MICHIGAN
ROSS SCHOOL OF BUSINESS

Working Paper

Consumer Search Behavior on the Mobile Internet: An Empirical Analysis

Stephan Daurer
Baden-Wuerttemberg Cooperative State University

Dominik Molitor
Ludwig-Maximilians-University of Munich

Martin Spann
Ludwig-Maximilians-University of Munich

Puneet Manchanda
Stephen M. Ross School of Business
University of Michigan

UNIVERSITY OF MICHIGAN

Consumer Search Behavior on the Mobile Internet:
An Empirical Analysis

Stephan Daurer[1]

Dominik Molitor[2]

Martin Spann[2]

Puneet Manchanda[3]

May 2014

This version: 5 May 2015

[1] DHBW Ravensburg, Baden-Wuerttemberg Cooperative State University, Marienplatz 2, 88212 Ravensburg, Germany, *daurer@daurer.net*

[2] Ludwig-Maximilians-University of Munich, Institute of Electronic Commerce and Digital Markets, Geschwister-Scholl-Platz 1, 80539 Munich, Germany, *molitor@bwl.lmu.de*, *spann@spann.de*

[3] Stephen M. Ross School of Business, University of Michigan, 701 Tappan St., Ann Arbor, MI 48109-1234, *pmanchan@umich.edu*

Consumer Search Behavior on the Mobile Internet:
An Empirical Analysis

Abstract

The increasing diffusion of smartphones and tablet computers has facilitated access to product information by providing Internet access anywhere and at any time. As a result, consumers are increasingly using the mobile Internet to search for product information to help them in their purchase decisions. However, there is very little documentation of how, where and when consumers actually carry out such search. Using location-based data from a leading European product information and barcode-scanning app that contains more than 80 million observations, this study provides insights using actual consumer search behavior. The results show that consumer search on the mobile Internet is not bound to store opening hours and is likely to happen to a large extent as ongoing search during consumption. Furthermore, consumers' geographic mobility is positively correlated and previous search experience is negatively correlated with their search intensity. Finally, access to more types of information via search results, especially product related information, reduces further search on price information, suggesting that product information content can lower price sensitivity.

Keywords: *Consumer Search, Product Information, Mobile Marketing, Location-Based Services, Big Data*

Introduction

The increasing diffusion of smartphones and tablet computers has facilitated access to product information by providing Internet access anywhere and at any time.[1] For consumers, the availability of product-related information is crucial, especially during the information search phase of the purchase decision process (Shugan 2004). Examples of relevant information are prices and discounts in online and offline stores (e.g., retail prices in their vicinity), detailed product information (e.g., provided by the manufacturer or by neutral third parties) and consumer reviews (i.e., user-generated content). A consumer can obtain this information by searching the mobile web, but the use of smartphone apps is becoming the default mechanism for such searches. These apps use a combination of barcode scanning and location-based services to provide relevant information, e.g., showing only stores near the consumer when she is carrying out a price comparison. These apps are thus suited to deliver context-specific and hence more relevant information to consumers, leading to lower information asymmetries between consumers and merchants. Another way to think about mobile search is that it aids the convergence of the offline and online worlds as barcode scanning apps help consumers while shopping offline at physical stores and also provide information on online retailers. Surveys suggest that most consumers carry out product information search at or close to the point of purchase, both in space and time (TNS infratest 2013), leading to significant changes in consumer behavior. For example, according to a recent study, 75% of all smartphone users cancelled a purchase in a store shortly before check-out to buy the product somewhere else (IntelliAd 2014).

---

[1] The number of mobile Internet users is expected to grow substantially over the next few years PwC (2013). The increasing importance of mobile aspects for marketers is currently reflected by the shift in advertising spending. eMarketer (2013) forecasts that 36.3 % of digital ad spending around the world will go toward mobile formats in 2017 (as opposed to just 4.6 % in 2011).

