

## Spatial Shopping Behavior in a Multi-Channel Environment: A Discrete Choice Model Approach

Thomas Wieland<sup>1</sup>

<sup>1</sup> Karlsruhe Institute of Technology, Karlsruhe, Germany

Received: 26 February 2021/Accepted: 22 June 2021

**Abstract.** Spatial impacts of online shopping are discussed frequently in retail geography. Here, online shopping is mostly regarded as a central driver of competition for physical retailing and its locations, such as town centers or malls. Due to its high popularity, cross-channel shopping is sometimes considered to be a support for physical retailing. However, traditional retail location theory does not consider shopping channels other than in-store shopping. Furthermore, although online shopping is far too important to be neglected in examining consumer spatial shopping behavior, there is an obvious lack in the previous literature towards incorporating multi- and cross-channel shopping into store choice models. The present study aims to identify the main drivers of store choice on the basis that both in-store and online shopping alternatives are available, as well as the opportunity for cross-channel shopping. Taking into account existing literature on both physical store choice and multi-channel shopping, hypotheses on the impact of different shopping transaction costs (such as travel time, delivery charges, or uncertainty with respect to the stores' assortment) were derived. Based on a representative consumer survey, real past shopping decisions in three retail sectors (groceries, consumer electronics [CE], and furniture) were collected. The econometric analysis of empirical store choices was performed using a nested logit model which includes both physical and online stores. The results confirm several assumptions of classical retail location theory as well as previous findings from single-firm studies and stated choice experiments on multi-channel shopping behavior. Travel time to physical stores reduces consumer utility and store choice probability, respectively. Consumer sensitivity towards travel time decreases with decreasing purchase frequency of the desired goods. Delivery charges also decrease the likelihood of choosing a store. The impact of cross-channel integration on store choice (assuming the reduction of consumer transaction costs) is considerably lower than expected and differs between retail sectors. While furniture retailers profit from enabling cross-channel shopping, there is no such competitive advantage found for grocery and CE retailers. The positive effect of assortment on condition of diminishing marginal utility is confirmed for grocery stores and CE stores, but not for furniture stores. From a theoretical perspective, this study shows that multi- and cross-channel shopping behavior does not contradict the main thoughts of classical retail location theory. From a practical perspective, the study is a contribution, as store choice models play a significant role in both business location planning and governmental land use planning.

## 1 Introduction

The e-commerce share of total global retail sales doubled from 7.4% in 2015 to 14.1% in 2019 (Statista 2020). In Germany, the market share of online retailing increased from 8.3% in 2015 to 10.1% in 2018. Whilst this share is very small in grocery retailing (1.2%), other retail sectors have reached considerably higher values (e.g., consumer electronics: 31.0%, clothing: 27.7%, furniture: 13.4%) (Handelsverband Deutschland, IFH Köln GmbH 2019). The rising economic relevance of e-shopping has increased competitive pressure for brick-and-mortar retail stores (called “physical stores” hereinafter) and, consequently, has impacted evolved retail locations (such as town centers), shopping malls as well as other retail agglomerations. Thus, it is not surprising that spatial impacts of online shopping are discussed frequently in retail geography, urban planning, and transportation science. In these discussions, online shopping is mostly regarded as a main driver of competition and as contributing towards falling demand for retail properties and rising vacancy rates in town centers (Doherty, Ellis-Chadwick 2010, Singleton et al. 2016, Stepper 2016).

However, this represents only one side of the argument, as the boundaries between shopping channels – in particular, in-store and online shopping – become blurred when considering the rising importance of multi- and cross-channel retailing. By definition, a *multi-channel* retailer sells products on at least two shopping channels *independent from each other*, e.g., via physical stores and online. A *cross-channel* retailer allows for combining different channels within the *same* purchasing process, such as by providing information about products online before buying in-store (“Research online, buy offline”), “Buy online, pick up in store” (called “click and collect” hereafter), or “Buy online, return offline” in the case of returns (Cao, Li 2015). Several consumer surveys have shown that retail customers engage in cross-channel shopping frequently, such as by browsing in one channel and buying in the other or by using “click and collect” (Creditreform Boniversum GmbH 2018, Handelsverband Deutschland, IFH Köln GmbH 2019, McKinsey & Company 2019). Due to its high popularity, cross-channel shopping is sometimes regarded as a “sheet anchor” for physical retailing (Flavián et al. 2020, Heinemann 2015).

Traditionally, store choice and spatial shopping behavior are key topics in retail geography and its related disciplines (such as retail marketing, urban planning, and transportation research). This importance extends from theoretical to empirical perspectives. These studies attempt to analyze and explain the drivers of consumer store choice in a competitive environment of physical shopping locations. However, traditional retail location theory does not consider other shopping channels apart from in-store shopping (Brown 1993, Reigadinha et al. 2017, Timmermans 1993). In the last years, online shopping has become far too important to be neglected in examining consumer decisions with respect to retail stores and locations. Consumers may choose between physical and online stores, whilst weighing the utilities of the available alternatives. Furthermore, consumers may use different shopping channels within the same purchasing process, e.g. by using “click and collect”. By consequence, *the present study aims to identify the main drivers of consumer store choice based on the condition that both in-store and online shopping alternatives are available, as well as the opportunity for cross-channel shopping.* For this purpose, traditional explanatory approaches towards physical store choice from retail location theory are combined with current research findings on multi-channel shopping behavior. The empirical analysis is performed using an econometric store choice model including both in-store and online shopping alternatives as well as cross-channel shopping features of both. The study is based on a representative consumer survey in which real past shopping decisions were observed.

The article is structured as follows. Section 2 provides a literature review of store choice models with respect to in-store shopping and multi-channel shopping behavior. Common features and research gaps of both fields of research with respect to the present research aims are outlined. In section 3, the current research approach is discussed, including the formulation of research hypotheses, description of the modeling approach, and undertaking of the survey. Section 4 presents and discusses the empirical results, including descriptive and modeling results, and continues with an examination of the hypotheses. Section 5 contains the main conclusions of this study and the related limitations.

## 2 Literature review

### 2.1 Consumer spatial shopping behavior

Consumer spatial shopping behavior details consumer choice of a specific retail store (or shopping center) from a given set of shopping alternatives during planning and performing a shopping trip. These alternatives are characterized by a set of attributes which are implicitly weighed by consumers while making the choice. In theories and models of spatial shopping behavior, consumer behavior is assumed to be utility-maximizing, although this does not imply full information and an objective assessment of the shopping alternatives (Timmermans 1993). Spatial shopping behavior is a key concern of classical retail location theory because it is assumed to be the main determinant of a retail store's or supply location's economic success. Two key concepts in retail location theory are predominantly focused on spatial shopping behavior, 1) the strain of *central place theory*, and 2) *spatial interaction models* (sometimes referred to as *market area models*) (Brown 1993, Reigadinha et al. 2017).

The theoretical strain of central place theory (CPT) was introduced by Christaller (1933) and describes the spatial structure of service locations (including retailing) based on microeconomic assumptions of consumer and supplier behavior, both of whom are assumed to act in a utility-maximizing and cost-minimizing manner. A generalization of CPT, including nearly all economic activities was formulated by Lösch (1944), whilst other advances in the literature have focused on specific aspects such as multipurpose shopping or time constraints (Ghosh 1986, Lange 1973). The most important aspect of this theoretical framework with respect to consumer behavior, is distance-dependent demand. Distance-dependent demand means that the demand for goods declines with distance from the supply locations as a result of increasing transport costs borne by the consumers themselves. According to Christaller (1933), consumer sensitivity to transport costs depends on the purchase frequency of the specific good. The maximum distance within which the good is purchased is called its *outer range*, whilst the demand threshold required for economic viability of the seller is its *inner range*. With respect to retailing, expensive and infrequently purchased goods (such as furniture) are assumed to have high outer and inner ranges, whilst daily consumer goods (such as groceries) should have low ranges (Brown 1993).

Independent from CPT, Reilly (1931) constructed a first mathematical market area model which defines the market shares of two competing supply locations (or cities). The explanatory variables are their attraction (size) and the transport costs between them (e.g., distance, travel time) and a consumer origin. This so-called *law of retail gravitation* was extended by Converse (1949) and his *breaking point formula* which calculates the maximum market area of both locations.

Building upon this, Huff (1962, 1964) created a probabilistic market area model based on microeconomic assumptions. The utility of a store (or shopping center)  $j$  for the consumers in area  $i$ ,  $U_{ij}$ , is explained by a multiplicative utility function with two explanatory variables (partial utilities):

$$U_{ij} = A_j^\gamma d_{ij}^{-\lambda} \quad (1)$$

where  $A_j$  is the size of shopping location  $j$ ,  $d_{ij}$  is the travel time from customer origin  $i$  to shopping location  $j$ , and  $\gamma$  and  $\lambda$  are weighting parameters.

According to Huff (1962), store size (measured by selling space) is a proxy variable for the assortment width and depth of the shopping location, which is assumed to increase the consumers' utility. The underlying idea is that consumers choose shopping locations on condition of uncertainty about the in-store availability of the desired products. The larger the number of articles, the higher the likelihood to get them. However, assortment is assumed to be affected by diminishing marginal utility because as the number of articles increases, also the consumers' search and decision costs increase. Thus, store size is assumed to have a positive but degressive impact on store utility ( $0 < \gamma < 1$ ). Applying the law of diminishing marginal utility to a shopping location's variety was originally formulated by Baumol, Ide (1956), to which Huff (1962) refers in his work. The disadvantage of variety in retailing due to increasing search and decision costs is

well known in retail marketing (Broniarczyk, Hoyer 2010, Desmeules 2002). Consumer transport costs are translated in terms of travel time to reflect the time effort of shopping trips rather than distance. Huff (1962) assumes that, due to opportunity costs of traveling to shopping locations, the negative effect of increasing travel time is superlinear ( $|\lambda| > 1$ ) and differs between shopping goods. According to Güßefeldt (2002), the  $\lambda$  parameter directly links to the range concept from CPT as it reflects good-specific sensitivity of consumers towards transport costs.

Following the decision-theoretic concept by Luce (1959), Huff (1962) regards store choice and market areas of shopping locations as probabilistic. The probability for consumers in area  $i$  to patronize shopping location  $j$ ,  $p_{ij}$  is equal to:

$$p_{ij} = \frac{U_{ij}}{\sum_{j=1}^J U_{ij}} \quad (2)$$

where  $J$  is the number of available shopping locations.

The probability is interpreted as the market share of shopping location  $j$  in customer origin  $i$ . The number of consumers from  $i$  patronizing  $j$  (or the expenditure flows from  $i$  to  $j$ ),  $E_{ij}$ , is equal to  $E_{ij} = p_{ij}C_i$ , where  $C_i$  is the total customer/expenditure potential in customer origin  $i$  (Huff 1962). The total market area (consumers or turnover) of a shopping location  $j$ ,  $T_j$  is defined as the sum of customer/expenditure flows over all  $I$  customer origins:  $T_j = \sum_{i=1}^I E_{ij}$  (Huff 1964).

Empirical consumer spatial shopping behavior is typically investigated by using econometric store choice models such as the *Multiplicative Competitive Interaction (MCI) Model* (Nakanishi, Cooper 1974, 1982) for market shares or the *Discrete Choice Model (DCM)* (McFadden 1974) for individual shopping decisions, with the latter representing an umbrella term for different types of logit models. Although not designed for the investigation of shopping behavior, both types of models can be linked to the Huff Model with respect to the underlying assumptions towards utility maximization and probabilistic choice behavior (Wieland 2015). The assumptions on the impact of store size/assortment (positive effect) and travel time (negative effect), as stated in the Huff Model, have been confirmed in many studies, especially with respect to grocery stores (Baviera-Puig et al. 2016, Briesch et al. 2009, Hillier et al. 2015, 2017, Lademann 2007, Mensing 2018, Orpana, Lampinen 2003, Popkowski Leszczyc et al. 2004, Reutterer, Teller 2009, Suárez-Vega et al. 2015, Tihi, Oruc 2012, Vroegrijk et al. 2013, Wieland 2015, 2018a).

