduction* (pp. 101–124). Cheltenham, UK. Northampton, MA, USA: Edward Elgar Publishing.

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

CBS. (2017). Minder fysieke, meer online kledingwinkels. Retrieved May 1, 2018, from https://www.cbs.nl/nl-nl/nieuws/2017/35/minder-fysieke-meer-online-kledingwinkels

Childers, T., Carr, C., Peck, J., & Carson, S. (2001). Hedonic and Utilitarian Motivations for Online Retail Shopping Behavior Hedonic and utilitarian motivations for online retail. *Journal of Retailing*, *77*(Winter), 511–535. https://doi.org/10.1016/S0022-4359(01)00056-2

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

Close, A. G., & Kukar-Kinney, M. (2010). Beyond buying: Motivations behind consumers' online shopping cart use. *Journal of Business Research*, *63*(9–10), 986–992. https://doi.org/10.1016/j.jbusres.2009.01.022

Eurostat. (2017). E-commerce statistics for individuals. Retrieved May 1, 2018, from http://ec.europa.eu/eurostat/statistics-explained/index.php/E-commerce_statistics_for_individuals

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

Garcia-Sierra, M., van den Bergh, J. C. J. M., & Miralles-Guasch, C. (2015). Behavioural economics, travel behaviour and environmental-transport policy. *Transportation Research Part D: Transport and Environment*, *41*, 288–305. https://doi.org/10.1016/j.trd.2015.09.023

Hanssen, T.-E. S., Mathisen, T. A., & Jørgensen, F. (2012). Generalized Transport Costs in Intermodal Freight Transport. *Procedia - Social and Behavioral Sciences*, *54*, 189–200. https://doi.org/10.1016/j.sbspro.2012.09.738

Herrmann, R. O., & Beik, L. L. (1968). Shoppers' Movements outside Their Local Retail Area. *Journal of Marketing*, *32*(4), 45–51. https://doi.org/10.2307/1249337

Hsiao, M. H. (2009). Shopping mode choice: Physical store shopping versus e-shopping. *Transportation Research Part E: Logistics and Transportation Review*, *45*(1), 86–95. https://doi.org/10.1016/j.tre.2008.06.002

Jiang, L. (Alice), Yang, Z., & Jun, M. (2013). Measuring consumer perceptions of online shopping convenience. *Journal of Service Management*, *24*(2), 191–214. https://doi.org/10.1108/09564231311323962

Jones, K., & Simmons, J. (1990). *The retail environment*. (Chapmann & Hall, Eds.). New York: Routledge.

Kahneman, D., & Tversky, A. (1979). Prospect Theory: An Analysis of Decision under Risk. *Econometrica*, *47*(2), 263–292. https://doi.org/10.2307/1914185

Lam, T. C., & Small, K. A. (2001). The Value of Time and Reliability: Measurement from a Value Pricing Experiment. *Transportation Research Part E: Logistics and Transportation Review*, *37*, 231–251. Retrieved from https://escholarship.org/content/qt47s4z7z5/qt47s4z7z5.pdf

Lee, R. J., Sener, I. N., Mokhtarian, P. L., & Handy, S. L. (2017). Relationships between the online and in-store shopping frequency of Davis, California residents. *Transportation Research Part A: Policy and Practice*, *100*, 40–52. https://doi.org/10.1016/j.tra.2017.03.001

Lo, H. K., Luo, X. W., & Siu, B. W. Y. (2006). Degradable transport network: Travel time budget of travelers with heterogeneous risk aversion. *Transportation Research Part B: Methodological*, *40*(9), 792–806. https://doi.org/10.1016/j.trb.2005.10.003

Marmorstein, H., Grewal, D., & Fishe, R. P. H. (1992). The Value of Time Spent in Price-Comparison Shopping: Survey and Experimental Evidence. *Journal of Consumer Research*, *19*(1), 52–61. https://doi.org/10.1086/209285

Mokhtarian, P. (2004). A conceptual analysis of the transportation impacts of

B2C e-commerce. *Transportation*, *31*(3), 257–284. https://doi.org/10.1023/B:PORT.0000025428.64128.d3

Mokhtarian, P. L., & Chen, C. (2004). TTB or Not TTB, that is the Question: A Review and Analysis of the Empirical Literature on Travel Time (and Money) Budget. *Transportation Research Part A-Policy and Practice*, *38*(9–10), 643–675.

Overby, J. W., & Lee, E. J. (2006). The effects of utilitarian and hedonic online shopping value on consumer preference and intentions. *Journal of Business Research*, *59*, 1160–1166. https://doi.org/10.1016/j.jbusres.2006.03.008

Park, C., & Kim, Y. (2003). Identifying key factors affecting consumer purchase behavior in an online shopping context. *International Journal of Retail & Distribution Management*, *31*(1), 16–29. https://doi.org/10.1108/09590550310457818

Rabin, A., Rabin, M., Bowman, D., Camerer, C., Dekel, E., Epstein, L., … Weber, R. (2000). Risk aversion and expected-utility theory: A calibration theorem. *Econometrica*, *68*(5), 1281–1292. https://doi.org/10.1111/1468-0262.00158

Rotem-Mindali, O., & Salomon, I. (2007). The impacts of E-retail on the choice of shopping trips and delivery: Some preliminary findings. *Transportation Research Part A: Policy and Practice*, *41*(2), 176–189. https://doi.org/10.1016/j.tra.2006.02.007

Rotem-Mindali, O., & Weltevreden, J. W. J. (2013). Transport effects of e-commerce: What can be learned after years of research? *Transportation*, *40*(5), 867–885. https://doi.org/10.1007/s11116-013-9457-6

Simon, H. A. (1955). A behaviorla model of rational choice. *The Quarterly Journal*, *69*(1), 99–118.

Tversky, A., & Kahneman, D. (1991). Loss Aversion in Riskless Choice : A Reference-Dependent Model Author (s): Amos Tversky and Daniel Kahneman Published by : Oxford University Press. *The Quarterly Journal of Economics*, *106*(4), 1039–1061. https://doi.org/10.2307/2937956

Van De Kaa, J. E. (2010). Prospect Theory and Choice Behaviour Strategies: Review and Synthesis of Concepts from Social and Transport sciences. *EJTIR*, *10*(4), 299–329. Retrieved from www.ejtir.tbm.tudelft.nl

van Wee, B., Annema, J. A., & Banister, D. (Eds). (2013). *The transport system and transport policy: an introduction*. Cheltenham, UK. Northampton, MA, USA: Edward Elgar Publishing.

van Wee, B., Rietveld, P., & Meurs, H. (2002). A constant travel time budget? In search for explanations for an increase in average travel. *Research Memorandum*, *31*.

van Wee, B., Rietveld, P., & Meurs, H. (2006). Is average daily travel time expenditure constant? In search of explanations for an increase in average travel time. *Journal of Transport Geography*, *14*(2), 109–122. https://doi.org/10.1016/j.jtrangeo.2005.06.003

Visser, E.-J., & Lanzendorf, M. (2004). Mobility and Accessibility effect of E-Commerce: a Literature Review. *Tijdschrift Voor Economische En Sociale Geografie*, *95*(2), 189–205. https://doi.org/10.1111/j.0040-747X.2004.00300.x