However, to the best of the author's current knowledge, there are only a few studies investigating other types of retail goods and supply locations, respectively. Whilst assuming a linear impact of store size ( $\gamma = 1$ ), Huff (1962) finds travel time weighting parameters equal to  $\lambda = 3.191$  for clothing and  $\lambda = 2.723$  for furniture based on his own consumer survey. Kubis, Hartmann (2007) confirm the aforementioned assumptions for shopping malls, as well as clothing and food, with the latter having the strongest (negative) impact of travel time. Wieland (2015) finds support for the Huff Model assumptions as applied to grocery and consumer electronics (CE) retailing; however, the effect of travel time is higher for groceries and the effect of store size is more relevant for CE stores. For do-it-yourself (DIY) stores, store size is found to have a superlinear positive effect, whilst the influence of travel time is stronger than for groceries and CE. Lademann (2007) shows that store size has a degressive influence on store choice in grocery and clothing retailing but a progressive effect for furniture stores. Consumer sensitivity to distance is found to be overproportionate in all cases.

Modeling approaches to store choice behavior have been successively extended by integrating a large number of further explanatory variables into the models. Several studies incorporate objective and/or subjective characteristics of both shopping locations and consumers, including pricing levels, agglomeration effects, chain images, and household attributes (Baviera-Puig et al. 2016, Briesch et al. 2009, Hillier et al. 2015, 2017, Mensing 2018, Orpana, Lampinen 2003, Popkowski Leszczyc et al. 2004, Reutterer, Teller 2009, Tihi, Oruc 2012, Vroegrijk et al. 2013, Wieland 2015, 2018a). However, up until now, store choice literature has been almost completely focused on physical stores and shopping locations, respectively. An exception is Melis et al. (2015), who utilize a store choice model for online stores only.

## 2.2 Spatial perspectives on multi-channel shopping behavior

Research into multi-channel shopping behavior is typically carried out using channel choice models rather than store choice models (as described in section 2.1). However, these models (e.g., binary logit) are simultaneously based on the (implicit or explicit) assumption of utility maximization (Chintagunta et al. 2012, Hsiao 2009, Schmid, Axhausen 2019). Spatial perspectives on channel choice frequently focus on consumer transaction costs related to the specific shopping channel. Transaction costs include all costs (in terms of efforts) which occur within a transaction due to e.g. imperfect information or spatial and temporal barriers. Following the main rationale of transaction costs economics, it is assumed that consumers choose the shopping channel which minimizes their transaction costs (Chintagunta et al. 2012, Teo, Yu 2005). According to Chintagunta et al. (2012), transaction costs occur in all stages of a purchasing process and include:

- Costs related to in-store shopping only: Opportunity costs of time (in particular, travel time to and from a physical store), and traffic expenditures (e.g., fuel costs, ticket purchase in public transport), as well as several other efforts when visiting a shopping location (e.g., carrying the purchased products in a shopping cart)
- Costs related to online shopping only: Waiting costs (due to the time between ordering and receiving the ordered products) and delivery charges
- Costs related to both in-store and online shopping, but to different extents: Search costs (because of consumers' imperfect information; including gathering information on product availability, price, or product attributes), adjustment costs (in particular, additional time efforts when a desired product is not available), and examination costs (depending on the consumers' ability to evaluate the bought products at the time and place of purchase)

Based on a stated choice experiment (fictional book purchase), Hsiao (2009) focuses on transaction costs related to traveling and delivery. The results show that increasing travel costs and travel time decrease the likelihood of choosing a physical store, whilst increasing delivery time discourages consumers from buying online. An analysis of the trade-off between time and cost attributes shows that saving one hour of travel time is worth 5.29 US\$ of travel costs and the average value of product delivery time equals 0.53 US\$ per day. Chintagunta et al. (2012) quantify transaction costs with respect to consumer channel choice based on observed purchases at a multi-channel grocery retailer. They find evidence that several types of transaction costs influence channel choice. Increasing travel time and traffic costs due to a longer distance to a physical store increase the likelihood of using the online channel, whilst delivery charges show the reverse effect. Examination costs discourage consumers from buying online. When purchasing a larger number of items which are difficult to transport, consumers tend to use the online channel because of increasing in-store shopping time and efforts related to item picking and basket carrying. Teo, Yu (2005) focus on information-related transaction costs and their influence on willingness to buy online based on a consumer survey of past shopping decisions. In their study, several types of uncertainty with respect to products and stores (e.g., performance of the product) increased perceived transaction costs, which discouraged consumers from online shopping. Kacen et al. (2013) explore consumer perceptions of the performance of online stores. Among other things, they identify delivery charges, aspects of returns policy and post-purchase service, as well as product uncertainty as perceived disadvantages of online stores.

Moreover, several channel choice studies find similar evidence, although not explicitly referring to the concept of transaction costs. It is frequently found that the likelihood of e-shopping depends on travel efforts (to physical stores) and delivery efforts (with respect to online stores). In particular, the better the spatial accessibility of physical shopping locations, the lower the tendency is towards online shopping, which has been confirmed in both stated choice experiments (Chocarro et al. 2013) and studies on real shopping behavior (Clarke et al. 2015, Ren, Kwan 2009, Zhai et al. 2017, Zhen et al. 2018). With respect to an Italian multi-channel furniture retailer, Marino et al. (2018) outline the large negative effect of increasing delivery time on the choice of the online channel. In their

stated choice experiment regarding grocery and CE shopping in Switzerland, [Schmid, Axhausen \(2019\)](#) demonstrated that increasing travel time decreases the likelihood of visiting a physical store, whilst increasing delivery time and delivery charges discourages consumers from online shopping. They also showed that subjective attitudes of consumers towards shopping risks and uncertainty with respect to e-shopping result in a lower choice probability of the online channel. In stated choice experiments towards channel choice with respect to grocery shopping in China and Norway, [Gatta et al. \(2021\)](#) and [Marcucci et al. \(2021\)](#) identified different influential attributes, including product and delivery service, pricing, assortment, and delivery time, as well as travel time.

For the sake of completeness, it should be noted that the literature on multi-channel shopping behavior also considers several other aspects which are non-spatial in nature (but may result in spatial effects), such as demographic characteristics and other types of shopping attitudes ([Beckers et al. 2018](#), [Clarke et al. 2015](#), [Schmid, Axhausen 2019](#), [Zhai et al. 2017](#), [Zhen et al. 2018](#)), or pricing ([Gatta et al. 2021](#), [Marcucci et al. 2021](#), [Schmid, Axhausen 2019](#)). However, channel-specific transaction costs (as defined above) may be regarded as the common denominator of the vast majority of studies with respect to multi-channel shopping behavior from a spatial perspective. Moreover, one has to keep in mind that several studies on channel choice behavior are based on stated choice experiments ([Chocarro et al. 2013](#), [Gatta et al. 2021](#), [Hsiao 2009](#), [Marcucci et al. 2021](#), [Schmid, Axhausen 2019](#)). The related results might not be transferable to real-world shopping behavior.

### 2.3 *Synthesis and research gaps*

Taking a look at both store choice and channel choice literature, there are obvious commonalities between the explanation patterns. First, model-based investigations are founded on the assumption that consumers pursue utility maximization, which is sometimes equated with cost minimization. Second, channel choice studies frequently investigate channel-specific transaction costs as borne by the consumers. This perspective may be regarded as a link to store choice studies which are also focused on specific transaction costs related to in-store shopping trips. In particular, the Huff Model ([Huff 1962](#)) is explicitly founded on microeconomic assumptions around search and information costs as well as opportunity costs with respect to shopping trips under the assumption of utility maximization. Different types of consumer transport costs are key determinants of store choice in retail location theory (see section 2.1). Following [Melis et al. \(2015\)](#), transport-related transaction costs in online shopping – in particular, delivery time and delivery charges – may be regarded as an equivalent to consumer transport costs occurring during physical shopping trips (However, unlike transport costs according to location theory, delivery time and delivery charges in online retailing are *not* distance-dependent from the consumer perspective. Costs associated with shipping are passed to the supply side, in particular, the retailer and/or the logistics service provider. Occurring costs may be minimized by rationalizing the delivery process. Moreover, large-scale retailers may provide free delivery and implicitly add shipping costs to product prices). Thus, both literature strains follow the same rationale with respect to shopping behavior.

However, considering the research questions of the present study, there are three obvious gaps in the existing store choice literature. First, surprisingly, there is a straight focus on grocery retailing, and thus, other goods and types of retailing are strongly underrepresented. Second, as [Suel, Polak \(2018\)](#) point out, there is a lack of store choice studies incorporating both physical and online stores. This gap is not surprising as the store choice literature is (explicitly or implicitly) related to classical retail location theory ([Christaller 1933](#), [Huff 1962](#)), which does not consider shopping channels other than in-store shopping. In particular, the focus on spatial attributes is not directly applicable to online stores. In short, the existing literature on store choice has not considered multi-channel shopping behavior yet. Third, there is also a lack of research with respect to cross-channel shopping. Cross-channel shopping features (such as “click and collect”) may be regarded as characteristics of retail stores (or chains) but have not yet been considered as explanatory variables in store choice models.

The existing channel choice literature does not fill this gap. First, the choice between

online and in-store shopping is typically investigated on the aggregated level (shopping channel chosen by consumers in general or with respect to a specific choice situation) rather than the store level. Thus, channel choice studies cannot consider spatial or nonspatial characteristics of specific retail stores: “The literature on channel choice for a specific choice occasion, crucially, has largely focused on modelling the choice between online and in-store using aggregated alternative categories. In reality, however, shoppers choose from a set of alternatives that include both online and in-store alternatives available to them. Using aggregated alternatives based on channel, therefore, ignores the heterogeneity within each channel” (Suel, Polak 2018). Second, in all conscience, no channel choice study was found that investigates possible effects of cross-channel opportunities. For example, the ability of “click and collect” might influence the choice of a specific cross-channel retailer rather than a single-channel retailer.

Consequently, there is a research gap with respect to including online stores into a store choice framework. This is all the more problematic as 1) consumers may choose between different shopping channels *and* stores, and 2) this decision is based on both the spatial structure of shopping locations and the characteristics of online stores.

### 3 Research approach

#### 3.1 Research hypotheses

The research hypotheses in the present study aim at the most frequent aspects of store *and* channel choice and their linkages, as discussed in section 2.3. The underlying rationale is that consumers may choose between both physical and online stores, whilst assuming utility maximization and minimization of transaction costs, respectively.

Taking into account the literature on spatial shopping behavior (see section 2.1), the first two hypotheses relate to the assortment of stores. Following Huff (1962), assortment size is assumed to increase consumer utility of a store. Although this basic assumption is stated for physical shopping locations, there is no reason why it should not be valid for online stores. Furthermore, Melis et al. (2015) show positive assortment effects with respect to the choice of online stores. Thus, the first hypothesis is:

*H<sub>1a</sub>: An increasing assortment increases the likelihood of choosing a store*

However, following the assumptions of Baumol, Ide (1956) and Huff (1962), and the corresponding empirical findings in previous store choice studies (see section 2.1), it is assumed that a larger assortment leads to increasing search and decisions costs on the consumer side. Thus, assortment size is expected to be affected by diminishing marginal utility. This suggests that the increase in store utility must be sublinear:

*H<sub>1b</sub>: The positive effect of assortment is degressive*

Another fundamental statement of both central place theory and market area models is distance-dependent demand (Christaller 1933, Lösch 1944, Huff 1962, Reilly 1931). This assumption was empirically confirmed many times, especially with respect to grocery stores (see section 2.1). Regarding channel choice studies, it was frequently found that spatial accessibility of physical shopping locations affects the choice probability of online stores (see section 2.2). Travel time, which will occur inevitably when shopping at a physical store, may be regarded as the main type of transaction costs during a shopping trip. Consequently, it is assumed that travel time to physical stores decreases their utility for the consumers:

*H<sub>2a</sub>: Increasing travel time decreases the likelihood of choosing a store*

One key message of CPT is that, during a shopping trip, consumer sensitivity towards distance and travel time depend, respectively, on the order of the purchased good. The more valuable and the less frequently purchased, the higher the transport effort consumers are willing to spend to acquire a specific good (Christaller 1933, Brown 1993). In empirical store choice studies, different estimates for distance or travel time sensitivity were found,

however, most results are focused on grocery stores only (see section 2.1). The present study incorporates three different retail sectors (groceries, consumer electronics, furniture), with groceries being the most frequently purchased. Following the CPT assumptions, with respect to the negative effect of travel time, it is hypothesized that:

*H<sub>2b</sub>: The negative effect of travel time declines with decreasing purchasing frequency*

Several channel choice studies based on both experiments and actual shopping behavior show that delivery-specific efforts reduce the consumer's likelihood of buying online (see section 2.2). Unfortunately, delivery time cannot be determined reliably in any case, as the effective waiting time depends on several factors such as the day of order or local characteristics of the postal service. However, delivery charges of online stores may be observed from the store's website. Following Melis et al. (2015), delivery charges of online stores might be regarded as the equivalent of travel time or distance with respect to physical stores. In any case, home delivery fees belong to transaction costs which are specific for the online channel (Chintagunta et al. 2012). Based on previous findings on multi-channel shopping behavior from single-firm studies or choice experiments, it is assumed that delivery charges reduce the utility of a store:

*H<sub>3</sub>: Increasing delivery charges decrease the likelihood of choosing a store*

In the Huff model (Huff 1962), store size is assumed to have a positive impact on store utility due to the consumer's imperfect information with respect to the store availability of the desired products (the larger the assortment, the more likely a "successful" shopping trip will be). However, this assumption cannot be transferred to e-shopping as online stores allow for browsing the whole assortment by visiting the store's website. In principle, the same is true for cross-channel retailers which provide access to product details (e.g., price) and store availability on their websites. If they also provide the "click and collect" option, consumers are given a guarantee to pick up the desired products in a specific store (on condition that the service works efficiently). Moreover, cross-channel integration may also facilitate product returns ("Buy online, return offline"). Consequently, from the transaction costs perspective, retail stores (or chains) providing cross-channel shopping reduce consumer uncertainty and search costs. Thus, it is assumed that cross-channel services increases consumer utility. As surveys show, many (German) retail customers engage in cross-channel shopping (see section 1). Accordingly, the last two hypotheses aim at the (assumably positive) effect of a retailer's cross-channel integration by providing 1) an integrated online shop (including information about product availability etc.), and 2) the "click and collect" service:

*H<sub>4</sub>: An integrated online shop increases the likelihood of choosing a store*

*H<sub>5</sub>: Providing the "click and collect" service increases the likelihood of choosing a store*

The concrete definition of an "integrated online shop" and the "click and collect" service can be found in section 3.2 (Data collection and processing).

### 3.2 Data collection and processing

#### 3.2.1 Survey approach and questionnaire design

The present analysis is founded on a *revealed preference approach*, which means that (shopping) preferences of individuals are inferred from their actual decisions in real-world situations, rather than on stated choice experiments (Train 2009). This kind of analysis requires 1) the observation of past shopping decisions, and 2) collecting information on the (physical and online) stores which are relevant for the consumers (*choice set*).

The observation of real-world shopping behavior was performed using a quantitative self-administered consumer survey. To account for actual shopping decisions, in the consumer survey, the respondents were asked about their last three purchases with respect



to groceries, consumer electronics, and furniture (“Where did you make your last three purchases of xxx?”). For any purchase, the specific shopping destination (e.g., “Aldi in street X of city Y”, “IKEA in commercial area Z”, “Amazon online”) was noted as well as the corresponding total expenditures (in EUR) incurred in the shopping trip (Wieland 2015, 2018a).

The respondents were also asked for the purchase frequency with respect to the considered retail sector (Groceries: number of purchases during the last four weeks, other goods: number of purchases during the last 12 months). Furthermore, the questionnaire included socio-demographic characteristics of the respondents.

### 3.2.2 Survey data collection and consumer sample

The consumer survey was conducted from March to June 2019 in two German regions. To cover both rural and urban regions in Germany, two spatially different study regions were chosen: 1) South Lower Saxony (pop. of 531,814 in 2018) with three rural counties (“sparsely populated rural counties”) and one large city, and 2) the Middle Upper Rhine Region (pop. of 1,043,465 in 2018) with three urbanized counties and one large city, according to the official German planning classification of regions (Bundesinstitut für Bau-, Stadt- und Raumforschung 2017). In order to achieve representativeness, the communication mode of the survey was a self-administered postal survey. The addresses of the contacted individuals were drawn as a random sample from official address data as held by the residents’ registration offices of all included municipalities. All respondents were given the alternative option to fill out the questionnaire in a web form. The target population was defined as all residents aged 15 or above, as this is the first category in German official population statistics (the range 15 to 18 years) in which some financial autonomy with respect to shopping appears reasonable.

In both survey areas, 9,109 randomly sampled individuals were contacted (South Lower Saxony: 3,109; Middle Upper Rhine Region: 6,000). The net sample equalled 1,375 individuals (South Lower Saxony: 297; Middle Upper Rhine Region: 1,078). Taking into account 355 neutral losses (in particular, invalid addresses and deceased persons), the response rate was equal to 15.7 % (South Lower Saxony: 10.0 %; Middle Upper Rhine Region: 18.6 %). Table 1 shows the socio-demographic characteristics of the respondents compared to the population data (if available), aggregated by survey area. Female respondents are slightly over-represented in the survey. Respondents in working life (between 25 and 65 years) are a little under-represented, which might be explained by a shorter time budget with respect to survey participation. About 11% in both survey areas decided to fill out the questionnaire online.

### 3.2.3 Collection of store information

The local physical stores and the online stores make up the choice sets (alternatives available to the consumers) for each survey area. All stores which are comparable with respect to their assortment width were included into the analysis. In grocery retailing, all grocery stores according to the store format definition by The Nielsen Company (Germany) GmbH (2016) were considered as relevant (small/large supermarkets, small/large hypermarkets, and discounters). The relevant CE stores include all physical stores and online shops providing the five following categories, as classified by GfK (2019): “Electrical Household Appliances, Lighting”, “Consumer Electronics, Electronic Media”, “Information Technology”, “Telecommunications”, and “Photography, Optics”. This definition applies to all big-box CE chains such as *Media Markt*, *Saturn*, *Euronics XXL*, and *Expert* (and their online shops), as well as several online retailers such as *Amazon*. Additionally, specialty stores which provide the first two *and at least* one more of the aforementioned categories were also included. These stores are flagged with a control variable in the models (see section 3.3). All furniture stores were regarded as relevant if they provided the GfK category “Furnishings”, including the sub-categories “Furniture” and “Textiles; wood-, basket- and cork goods” (GfK 2019). This applies to common furniture chains such as *IKEA*, *XXXLutz*, and *Roller*, as well as to several independent furniture stores. In addition, departments of large hypermarkets and department stores

Table 1: Socio-demographic characteristics of the respondents by survey area

Variables	Categories	Survey area 1		Survey area 2		Pop. 2018 %	Pop. 2018 %
		Sample 2019 n	2018 %	Sample 2019 n	2018 %		
Gender	Female	155	52.7	51.1	598	56.5	50.1
	Male	138	46.9	48.9	448	42.3	49.9
	No information	1	0.3	–	12	1.1	–
Age	15 – <18	10	3.4	3.2	22	2.1	3.1
	18 – <25	37	12.7	9.8	89	8.4	9.8
	25 – <45	55	18.8	25.8	247	23.4	29.7
	45 – <65	91	31.2	34.2	425	40.2	33.7
	65 – <75	57	19.5	12.6	158	15.0	11.2
	>= 75	42	14.4	14.5	115	10.9	12.6
Household size	1	56	19.5	n/a	165	15.7	n/a
	2	144	50.2	n/a	478	45.5	n/a
	3	45	15.7	n/a	190	18.1	n/a
	4	36	12.5	n/a	155	14.7	n/a
	>4	6	2.1	n/a	63	6.0	n/a
Working status	Employed/self-employed	131	45.2	n/a	601	57.2	n/a
	Retired	100	34.5	n/a	291	27.7	n/a
	School or university	41	14.1	n/a	91	8.7	n/a
	Homemaker (m/f)	6	2.1	n/a	40	3.8	n/a
	Unemployed	5	1.7	n/a	7	0.7	n/a
	Other	7	2.4	n/a	21	2.0	n/a
Type of survey	Written survey (mail)	265	89.2	–	957	88.8	–
	Online survey	32	10.8	–	121	11.2	–

*Note:* Because of missing values, the sample sizes differ for each characteristic.

were included given that they met the criteria. These included CE departments of *Real* or *Karstadt*.

The relevant stores were collected via desktop research using the websites of the retail chains and stores, respectively, in March 2019. After finishing the consumer survey, the choice sets for each survey area were corrected and supplemented according to the requirements of Discrete Choice analysis (Train 2009). Additional physical stores outside the survey areas, but relevant in the survey (i.e., mentioned in the survey in at least two municipalities) and further online stores which had not been considered before were added. Furthermore, stores with no observed purchases were excluded from the choice sets. Data on all physical stores was collected including their street address and store size (selling space in sqm) as well as information about cross-channel integration of the store/chain. According to the hypotheses, an “integrated online shop” was defined as a web platform of a cross-channel retailer which provides 1) information about the full assortment of both the online shop and the associated outlets, and 2) an availability check for each product in a given store, and 3) information on in-store prices as well as some product details. Providing the “click and collect” option means that retail customers may order the desired product online and pick it up in a selected store. Street addresses and information with respect to cross-channel integration were obtained from the store/chain websites. The store size values were retrieved from the corresponding companies and publicly available information such as newspapers and urban land use plans.

Huff (1962) argues that store size is much easier to ascertain than the number of products provided by shopping locations, and for this reason, used selling space as a proxy for assortment. In the present case, however, this argument helps little because online stores have no “selling space”. Instead, similar to Briesch et al. (2009), the present study uses the number of products provided by each store as the assortment indicator. For the grocery stores, the number of articles was available for 50 supermarkets on their websites (mostly from the chains *Edeka* and *Rewe*). Based on this data and the corresponding selling spaces, chain-specific average values (articles per sqm) were calculated and used to estimate the total number of products for the remaining supermarkets. For the grocery discounters, available average values for store sizes and numbers of products (Hahn

Table 2: Estimation results for the assortment model for consumer electronics stores – Dependent variable:  $\ln(\text{number of articles})$ 

variable	coefficient (std. error)
$\ln(\text{store size in sqm})$	0.914*** (0.069)
Dummy EP	1.448*** (0.459)
Dummy Expert	2.651*** (0.548)
Dummy Euronics	1.306** (0.521)
Dummy Media Markt	2.251*** (0.580)
Dummy Saturn	2.427*** (0.604)
Dummy not full range	2.588*** (0.326)
Observations	42
$R^2$	0.997
Adjusted $R^2$	0.996
Residual Std. Error	0.528 (df = 35)
F Statistic	1,447.725*** (df = 7; 35)

Note: \* $p < 0.1$ ; \*\* $p < 0.05$ ; \*\*\* $p < 0.01$

Gruppe 2019, Statista 2019) were used to estimate the assortment size based on the existing data on selling space. The number of articles offered in CE and furniture stores was obtained from the online shops of the corresponding companies (e.g., *Media Markt*, *IKEA*) and independent stores. As most of the chains provide an integrated online shop including an availability check for each store, this procedure was done using own functions and scripts for web scraping with the help of the package *httr* (Wickham 2019) in *R* (R Core Team 2019).

However, there was no information on the number of articles available for all stores, and thus, missing values had to be interpolated in the following way. Based on the available store information on 42 CE stores and 45 furniture stores, regression models were estimated with the number of articles as dependent variable and store size (in sqm) as well as chain dummies as independent variables. Both assortment size and selling space were transformed by natural logarithm. See tables 2 and 3 for the estimations. Both models show a significant positive relationship between store size and the number of articles, which is in a similar order of magnitude and nearly proportionate (coefficients of 0.914 and 0.873, respectively). However, there are additional chain-specific differences in the number of articles, because most of the chain dummies are significant as well. These models were used to interpolate the number of articles of the remaining stores. The assortment of the online shops was retrieved by visiting their web platforms.

### 3.2.4 Subsequent processing of spatial data

Both the street addresses of the survey respondents (residential address) and the physical stores were geocoded. Travel times between all respondents and stores were calculated. These steps were performed in *R* (R Core Team 2019) using the packages *osrm* (Giraud 2019), *tmtools* (Tennekes 2019), and *MCI2* (Wieland 2019), which provide access to *OSM Nominatim* (*OpenStreetMap* address database) and *OSRM* (*OpenStreetMap Routing Machine*). The geocoding of address locations was checked and corrected manually. Travel time is defined here as the fastest route between origins and destinations in terms of car driving time in minutes.

Table 3: Estimation results for the assortment model for furniture stores – Dependent variable:  $\ln(\text{number of articles})$ 

variable	coefficient std. error
$\ln(\text{store size in sqm})$	0.873*** (0.019)
Dummy Dänisches Bettenlager	1.110*** (0.160)
Dummy IKEA	0.955*** (0.238)
Dummy Möbel Heinrich	1.857*** (0.439)
Dummy Mömax	0.255 (0.285)
Dummy Poco	1.485*** (0.284)
Dummy Porta	0.386 (0.343)
Dummy Roller	1.137*** (0.331)
Dummy SB Möbel Boss	0.424 (0.427)
Dummy XXXLutz	1.439*** (0.345)
Dummy VME	0.796** (0.333)
Observations	45
$R^2$	0.998
Adjusted $R^2$	0.998
Residual Std. Error	0.398 (df = 34)
F Statistic	1,779.251*** (df = 11; 34)

Note: \* $p < 0.1$ ; \*\* $p < 0.05$ ; \*\*\* $p < 0.01$

### 3.3 Modeling approach

The present analysis operates on a disaggregated level, which means that individual (shopping) decisions are investigated. Thus, a Discrete Choice Model (DCM) is utilized (Hillier et al. 2015, 2017, Lademann 2007, Popkowski Leszczyc et al. 2004, Reutterer, Teller 2009, Vroegrijk et al. 2013). DCM is an umbrella term for a number of logit models which differ with respect to the sets of explanatory variables they include. The common denominator is that they describe the choice of an alternative  $j$  by an individual  $i$  ( $i = 1, \dots, I$ ), whilst the alternatives stem from a *choice set*, which is a collection of  $J$  options available to the individual decision-maker ( $j = 1, \dots, J$ ). Based on observed individual decisions in artificial or real-world situations (*stated preference approach* and *revealed preference approach*, respectively), the choice probabilities are derived from the assumption of utility-maximizing behavior (Greene 2012, Train 2009). The present analysis focuses on the attributes of the choice alternatives, and the related logit model compares the utilities of the alternatives (*conditional logit model*).

The following description is based on the presentation of choice models in Greene (2012) and Train (2009), which is applied to choice situations with respect to retail stores. The utility of store  $j$  for consumer  $i$ ,  $U_{ij}$  is described by a linear utility function, which includes an explained part (representative utility) and an unexplained part (error term):

$$U_{ij} = \beta \mathbf{x}'_{ij} + \varepsilon_{ij} \quad (3)$$

where  $\mathbf{x}'_{ij}$  is a set of explanatory variables,  $\beta$  is a set of the corresponding regression

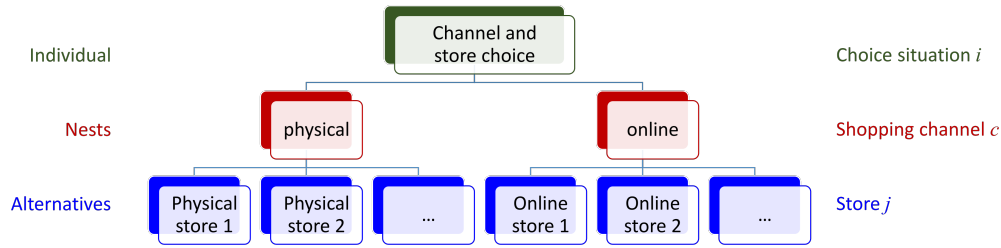


Figure 1: Nested logit approach for store choice including physical and online stores  
Source: own illustration.

coefficients which are estimated empirically, and  $\varepsilon_{ij}$  is the error term.

The choice probability is the highest for the alternative with the largest utility relative to the others. The probability that consumer  $i$  chooses store  $j$  (which means that the dependent variable in choice situation  $i$ ,  $Y_i$  equals alternative  $j$ ) is equal to:

$$\Pr(Y_i = j) = p_{ij} = \frac{\exp(\beta \mathbf{x}'_{ij})}{\sum_{j=1}^J \exp(\beta \mathbf{x}'_{ij})} \quad (4)$$

Estimation of DCM/logit models is performed using the Maximum Likelihood (ML) technique. The log-likelihood of the conditional logit model is (Greene 2012):

$$\ln L = \sum_{i=1}^I \sum_{j=1}^J d_{ij} \ln p_{ij} \quad (5)$$

where  $d_{ij} = 1$  if  $Y_i = j$  and 0 otherwise.

However, competing stores of two shopping channels (in-store and online shopping) are included, and it is unlikely that all consumers have identical patterns of substitution. More precisely, it is to be expected that some consumers are more likely to buy online than others. If one choice alternative is removed from the choice set, e.g., an online retailer, it is questionable whether the choice probabilities of *all* other alternatives, including physical stores, rise proportionately. Thus, the IIA (independence from irrelevant alternatives) assumption of the standard conditional logit model (referring to the ratio of the probabilities of two alternatives  $j$  and  $k$ , which is assumed to be independent from alternatives other than  $j$  and  $k$ ) might not hold true for store choices with respect to different channels. This assumption leads to using a *nested logit model*, which is to be preferred when the alternatives in the choice set can be partitioned into subsets (nests). This means that the IIA assumption holds for the alternatives within each nest but not for alternatives from different nests (Greene 2012, Train 2009). With respect to store choice in the multi-channel context, we define two nests, online and physical shopping alternatives, and each  $j$  store belongs to one of them. Originating from choice situation/individual  $i$ , the first choice is the shopping channel  $c$ , and the second choice is the store  $j$  (see figure 1). However, note that the nested logit model approach is motivated statistically (correlation among alternatives) and not designed as a model for behavior (Greene 2012). In fact, the nested structure does not imply that the *real* consumer decisions are made consecutively.

Following Greene (2012) and Train (2009), the nested logit model for channel and store choice can be defined as follows. The choice probability of store  $j$  ( $j = 1, \dots, J$ ), which belongs to channel  $c$  ( $c = 1, \dots, C$ ), in choice situation  $i$  ( $i = 1, \dots, I$ ),  $p_{ijc}$ , is the product of two logit choice probabilities:

$$p_{ijc} = p_c p_{ij|c} \quad (6)$$

where  $p_c$  is the marginal probability of choosing an alternative (store) out of nest (channel)  $c$ , and  $p_{ij|c}$  is the conditional probability of choosing alternative (store)  $j$  on condition that an alternative from nest (channel)  $c$  is chosen. The probability that consumer  $i$

decides for store  $j$  is:

$$p_{ij|c} = \frac{\exp(\beta \mathbf{x}'_{ij|c})}{\sum_{j=1}^J \exp(\beta \mathbf{x}'_{ij|c})} \quad (7)$$

where  $\mathbf{x}'_{ij|c}$  is a set of explanatory variables, and  $\beta$  is a set of the corresponding regression coefficients. The probability that consumer  $i$  chooses a store which belongs to channel  $c$  is equal to:

$$p_c = \frac{\exp(\gamma \mathbf{z}'_{ic} + \lambda_c IV_{ic})}{\sum_{c=1}^C \exp(\gamma \mathbf{z}'_{ic} + \lambda_c IV_{ic})} \quad (8)$$

where  $\mathbf{z}'_{ic}$  is a set of explanatory variables,  $IV_{ic}$  is the “inclusive value” for channel  $c$ , and  $\gamma$  and  $\lambda_c$  are regression coefficients.

The inclusive value of nest  $c$ ,  $IV_{ic}$  equals the log-sum of the representative utilities of all alternatives (stores) which belong to nest (channel)  $c$ .

$$IV_{ic} = \ln \left( \sum_{j=1}^{J_c} \exp(\beta \mathbf{x}'_{ij|c}) \right) \quad (9)$$

The term  $\lambda_c IV_{ic}$  represents the utility which decision-maker  $i$  receives from the choice of an alternative from nest  $c$  (Train 2009). Therefore, this term is called “channel utility” hereafter.

The log-likelihood of a nested logit model which is estimated with a full information Maximum Likelihood estimation is equal to (Greene 2012):

$$\ln L = \sum_{i=1}^I \ln [p_c p_{ij|c}] \quad (10)$$

The utility function relates to the aforementioned hypotheses (see section 3.1) but includes specific control variables for each retail sector, as described below. The utility of store  $j$  for individual consumer  $i$ ,  $U_{ij}$ , is equal to:

$$U_{ij} = \alpha_{ij} + \beta_1 \ln A_j + \beta_2 t_{ij} + \beta_3 d_j + \beta_4 OS_j + \beta_5 CC_j + \sum_{v=1}^V \delta_v K_{vj} + \varepsilon_{ij} \quad (11)$$

where  $A_j$  is the assortment size of store  $j$ ,  $t_{ij}$  is the travel time between consumer  $i$  and store  $j$ ,  $d_j$  equals the delivery charges of store  $j$ ,  $OS_j$  is a dummy variable indicating whether store  $j$  has an integrated online shop ( $OS_j = 1$ ) or not ( $OS_j = 0$ ),  $CC_j$  is a dummy variable indicating whether store  $j$  provides “click and collect” ( $CC_j = 1$ ) or not ( $CC_j = 0$ ),  $K_{vj}$  is the  $v$ -th control variable ( $v = 1, \dots, V$ ) with respect to store  $j$ ,  $\alpha_{ij}$ ,  $\beta_1, \beta_2, \beta_3, \beta_4, \beta_5$ , and  $\delta_v$  are regression coefficients, and  $\varepsilon_{ij}$  is the error term.

Effects which are related to specific chains or store formats are outside the scope of this study. However, taking into consideration such effects, the utility functions for each retail sector are extended by control variables. These include dummy variables for store chains. For grocery retailing, all chains are included with at least two sites in the respective survey area (e.g., *Aldi*, *Edeka*, *Rewe*). Furthermore, a dummy variable indicating whether store  $j$  is an organic food store is included. The models for CE and furniture stores are extended by dummy variables for the multi-channel retailers (e.g., *Media Markt*, *IKEA*), on condition that these chains are available in the form of physical stores in the respective survey area, and for *Amazon* and *eBay*, which may be regarded as main sources of online shopping in both sectors. Additionally, the models for CE stores also include a dummy variable which indicates whether store  $j$  is a specialty store which avoids specific categories such as digital media or digital cameras but still provides the main CE categories (see section 3.2). These stores are identified as “not full range” providers in the present study.

For each retail sector and survey area, two models were estimated, the conditional (non-nested) logit model and the nested logit model. As there are no grocery stores with integrated online shops, the dummy variable indicating this (term  $\beta_4 OS_j$  in the utility function) is left out in the grocery store models. Data preparation and model estimations

Table 4: Included stores by survey area, retail sector and store format

	Survey area 1				Survey area 2			
	No.	Number of articles [mean]	W/ int. online shop [No.]	W/ click and collect [No.]	No.	Number of articles [mean]	W/ int. online shop [No.]	W/ click and collect [No.]
Grocery stores								
SM 100-399 sqm	4	2,346.4	–	0	13	3,733.6	–	0
SM 400-999 sqm	17	8,539.1	–	0	44	9,086.3	–	0
HM 1,000-<2,500 sqm	53	19,091.1	–	4	67	18,207.8	–	7
HM ≥ 2,500 sqm	19	40,021.7	–	0	26	47,449.0	–	1
Discounter	76	4,831.5	–	0	163	4,867.2	–	0
Department	–	–	–	–	1	23,292.8	–	0
Online stores	–	–	–	–	2	5,406.5	–	1
Consumer electronics stores								
Specialty store	9	681.8	2	5	13	1,581.6	2	2
Big-box store	13	12,282.0	13	13	11	15,012.4	11	11
Department	4	770.4	1	3	10	681.1	1	9
Online stores	11	14,118,316.5	3	3	19	8,221,947.0	4	5
Furniture stores								
<1,000 sqm	7	1,011.1	7	7	12	605.9	6	6
1,000 -<5,000 sqm	7	1,220.7	4	4	14	788.1	1	1
5,000 -<10,000 sqm	4	5,954.5	4	2	10	4,066.5	4	4
≥10,000 sqm	11	10,545.7	7	4	16	15,088.1	8	8
Online stores	11	5,852,952.0	3	3	27	3,173,645.9	5	6

Note: SM = supermarket, HM = hypermarket

were performed in  $R$  (R Core Team 2019) using the *mlogit* package (Croissant 2020). As the analysis only includes two shopping channels (in-store and online), the coefficient  $\lambda_c$  was left out for physical stores. Thus, only  $\lambda_c IV_{ic}$  for online stores is included, which reflects the channel utility of e-shopping. For the ML estimation of the models, the default method of the *mlogit* package (*bfgs*) was used. The significance level was set to 10% ( $p < 0.1$ ), which is the default threshold of regression models in  $R$ . Model visualization was done with the *stargazer* package (Hlavac 2018).

## 4 Results and discussion

### 4.1 Descriptive statistics

Table 4 shows both the physical and online stores which were included in the analysis according to survey area, retail sector and store format. These stores form the choice sets for the consumers in both survey areas. The grocery stores are classified according to the store format definition by The Nielsen Company (Germany) GmbH (2016). In the first survey area (South Lower Saxony), no purchases at grocery online retailers were observed and the delivery service by *Rewe* and *Amazon* (which is provided in survey area 2) was not available there during the time period of the survey. For consumer electronics and furniture stores, there is no established store format definition. Thus, CE stores are classified with respect to distinguishing between big-box stores (such as *Media Markt* or *Saturn*) and specialty shops more focused on assortment depth (such as *EP*). Furniture stores are sorted with respect to their selling space. The classification “department” includes departments of hypermarkets and department stores which match the aforementioned criteria of assortment width (see section 3.2). The table presents the number of stores, the average number of articles (arithmetic mean), and the number of stores with integrated online shop and “click and collect” service, respectively.

Although, theoretically, online stores are available from elsewhere, the number of online stores in the choice sets differ between the survey regions. This is because a different set of online shops were reported in the survey. The average number of articles differs between store formats, with the online stores, as to be expected, having a quite larger assortment than physical stores (e.g., 14 and 8 millions of articles in online CE stores compared to 12,282 and 15,012 articles in big-box-CE stores). The cross-channel

Table 5: Observed purchases and purchase frequency by survey area and retail sector

	Survey area 1						Survey area 2					
	Purcha- ses [No.]	Mean	SD	Median	Q2.5	Q97.5	Purcha- ses [No.]	Mean	SD	Median	Q2.5	Q97.5
Groceries*	779	8.8	4.9	8	2	20	2,890	9.2	5.7	8	2	20
Consumer electronics <sup>+</sup>	549	4.7	4.5	3	1	17.28	1,996	5.1	7.7	3	1	20
Furniture <sup>+</sup>	317	1.9	1.6	2	0	5.8	1,258	2.2	1.9	2	0	8

Note: Purchase frequency: \* = last 4 weeks, <sup>+</sup> = last 12 months; SD = standard deviation, Q = quantile

integration obviously differs between store formats. In CE retailing, all big-box stores belong to chains which provide an integrated online shop and the “click and collect” service. However, this does not hold true for furniture stores which include several independent retailers. No grocery stores with integrated online shop were obtained. Online stores with cross-channel integration mostly belong to the same chains as the physical stores providing cross-channel shopping (such as *Media Markt*, *Saturn*, *IKEA*, *XXXLutz*).

In table 5, descriptive statistics of all observed purchases (including those which are not relevant for the Discrete Choice analysis in section 4.2) by survey area and retail sector are shown. The table presents the number of observed purchases and measures of central tendency and distribution (arithmetic mean, standard deviation, median, quantile) for the observed purchase frequencies. All in all, 7,789 purchases were reported, with 1,645 in survey area 1 (South Lower Saxony) and 6,144 in survey area 2 (Middle Upper Rhine Region). Arithmetic means and median values for purchase frequency are quite similar within the two survey areas (e.g., arithmetic means of 8.8 and 9.2 grocery purchases, respectively, in the last four weeks, with corresponding median values equal to 8). Obviously, as to be expected, groceries are much more frequently purchased than consumer electronics (arithmetic mean of 4.7 resp. 5.1 purchases during the last four months; median equal to 3) and furniture, with the latter being the least frequently purchased (on average, 1.9 and 2.2 furniture purchases in the last four months; median equal to 2).

Out of the 7,789 purchases, 5,700 (73.2 %) are relevant for the Discrete Choice analysis (survey area 1: 1,186; survey area 2: 4,514). Table 6 presents the included purchases (choice situations) and the corresponding expenditures by survey area, retail sector, and store format. For the purchases, the absolute and relative frequencies (share of purchases in %) are shown. The expenditures are presented by relative frequencies (share of expenditures in %) and arithmetic means (average expenditures). The classification is according to table 4. With respect to all store types, there are obviously differences between the share of purchases and the share of expenditures between store formats and shopping channels, respectively. These differences are also represented in the observed average expenditures. In grocery retailing, supermarkets are visited for rather small purchases (fill-in trips), whilst the expenditure shares of large hypermarkets and discounters exceed the corresponding shares of purchases. The share of online purchases and expenditures, respectively, in survey area 2 is remarkably small (0.3 and 0.4 %, respectively).

The market share of online stores is the largest for consumer electronics. However, the shares of purchases are much larger than the expenditure shares due to substantially lower expenditures per purchase. In both survey areas, about half of both purchases (47.9 and 49.1%, respectively) and expenditures (49.9 and 50.9%, respectively) relate to big-box stores (e.g., *Media Markt*, *Saturn*, *Expert*). Note that in the reported CE online purchases, *Amazon* is extremely dominant as this supplier represents about three quarters of online purchases and expenditures (Survey area 1: 30.9% of all purchases and 23.4% of total expenditures; survey area 2: 30.8 and 22.1%, respectively). In furniture retailing, the expenditure shares of the online channel (5.3 and 5.8%, respectively) is about two thirds smaller than the shares of purchases (16.8 and 15.2%, respectively). Furniture stores of the largest category ( $\geq 10,000$  sqm) account for the vast majority of both purchases and expenditures (e.g., *IKEA*, *XXXLutz*). The most dominant companies in the observed furniture purchases are *IKEA* in survey area 1 (41.7% of all purchases and 50.2% of total



Table 6: Observed purchases (only included) by survey area, retail sector and store format

	Survey area 1				Survey area 2			
	Purchases		Expenditures		Purchases		Expenditures	
	[No.]	[%]	[%]	[EUR], mean	[No.]	[%]	[%]	[EUR], mean
<i>Grocery stores</i>								
SM 100-399 sqm	8	1.3	0.4	13.62	48	2.1	1.7	39.31
SM 400-999 sqm	59	9.4	6.0	25.44	265	11.4	10.1	42.65
HM 1,000-<2,500 sqm	215	34.2	31.8	36.84	618	26.5	26.7	48.27
HM $\geq$ 2,500 sqm	137	21.8	25.7	46.64	460	19.7	25.7	62.51
Discounter	209	33.3	36.1	42.95	929	39.8	35.1	42.25
Department	–	–	–	–	7	0.3	0.2	37.43
Online store	–	–	–	–	6	0.3	0.4	68.67
<i>Consumer electronics stores</i>								
Specialty store	19	5.6	15.9	609.42	46	3.6	8.6	570.26
Big-box store	162	47.9	49.9	224.41	623	49.1	50.9	250.31
Department	14	4.1	2.4	124.79	47	3.7	2.5	164.66
Online store	143	42.3	31.8	161.70	553	43.6	38.0	210.70
<i>Furniture stores</i>								
<1,000 sqm	17	7.7	1.1	134.41	37	4.1	4.0	908.89
1,000 -<5,000 sqm	28	12.7	24.5	1,782.00	64	7.0	13.9	1,847.30
5,000 -<10,000 sqm	20	9.1	2.7	277.75	151	16.6	18.8	1,061.01
$\geq$ 10,000 sqm	118	53.6	66.3	1,143.37	521	57.1	57.5	938.93
Online stores	37	16.8	5.3	289.54	139	15.2	5.8	354.22

Note: SM = supermarket, HM = hypermarket

expenditures) and *XXXLutz* in survey area 2 (23.4 and 39.3%, respectively), including both in-store and online purchases.

#### 4.2 Discrete Choice Model results

The results of the DCM analyses are presented in tables 7 (grocery stores), 8 (consumer electronics stores), and 9 (furniture stores). Table 10 shows the corresponding hypotheses check according to the hypotheses formulated in section 3.1. Except for grocery stores in survey area 1 (South Lower Saxony), the tables show both the conditional logit model (CL) and the nested logit model (NL). In all cases, the log likelihood of the nested logit model is (at least slightly) higher than for the conditional logit model. Furthermore, the inclusive value of the online channel (channel utility) is significant in the CE and furniture models. Thus, the nested logit results may be regarded as the more reliable, except for the grocery stores in South Lower Saxony, where no nested logit model was estimated as no online purchases were recorded.

With respect to grocery stores in both survey areas, the coefficient of the (naturally logged) number of articles is significant and positive but below one ( $\beta_1 = 0.645$  in survey area 1,  $\beta_1 = 0.893$  in survey area 2). Thus, an increase in a store's assortment increases its choice probability but in a disproportionate (degressive) way. This may be interpreted as diminishing marginal utility of assortment due to increasing search and decision costs for the consumers, as outlined by Huff (1962), and confirms previous findings with respect to physical store choice (see section 2.1). Thus, hypotheses  $H_{1a}$  (“An increasing assortment increases the likelihood of choosing a store”) and  $H_{1b}$  (“The positive effect of assortment is degressive”) are confirmed for grocery stores. The same is true for hypothesis  $H_{2a}$  (“Increasing travel time decreases the likelihood of choosing a store”), as the travel time coefficient is significant and negative ( $\beta_2 = -0.268$  and  $-0.335$ , respectively). Increasing travel time decreases the likelihood of choosing a grocery store, which is also congruent with previous findings. As, at the time of the survey, grocery online stores were only available in survey area 2, hypothesis  $H_3$  (“Increasing delivery charges decrease the

Table 7: Discrete Choice analysis for grocery stores by survey area – Dependent variable: Store choice probability

	Survey area 1		Survey area 2	
	CL		CL	NL
ln(number of articles)	0.645*** (0.098)		0.893*** (0.051)	0.893*** (0.050)
Travel time	-0.268*** (0.009)		-0.335*** (0.005)	-0.335*** (0.004)
Delivery charges			-1.258*** (0.102)	-1.138*** (0.164)
Click and collect	-0.277 (0.279)		0.102 (0.125)	0.101 (0.128)
<i>Control variables</i>				
Dummy Aldi	0.587** (0.244)		1.297*** (0.116)	1.298*** (0.124)
Dummy Edeka	-0.364* (0.191)		-0.071 (0.104)	-0.070 (0.100)
Dummy Kaufland	0.564** (0.222)		0.527*** (0.142)	0.528*** (0.138)
Dummy Lidl	0.179 (0.270)		0.942*** (0.125)	0.943*** (0.132)
Dummy Nahkauf	-0.415 (0.325)		0.031 (0.180)	0.032 (0.188)
Dummy Netto	-0.047 (0.263)		-0.222 (0.146)	-0.222 (0.152)
Dummy Norma			0.171 (0.327)	0.172 (0.328)
Dummy Penny	0.133 (0.291)		0.271** (0.132)	0.272* (0.140)
Dummy Real	0.511* (0.283)		-0.012 (0.139)	-0.012 (0.133)
Dummy Rewe	-0.077 (0.183)		0.106 (0.112)	0.107 (0.112)
Dummy organic food retailer	0.624* (0.323)		0.536*** (0.144)	0.537*** (0.146)
<i>Channel utility</i>				
IV online				0.146 (0.692)
Observations	628		2,333	2,333
Log Likelihood	-1,618.208		-6,974.784	-6,974.363

Note: \*p<0.1; \*\*p<0.05; \*\*\*p<0.01; CL = conditional logit, NL = nested logit

likelihood of choosing a store”) is only checkable in the Middle Upper Rhine Region. As the corresponding coefficient is significant and negative ( $\beta_3 = -1.138$ ), this hypothesis is also confirmed, which is in line with previous single-firm and experimental studies (see section 2.2). Thus, both considered transport-related transaction costs during grocery shopping are found to have the expected negative influence on (spatial) shopping behavior.

In absence of physical grocery stores providing an integrated online shop (according to the present definition), hypothesis  $H_4$  (“An integrated online shop increases the likelihood of choosing a store”) is dropped for grocery stores. However, in contrast to expectations, no positive effect of the availability of the “click and collect” service on store choice is found, as the corresponding coefficients are not significant in both survey areas ( $\beta_5 = -0.277$  and 0.101, respectively). Thus, hypothesis  $H_5$  (“Providing the ‘click and collect’ service increases the likelihood of choosing a store”) is rejected for grocery stores. Providing “click and collect” does not significantly increase a grocery store’s utility, which might be explained in a twofold way. First, it is likely that consumers want to inspect grocery products before buying (at least fresh food such as fruits and vegetables), which is not enabled by the “click and collect” service. Second, whilst the online shop for pre-ordering is open all the time, the physical stores are not, and thus picking up the order is bound by the opening hours of the respective store. Therefore, this service does not increase consumer flexibility with respect to grocery shopping trips (Dannenberg et al. 2016). Consequently, “click and collect” does not reduce consumer transaction costs to a great

Table 8: Discrete Choice analysis for consumer electronics stores by survey area – Dependent variable: Store choice probability

	Survey area 1		Survey area 2	
	CL	NL	CL	NL
ln(number of articles)	0.433*** (0.099)	0.665** (0.277)	0.196*** (0.060)	0.211*** (0.046)
Travel time	-0.133*** (0.009)	-0.134*** (0.010)	-0.121*** (0.004)	-0.119*** (0.006)
Delivery charges	-1.263*** (0.111)	-2.403* (1.310)	-0.764*** (0.045)	-0.705*** (0.075)
Int. online shop	0.828** (0.418)	0.620 (0.518)	0.454 (0.384)	0.621 (0.386)
Click and collect	-1.110*** (0.382)	-1.120** (0.446)	0.262 (0.221)	0.077 (0.253)
<i>Control variables</i>				
Dummy EP	-1.050* (0.609)	-0.816 (0.716)		
Dummy Euronics	-1.012** (0.424)	-1.218* (0.641)	0.515* (0.293)	0.310 (0.308)
Dummy Expert	-0.812** (0.414)	-1.323 (0.821)	1.662*** (0.336)	1.428*** (0.374)
Dummy Media Markt	-0.433 (0.409)	-0.921 (0.828)	1.724*** (0.315)	1.479*** (0.353)
Dummy Saturn	-0.445 (0.419)	-0.960 (0.862)	1.926*** (0.316)	1.661*** (0.357)
Dummy Amazon	-6.343*** (1.280)	-9.623** (3.975)	-0.231 (0.636)	-0.516 (0.584)
Dummy eBay	-1.930** (0.793)	-2.055 (1.921)	1.179** (0.458)	0.843* (0.434)
Dummy not full range	-1.132*** (0.331)	-1.186*** (0.424)	0.094 (0.300)	-0.067 (0.319)
<i>Channel utility</i>				
IV online		2.302 (1.542)		0.907*** (0.089)
Observations	338	338	1,269	1,269
Log Likelihood	-721.994	-720.269	-3,118.658	-3,118.292

Note: \*p<0.1; \*\*p<0.05; \*\*\*p<0.01; CL = conditional logit, NL = nested logit

extent. In the store choice model for survey area 2, the channel utility (IV online) is not significant. However, as some dummy control variables are significant (e.g., *Aldi*, *Kaufland*), these model results also suggest chain-specific utilities, which are outside the scope of this study.

The models for consumer electronics stores show similar results with respect to the hypotheses  $H_{1a}$ ,  $H_{1b}$ ,  $H_{2a}$ , and  $H_3$ . The assortment indicator is also significant and below one ( $\beta_1 = 0.665$  and  $0.211$ , respectively), however, the coefficient is much smaller than for grocery stores. This might be explained by the large variance of the number of articles between physical stores especially when comparing physical stores with online stores (see section 4.1). Confronted with shops providing several hundred thousand or even millions of products, consumer search and decision costs may increase heavily, and there might be no difference in utility as to whether an online shop provides, say, 3,000,000 or 4,000,000 products. The influence of transport-related transaction costs borne by consumers is found to be negative as well. Increasing travel time ( $\beta_2 = -0.134$  and  $-0.119$ , respectively) and delivery charges ( $\beta_3 = -2.403$  and  $-0.705$ , respectively) decrease store choice probability. Surprisingly, there is no significant and positive effect for either an integrated online shop, or for providing “click and collect”. Thus, hypotheses  $H_4$  and  $H_5$  must be rejected for CE stores. This cannot be explained in the same way as with respect to grocery stores. Perhaps these services cannot be regarded as a competitive advantage in CE retailing anymore, as cross-channel integration is extremely pervasive in this retail sector. For instance, the dominant big-box chains (e.g., *Media Markt*, *Saturn*, *Expert*) are cross-channel retailers without exception (see table 4 in section 4.1). The coefficient of the channel utility is positive in both survey areas and significant in the second survey area. There are also significant control variables, especially with respect

Table 9: Discrete Choice analysis for furniture stores by survey area – Dependent variable: Store choice probability

	Survey area 1		Survey area 2	
	CL	NL	CL	NL
ln(number of articles)	0.004 (0.066)	0.044 (0.089)	-0.006 (0.022)	0.058* (0.030)
Travel time	-0.060*** (0.005)	-0.058*** (0.006)	-0.065*** (0.003)	-0.067*** (0.003)
Delivery charges	-0.089*** (0.008)	-0.074*** (0.014)	-0.082*** (0.003)	-0.063*** (0.004)
Int. online shop	0.843*** (0.259)	0.713*** (0.258)	2.948*** (0.537)	1.964*** (0.437)
Click and collect	0.904* (0.478)	0.860* (0.484)	-1.205*** (0.414)	-0.302 (0.272)
<i>Control variables</i>				
Dummy Dänisches Bettenlager	-1.337*** (0.477)	-1.241** (0.495)	-1.322*** (0.394)	-1.689*** (0.392)
Dummy IKEA	2.176*** (0.459)	2.066*** (0.488)	3.293*** (0.384)	2.744*** (0.380)
Dummy Möbel Heinrich	1.214* (0.665)	1.077 (0.663)		
Dummy Möbel Höffner			2.482*** (0.378)	1.941*** (0.376)
Dummy Mömax			0.385 (0.375)	-0.004 (0.371)
Dummy Poco	-0.252 (0.664)	-0.301 (0.685)	-2.132*** (0.679)	-1.971*** (0.587)
Dummy SB Möbel Boss	-1.692*** (0.597)	-1.639*** (0.605)		
Dummy Sconto	-1.232** (0.599)	-1.178* (0.607)		
Dummy Roller			-0.175 (0.398)	-0.643 (0.394)
Dummy XXXLutz	-0.405 (0.876)	-0.239 (0.741)	1.766*** (0.352)	1.129*** (0.356)
Dummy Amazon	-1.148* (0.669)	-1.056 (0.756)	0.618** (0.268)	0.279 (0.281)
Dummy eBay	0.304 (0.631)	0.105 (0.770)	0.705** (0.327)	0.414 (0.304)
<i>Channel utility</i>				
IV online		0.781*** (0.151)		0.629*** (0.032)
Observations	220	220	912	912
Log Likelihood	-654.873	-653.437	-2,932.960	-2,883.416

Note: \*p<0.1; \*\*p<0.05; \*\*\*p<0.01; CL = conditional logit, NL = nested logit

to specific chains, however, these effects obviously differ between the survey areas. This might indicate differences of perceived chain utility between the two regions, which will not be examined in the present investigation.

The results for furniture stores differ significantly from those for grocery and CE stores. The coefficients of assortment are near zero ( $\beta_1 = 0.044$  and  $0.058$ , respectively) and not significant in survey area 1. Thus, hypotheses  $H_{1a}$  and  $H_{1b}$  may be confirmed only partially. Obviously, with respect to furniture stores, the number of articles does not play a decisive role in (spatial) shopping behavior. Instead, most chain dummies are significant, with partially consistent influences (signs) in both survey areas, which suggests that the effect of chain utilities exceeds the effect of assortment. This might be explained on the basis of the main difference between most furniture retailers and the other retail sectors in this study. Most furniture retailers are furniture designers and manufacturers at the same time, and thus, sell products of their own individual brands (e.g., *IKEA*, *Höffner*). Competition in furniture retailing might be driven primarily by consumer preferences for a specific brand rather than assortment size. Additionally, as the presence of an integrated online shop has a significant and positive impact on store choice ( $\beta_4 = 0.731$  and  $1.964$ , respectively), there is another obvious difference concerning the role of cross-channel integration. Furniture retailing companies profit from being cross-channel retailers. Thus,

Table 10: Hypothesis tests by retail sector

	Related variable	Related coefficient	Expected sign/value	Retail sector		
				GRO stores	CE stores	FURN stores
$H_{1a}$	$A_j$	$\beta_1$	+	confirmed	confirmed	partly
$H_{1b}$	$A_j$	$\beta_1$	$0 < \beta_1 < 1$	confirmed	confirmed	partly
$H_{2a}$	$t_{ij}$	$\beta_2$	-	confirmed	confirmed	confirmed
$H_{2b}$	$t_{ij}$	$\beta_2$	$ \beta_2  \text{ GRO} >$ $ \beta_2  \text{ CE} >$ $ \beta_2  \text{ FURN}$		confirmed	
$H_3$	$d_j$	$\beta_3$	-	confirmed	confirmed	confirmed
$H_4$	$OS_j$	$\beta_4$	+	not relevant	rejected	confirmed
$H_5$	$CC_j$	$\beta_5$	+	rejected	rejected	partly

Note: GRO = groceries, CE = consumer electronics, FURN = furniture

hypothesis  $H_4$  is explicitly confirmed for furniture retailers. However, there is no such impact of providing “click and collect”, as the corresponding coefficient is significant and positive only in survey area 1 ( $\beta_5 = 0.860$  and  $-0.302$ , respectively). Thus, hypothesis  $H_5$  is confirmed only partially. However, this might be a collinearity effect as, in survey area 2, most furniture stores in the choice set provide both an integrated online shop and “click and collect”. In survey area 1, the two related dummy variables show a middle to strong relationship ( $\chi^2 = 5,276.8$ , Cramers  $V = 0.775$ ,  $p < 0.01$ ), but a nearly perfect correlation in survey area 2 ( $\chi^2 = 64,083.0$ , Cramers  $V = 0.943$ ,  $p < 0.01$ ). The negative influences of travel time ( $\beta_2 = -0.058$  and  $-0.067$ , respectively) and delivery charges ( $\beta_3 = -0.074$  and  $-0.063$ , respectively) are found for furniture stores as well, which leads to the confirmation of hypotheses  $H_{2a}$  and  $H_3$ . The channel utility is significant in both survey regions.

In order to check hypothesis  $H_{2b}$  (“*The negative effect of travel time declines with decreasing purchasing frequency*”), the coefficients of the models for all three retail sectors have to be compared, as they refer to the same unit (car driving time in minutes). Taking into account shopping frequency (see section 4.1), the negative effect of travel time must be the highest for grocery stores, followed by consumer electronics and furniture. As the absolute values of the grocery store coefficients ( $\beta_2 = -0.268$  and  $-0.335$ , respectively) are higher than for CE ( $\beta_2 = -0.134$  and  $-0.119$ , respectively) and the lowest values are found for furniture stores ( $\beta_2 = -0.058$  and  $-0.067$ , respectively), hypothesis  $H_{2b}$  is clearly confirmed. Thus, as outlined in central place theory (Christaller 1933, Brown 1993), consumer sensitivity towards travel time is found to be dependent on shopping frequency of the considered goods.

## 5 Conclusions and limitations

The aim of the present study was to identify the main drivers of store choice given the availability of both in-store and online shopping alternatives. The econometric analysis of store choice confirms several assumptions of classical retail location theory. As stated in central place theory and the Huff model, travel time to physical stores reduces consumer utility and store choice probability, respectively. As assumed in central place theory, the consumer sensitivity towards travel time decreases with decreasing purchase frequency of the desired goods. Delivery charges also decrease the likelihood of choosing a store. Regarding both types of channel-specific transaction costs, the results also confirm previous findings from single-firm studies and stated choice experiments towards multi-channel shopping behavior. However, additional results differ between retail sectors. The positive effect of assortment on condition of diminishing marginal utility, as stated in the Huff model, is confirmed for grocery stores and consumer electronics stores. With respect to furniture stores, assortment size plays a negligible role. Instead, store choice in this sector appears to be primarily driven by chain or brand preferences. The impact of cross-channel integration on store choice (assuming the reduction of consumer transaction

costs) is considerably lower than expected. While furniture retailers profit from enabling cross-channel shopping, there is no such competitive advantage found for grocery and CE retailers.

The results demonstrate that online shopping in grocery retailing still plays a subordinate role, although its market share is increasing yearly ([Handelsverband Deutschland, IFH Köln GmbH 2019](#)) and might be boosted in the context of the SARS-CoV-2 pandemic ([Bundesverband E-Commerce und Versandhandel Deutschland e.V. 2020](#)). However, the important role of spatial accessibility and delivery charges as well as the current lack of impact of the “click and collect” service question whether multi- and cross-channel shopping in grocery retailing will, at least in Germany, become standard in the near future. In contrast, online shopping is well established in consumer electronics retailing. Looking in more detail at the results, this is primarily due to one online-only retailer (*Amazon*), whilst the online shops of established CE chains (e.g., *Media Markt*, *Saturn*, *Expert*) play a minor role and their cross-channel services obviously do not affect store choice significantly. These results match the findings of a model-based location analysis of German CE retailers, in which cross-channel integration, ceteris paribus, was not found to increase store turnovers ([Wieland 2018b](#)). This suggests that, in CE retailing, pure *online* shopping is of high significance, but *cross-channel* shopping is not. On the contrary, cross-channel shopping opportunities are very relevant in furniture retailing, where competition appears to be primarily driven by qualitative aspects (brands and chains, respectively). However, the common denominator of all three retail sectors is the important influence of travel time and delivery charges when consumers choose a channel and store, respectively.

The current study contributes to the literature on both spatial shopping behavior and multi-channel shopping behavior; most notably, as there has been an obvious lack in the previous literature towards incorporating multi- and cross-channel shopping behavior into store choice models. This is surprising given the fact, as outlined in this study, that both approaches towards explaining consumer behavior are implicitly based on the same rationale, namely utility maximization and minimization of shopping transaction costs, respectively. Thus, from the theoretical perspective, one has to conclude that multi- and cross-channel shopping behavior does not contradict the main thoughts of classical retail location theory, although this family of theories originally does not consider shopping channels other than physical stores. Spatial shopping behavior, which is traditionally a key element of retail geography and its related disciplines, should be regularly explored, incorporating other shopping channels, especially e-shopping, in the future.

From a practical perspective, incorporating multi- and cross-channel shopping into store choice models is important because these models play an important role in different fields of applied geography. First, such models are utilized in retail location planning for estimating potential sales of new store locations ([Levy et al. 2019](#), [Reynolds, Wood 2010](#)). Second, in several European countries, in the case of proposed new large-scale retail stores (or shopping malls), spatial planning authorities request a *retail impact assessment*. These studies evaluate the impact of the proposed retail project on town centers by estimating customer or purchasing power flows induced by the new retail location. Typically, quantitative store choice models are utilized for this estimation ([Khawaldah et al. 2012](#), [Müller-Hagedorn 2020](#)). As the online channel has a noticeable share in most retail sectors, the explanatory power of these models may be improved substantially by including the online shopping alternatives.

However, the present study faces some limitations. First, the study focuses on a collection of explanatory variables which represent different types of channel-and store-specific shopping transaction costs. Of course, this examination does not cover all aspects of transaction costs and leaves other aspects of shopping behavior aside. Subsequent studies should perform deeper investigations of further explanatory variables, such as the influence of chains, pricing strategies or subjective characteristics of consumers (motives and attitudes). Second, even though the econometric analysis is based on a nested model, the analysis does not account for the true sequential order of the purchasing process, as the nested logit model has a purely statistical motivation rather than a behavioral. In fact, a purchasing process consists of gathering information, comparing the alternatives, and

the shopping decision itself, as well as post-purchase behavior. The present study only incorporates variables which reflect the store-based opportunities of information search (such as an integrated online shop) but does not examine whether these opportunities are taken up by the consumers. Subsequent studies should investigate the whole consumer purchasing process which may include both spatial and non-spatial elements.

Third, there are several retail sectors which would benefit from a repetition of the analysis. Fourth, although this study includes two survey areas covering both urban and rural municipalities, not all types of spatial configurations were incorporated. Such analyses should be repeated in more extreme spatial settings, such as megacities or rural, deprived regions. Fifth, as in any study utilizing Discrete Choice models, the impact of explanatory variables relates to store choices and shopping trips, respectively, rather than expenditures. As the descriptive results show, purchase and expenditure shares are not to be confused. As a consequence, further studies should distinguish between shopping decisions and the corresponding expenditure and sales, respectively. This is even more important as average expenditures in online stores are mostly substantially lower than in physical stores.

### Funding

Funded by the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation) – 402130768.

### Acknowledgements

The author would like to thank Anne Auinger and Anjolie Kappler for their work with respect to data collection and data processing as well as to Jeff Burrowes for proofreading. Many thanks also to several retail companies and public administrations who provided store data for the data analysis.

### References

- Baumol WJ, Ide EA (1956) Variety in Retailing. *Management Science* 3: 93–101. [CrossRef](#).
- Baviera-Puig A, Buitrago-Vera J, Escriba-Perez C (2016) Geomarketing Models in Supermarket Location Strategies. *Journal of Business Economics and Management* 17: 1205–1221. [CrossRef](#).
- Beckers J, Cárdenas I, Verhetsel A (2018) Identifying the geography of online shopping adoption in Belgium. *Journal of Retailing and Consumer Services* 45: 33–41. [CrossRef](#).
- Briesch RA, Chintagunta PK, Fox EJ (2009) How Does Assortment Affect Grocery Store Choice? *Journal of Marketing Research* 46: 176–189. [CrossRef](#).
- Broniarczyk SM, Hoyer WD (2010) Retail Assortment: More  $\neq$  Better. In: Krafft M, Mantrala MK (eds), *Retailing in the 21st Century: Current and Future Trends* (2nd ed.). American Marketing Association, 271–284. [CrossRef](#).
- Brown S (1993) Retail location theory: evolution and evaluation. *The International Review of Retail, Distribution and Consumer Research* 3: 185–229. [CrossRef](#).
- Bundesinstitut für Bau-, Stadt- und Raumforschung (2017) Laufende Raumbeobachtung – Raumabgrenzungen – Siedlungsstrukturelle Kreistypen. [https://www.bbsr.bund.de/BBSR/DE/forschung/raumbeobachtung/Raumabgrenzungen/deutschland/kreise/Kreistypen4/kreistypen\\_node.html](https://www.bbsr.bund.de/BBSR/DE/forschung/raumbeobachtung/Raumabgrenzungen/deutschland/kreise/Kreistypen4/kreistypen_node.html)
- Bundesverband E-Commerce und Versandhandel Deutschland e.V. (2020) E-Commerce-Plus von 9,2 Prozent im 1. Halbjahr 2020 – dauerhaft mehr E-Commerce beim "Täglichen Bedarf". Press release from July 5, 2020. [https://www.bevh.org/fileadmin/content/05\\_presse/Pressemitteilungen.2020/200705'\\_-PM\\_Zahlen.2.\\_Quartal.2020\\_-im\\_Online-Handel.pdf](https://www.bevh.org/fileadmin/content/05_presse/Pressemitteilungen.2020/200705'_-PM_Zahlen.2._Quartal.2020_-im_Online-Handel.pdf)

- Cao L, Li L (2015) The Impact of Cross-Channel Integration on Retailers' Sales Growth. *Journal of Retailing* 91: 198–216. [CrossRef](#).
- Chintagunta PK, Chu J, Cebollada J (2012) Quantifying Transaction Costs in Online/Offline Grocery Channel Choice. *Marketing Science* 31: 96–114. [CrossRef](#).
- Chocarro R, Cortiñas M, Villanueva ML (2013) Situational variables in online versus offline channel choice. *Electronic Commerce Research and Applications* 12: 347–361. [CrossRef](#).
- Christaller W (1933) *Die zentralen Orte in Süddeutschland: eine ökonomisch-geographische Untersuchung über die Gesetzmäßigkeit der Verbreitung und Entwicklung der Siedlungen mit städtischen Funktionen*. Gustav Fischer
- Clarke G, Thompson C, Birkin M (2015) The emerging geography of e-commerce in British retailing. *Regional Studies, Regional Science* 2: 371–391. [CrossRef](#).
- Converse PD (1949) New Laws of Retail Gravitation. *Journal of Marketing* 14: 379–384. [CrossRef](#).
- Creditreform Boniversum GmbH (2018) Boniversum Verbraucherumfrage 11/2018. "Click & Collect" – Verbreitung und Nutzung. [https://www.boniversum.de/wp-content/uploads/2018/11/Boniversum\\_bevh\\_Studie\\_Click-Collect.pdf](https://www.boniversum.de/wp-content/uploads/2018/11/Boniversum_bevh_Studie_Click-Collect.pdf)
- Croissant Y (2020) Estimation of Random Utility Models in R: The mlogit Package. *Journal of Statistical Software* 95. [CrossRef](#).
- Dannenber P, Franz M, Lepper A (2016) Online einkaufen gehen – Einordnung aktueller Dynamiken im Lebensmittelhandel aus Perspektive der geographischen Handelsforschung. In: Franz M, Gersch I (eds), *Online-Handel ist Wandel*, Volume 24 of *Geographische Handelsforschung*. MetaGIS, 133–156. [https://opus.bibliothek.uni-wuerzburg.de/files/18076/GHF24\\_Franz\\_Gersch.Online-Handel\\_Wandel.pdf](https://opus.bibliothek.uni-wuerzburg.de/files/18076/GHF24_Franz_Gersch.Online-Handel_Wandel.pdf)
- Desmeules R (2002) The Impact of Variety on Consumer Happiness: Marketing and the Tyranny of Freedom. *Academy of Marketing Science Review* 12: 1–33
- Doherty NF, Ellis-Chadwick F (2010) Internet retailing: the past, the present and the future. *International Journal of Retail and Distribution Management* 38: 943–965. [CrossRef](#).
- Flavián C, Gurrea R, Orús C (2020) Combining channels to make smart purchases: The role of webrooming and showrooming. *Journal of Retailing and Consumer Services* 52: 101923. [CrossRef](#).
- Gatta V, Marcucci E, Maltese I, Iannaccone G, Fan J (2021) E-Groceries: A Channel Choice Analysis in Shanghai. *Sustainability* 13. [CrossRef](#).
- Güßefeldt J (2002) Zur Modellierung von räumlichen Kaufkraftströmen in unvollkommenen Märkten. *Erdkunde* 56: 351–370. [CrossRef](#).
- GfK (2019) GfK Purchasing Power for Retail Product Lines. <https://geodata.gfk.com/-availability/gfk-purchasing-power-for-retail-product-lines/>
- Ghosh A (1986) The Value of a Mall and Other Insights from a Revised Central Place Model. *Journal of Retailing* 62: 79–97
- Giraud T (2019) osrm: Interface Between R and the OpenStreetMap-Based Routing Service OSRM. R package version 3.3.2. <https://CRAN.R-project.org/package=osrm>
- Greene WJ (2012) *Econometric Analysis* (7th ed.). Pearson
- Hahn Gruppe (2019) Retail Real Estate Report Germany 2019/2020. [https://www.hahnag.de/index.php?eID=tx\\_nawsecuredl&u=0&g=0&t=1612448394&hash=b3f16edebe8a2ead25aba918c63eb8adc91a95aa&file=media/hahn/Downloads\\_PDF/-Publikationen/Research/Hahn\\_Retail\\_Real\\_Estate\\_Report\\_2019\\_web.pdf](https://www.hahnag.de/index.php?eID=tx_nawsecuredl&u=0&g=0&t=1612448394&hash=b3f16edebe8a2ead25aba918c63eb8adc91a95aa&file=media/hahn/Downloads_PDF/-Publikationen/Research/Hahn_Retail_Real_Estate_Report_2019_web.pdf)



- Handelsverband Deutschland, IFH Köln GmbH (2019) Online Monitor 2019. [https://einzelhandel.de/images/publikationen/Online\\_Monitor\\_2019\\_HDE.pdf](https://einzelhandel.de/images/publikationen/Online_Monitor_2019_HDE.pdf)
- Heinemann GM (2015) Location-based Services – Rettungsanker für den stationären Einzelhandel? *Marketing Review St. Gallen* 3: 58–66. [CrossRef](#).
- Hillier A, Smith T, Cannuscio CC, Karpyn A, Glanz K (2015) A Discrete Choice Approach to Modeling Food Store Access. *Environment and Planning B: Planning and Design* 42: 263–278. [CrossRef](#).
- Hillier A, Smith TE, Whiteman ED, Chrisinger BW (2017) Discrete Choice Model of Food Store Trips Using National Household Food Acquisition and Purchase Survey (FoodAPS). *International Journal of Environmental Research and Public Health* 14: 1133. [CrossRef](#).
- Hlavac M (2018) stargazer: Well-Formatted Regression and Summary Statistics Tables. R package version 5.2.2. <https://CRAN.R-project.org/package=stargazer>
- Hsiao MH (2009) Shopping mode choice: Physical store shopping versus e-shopping. *Transportation Research Part E: Logistics and Transportation Review* 45: 86–95. [CrossRef](#).
- Huff DL (1962) *Determination of Intra-Urban Retail Trade Areas*. University of California
- Huff DL (1964) Defining and Estimating a Trading Area. *Journal of Marketing* 28: 34–38. [CrossRef](#).
- Kacen JJ, Hess JD, Kevin Chiang WY (2013) Bricks or Clicks? Consumer Attitudes toward Traditional Stores and Online Stores. *Global Economics and Management Review* 18: 12–21. [CrossRef](#).
- Khawaldah H, Birkin M, Clarke G (2012) A review of two alternative retail impact assessment techniques: the case of Silverburn in Scotland. *The Town Planning Review* 83: 233–260. [CrossRef](#).
- Kubis A, Hartmann M (2007) Analysis of Location of Large-area Shopping Centres. A Probabilistic Gravity Model for the Halle–Leipzig Area. *Jahrbuch für Regionalwissenschaft* 27: 43–57. [CrossRef](#).
- Lademann RP (2007) Zum Einfluss von Verkaufsfläche und Standort auf die Einkaufswahrscheinlichkeit. In: Schuckel M, Toporowski W (eds), *Theoretische Fundierung und praktische Relevanz der Handelsforschung*. DUV, 143–162. [CrossRef](#).
- Lange S (1973) *Wachstumstheorie zentralörtlicher Systeme : eine Analyse der räumlichen Verteilung von Geschäftszentren*, Volume 5 of *Beiträge zum Siedlungs- und Wohnungswesen und zur Raumplanung*. Institut für Siedlungs- und Wohnungswesen der Universität Münster
- Levy M, Weitz BA, Grewal D (2019) *Retailing Management* (10th ed.). McGraw-Hill
- Lösch A (1944) *Die räumliche Ordnung der Wirtschaft* (2nd ed.). Gustav Fischer
- Luce RD (1959) *Individual choice behavior: a theoretical analysis*. Wiley
- Marcucci E, Gatta V, Le Pira M, Chao T, Li S (2021) Bricks or clicks? Consumer channel choice and its transport and environmental implications for the grocery market in Norway. *Cities* 110: 103046. [CrossRef](#).
- Marino G, Zotteri G, Montagna F (2018) Consumer sensitivity to delivery lead time: a furniture retail case. *International Journal of Physical Distribution and Logistics Management* 48: 610–629. [CrossRef](#).
- McFadden D (1974) Conditional logit analysis of qualitative choice behavior. In: Zarembka P (ed), *Frontiers in econometrics*. Academic Press, 105–142
- McKinsey & Company (2019) Our new research shows more than 70% of shoppers willing to try cross-channel shopping. Press release from March 19, 2019.

- <https://www.mckinsey.com/business-functions/marketing-and-sales/solutions/periscope/news/press-releases/more-than-70-percent-of-shoppers-willing-to-try-cross-channel-shopping-methods-in-new-research-from-periscope-by-mckinsey#>
- Melis K, Campo K, Breugelmans E, Lamey L (2015) The Impact of the Multi-channel Retail Mix on Online Store Choice: Does Online Experience Matter? *Journal of Retailing* 91: 272–288. [CrossRef](#).
- Mensing M (2018) Lebensmittel-Onlinehandel – Alternative zur zukünftigen Versorgung der Bevölkerung ländlicher Räume? PhD thesis, Rheinisch-Westfälische Technische Hochschule Aachen, <https://publications.rwth-aachen.de/record/758124>
- Müller-Hagedorn L (2020) Einzelhandelsgutachten sind eine schwierige Dienstleistung. In: Roth S, Horbel C, Popp B (eds), *Perspektiven des Dienstleistungsmanagements. Aus Sicht von Forschung und Praxis*. Springer Gabler, 105–125. [CrossRef](#).
- Nakanishi M, Cooper LG (1974) Parameter Estimation for a Multiplicative Competitive Interaction Model – Least Squares Approach. *Journal of Marketing Research* 11: 303–311. [CrossRef](#).
- Nakanishi M, Cooper LG (1982) Simplified estimation procedures for MCI models. *Marketing Science* 1: 314–322. [CrossRef](#).
- Orpana T, Lampinen J (2003) Building Spatial Choice Models from Aggregate Data. *Journal of Regional Science* 43: 319–348. [CrossRef](#).
- Popkowski Leszczyc PTL, Sinha A, Sahgal A (2004) The effect of multi-purpose shopping on pricing and location strategy for grocery stores. *Journal of Retailing* 80: 85–99. [CrossRef](#).
- R Core Team (2019) R: A Language and Environment for Statistical Computing. <https://www.R-project.org/>, Vienna, Austria
- Reigadinha T, Godinho P, Dias J (2017) Portuguese food retailers – Exploring three classic theories of retail location. *Journal of Retailing and Consumer Services* 34: 102–116. [CrossRef](#).
- Reilly WJ (1931) *The Law of Retail Gravitation*. Knickerbocker Press
- Ren F, Kwan MP (2009) The Impact of Geographic Context on E-Shopping Behavior. *Environment and Planning B: Planning and Design* 36: 262–278. [CrossRef](#).
- Reutterer T, Teller C (2009) Store format choice and shopping trip types. *International Journal of Physical Distribution and Logistics Management* 37: 695–710. [CrossRef](#).
- Reynolds J, Wood S (2010) Location decision making in retail firms: evolution and challenge. *International Journal of Retail and Distribution Management* 38: 828–845. [CrossRef](#).
- Schmid B, Axhausen KW (2019) In-store or online shopping of search and experience goods: A hybrid choice approach. *Journal of Choice Modelling* 31: 156–180. [CrossRef](#).
- Singleton AD, Dolega L, Riddlesden D, Longley PA (2016) Measuring the spatial vulnerability of retail centres to online consumption through a framework of e-resilience. *Geoforum* 69: 5–18. [CrossRef](#).
- Statista (2019) Anzahl der Einzelartikel in Lebensmittel-Discountern in Deutschland nach Unternehmen im Jahr 2019. <https://de.statista.com/statistik/daten/studie/479391/-umfrage/artikel-in-lebensmittel-discountern-in-deutschland/>
- Statista (2020) E-commerce share of total global retail sales from 2015 to 2023. <https://www.statista.com/statistics/534123/e-commerce-share-of-retail-sales-worldwide/>
- Stepper M (2016) Innenstadt und stationärer Einzelhandel – ein unzertrennliches Paar? Was ändert sich durch den Online-Handel? *Raumforschung und Raumordnung Spatial Research and Planning* 74: 151–163. [CrossRef](#).

- Suel E, Polak JW (2018) Incorporating online shopping into travel demand modelling: challenges, progress, and opportunities. *Transport Reviews* 38: 576–601. [CrossRef](#).
- Suárez-Vega R, Gutiérrez-Acuña JL, Rodríguez-Díaz M (2015) Locating a supermarket using a locally calibrated Huff model. *International Journal of Geographical Information Science* 29: 217–233. [CrossRef](#).
- Tennekes M (2019) tmaptools: Thematic Map Tools. R package version 2.0-1. <https://CRAN.R-project.org/package=tmaptools>
- Teo TS, Yu Y (2005) Online buying behavior: a transaction cost economics perspective. *Omega* 33: 451–465. [CrossRef](#).
- The Nielsen Company (Germany) GmbH (2016) Germany 2016. Retail, Consumers, Advertising. [https://www.nielsen.com/wp-content/uploads/sites/3/2019/04/Nielsen-Booklet-Handel-Verbraucher-Werbung\\_2016.pdf](https://www.nielsen.com/wp-content/uploads/sites/3/2019/04/Nielsen-Booklet-Handel-Verbraucher-Werbung_2016.pdf)
- Tihi B, Oruc N (2012) Competitive Location Assessment – the MCI Approach. *South East European Journal of Economics and Business* 7: 35–49. [CrossRef](#).
- Timmermans H (1993) Retail Environments and Spatial Shopping Behavior. In: Gärling T, Golledge RG (eds), *Behavior and Environment*, Volume 96 of *Advances in Psychology*. North-Holland, 342–377
- Train K (2009) *Discrete Choice Models with Simulation* (2nd ed.). Cambridge University Press
- Vroegrijk M, Gijsbrechts E, Campo K (2013) Close Encounter with the Hard Discounter: A Multiple-Store Shopping Perspective on the Impact of Local Hard-Discounter Entry. *Journal of Marketing Research* 50: 606–626. [CrossRef](#).
- Wickham H (2019) httr: Tools for Working with URLs and HTTP. R package version 1.4.1. <https://CRAN.R-project.org/package=httr>
- Wieland T (2015) *Räumliches Einkaufsverhalten und Standortpolitik im Einzelhandel unter Berücksichtigung von Agglomerationseffekten: Theoretische Erklärungsansätze, modellanalytische Zugänge und eine empirisch-ökonomische Marktgebietsanalyse anhand eines Fallbeispiels aus dem ländlichen Raum Ostwestfalens/Südniedersachsens*, Volume 23 of *Geographische Handelsforschung*. MetaGIS. [https://opus.bibliothek.uni-wuerzburg.de/files/18075/GHF\\_Buch23.pdf](https://opus.bibliothek.uni-wuerzburg.de/files/18075/GHF_Buch23.pdf)
- Wieland T (2018a) A Hurdle Model Approach of Store Choice and Market Area Analysis in Grocery Retailing. *Papers in Applied Geography* 4: 370–389. [CrossRef](#).
- Wieland T (2018b) Standorterfolg in Zeiten des Onlinehandels – Aufbau, Ergebnisse und planungsbezogene Implikationen einer modellgestützten Standortanalyse für die Elektrofachmärkte in der Region Mittlerer Oberrhein. *Berichte. Geographie und Landeskunde* 92: 5–26
- Wieland T (2019) MCI2: Market Area Models for Retail and Service Locations. R package version 1.1.2. <https://CRAN.R-project.org/package=MCI2>
- Zhai Q, Cao X, Mokhtarian PL, Zhen F (2017) The interactions between e-shopping and store shopping in the shopping process for search goods and experience goods. *Transportation* 44: 885–904. [CrossRef](#).
- Zhen F, Du X, Cao J, Mokhtarian PL (2018) The association between spatial attributes and e-shopping in the shopping process for search goods and experience goods: Evidence from Nanjing. *Journal of Transport Geography* 66: 291–299. [CrossRef](#).