While there is a lot of anecdotal and survey-based knowledge on product information search via mobile devices, there is very little documentation of how, where and when consumers actually carry out search. In this paper, we use a very large and unique behavioral dataset from a leading provider of mobile product information applications. The dataset contains more than 80 million observations across more than 2.5 million individual consumers. In addition, the dataset provides us granular detail on consumer location, allowing us to analyze location-based search behavior at an individual level. The focus of our research is to investigate how closely product information search via mobile devices is linked to potential purchases, both on the time and space dimensions. We are also able to zero in on both the quantity and quality of product information searched within the mobile app.

Our paper adds to the small but growing field of empirical research dedicated to analyzing consumer behavior on the mobile Internet.[2] This research has documented that mobile devices are used for both research and purchase (Ghose, Han, and Xu 2013) and that their use has both benefits and costs. For example, the form factor of mobile devices (smaller screen size, more difficult text entry, limited battery life) could lead to higher search costs (Ghose and Han 2011) while their portability could lead to lower search costs (Ghose, Goldfarb, and Han 2013). Other research has shown that context matters for mobile computing in general (Chen and Kotz 2000). Location tends to be the most important contextual factor (Ilarri, Mena, and Illarramendi 2010) as it affects information needs (Hinze, Chang, and Nichols 2010). Other research also suggests that location matters during mobile search as search results that are close to consumers' home (Ghose, Goldfarb, and Han 2013) or to their current location (Liu, Rau, and Gao 2010, p. 370) are more relevant than

---

[2] Our focus is search on the mobile Internet and its relationship to the context. There is a very large extant literature on search, especially in economics, that focuses on quantifying search costs and their impact e.g., Stigler (1961), Weitzmann (1979), Stiglitz (1989). More recent work has focused on search in traditional (computer based) online settings e.g., Brynjolfsson and Smith (2000), Spann and Tellis (2006), Ratchford et al. (2003), Kim et al. (2010).

other search results, leading to changing purchase intentions (Daurer, Molitor, and Spann 2012). There is also some evidence that mobile search behavior impacts the response to mobile advertising behavior (Goh, Chu, and Soh 2009). For the firm, location of mobile consumers can be beneficial for targeting (Luo et al. 2014).

Our research complements and extends the above research by using behavioral data comprising search behavior on the mobile Internet to examine how context – location and time – affect search. Specifically, we describe where and when mobile search for product-related information using location-based barcode scanning apps takes place. We also analyze factors related to the search intensity. Finally, we focus on the outcome of search by describing how consumers react to the search results and how individual information choice behavior differs by information type. We do this via a series of panel-data models, first focused on search intensity and then on information choice after consumer search.

We have several novel findings. First, mobile search does not seem to be focused at the point of purchase. Specifically, we find that search volumes are not different between Sundays (when stores in our market are closed) and weekdays. This suggests consumers also carry out mobile search in many situations other than shopping (e.g., while consuming the product). Second, geographic travel (mobility), the availability of specific types of product information and contextual factors (e.g., economic surroundings, competition, and weather) influence search intensity. Third, we show that consumers search using different types of information. This information choice depends on various factors including context, product category, user experience and availability of (other) information. A novel finding here is that access to more types of information, especially product-related information, reduces search on price information, suggesting that information content can lower price sensitivity.

The remainder of this study is organized as follows. Section 2 describes the large behavioral dataset and additional (contextual) data that we include in our analyses, including a brief description of the data generating process by illustrating the behavior of a typical user on a given day. Section 3 outlines our model. Section 4 details the results and their managerial implications. We conclude with a discussion of the general opportunities for research on mobile search in the future in Section 5.

Setting and Data

Setting and Variables

Our setting is the Northern European consumer market with data being provided by one of Europe's largest and leading providers of product information and barcode scanning apps.[3] The provider offers product information that is available via a location-based smartphone app as well as on a website. In this study, we focus on user behavior related to the app. As depicted in figure 1, the smartphone app provides a search screen (see figure 1 step 1) where product barcodes can be scanned. On the results screen, customers are able to browse through three types of available information (which varies across products). First, there is product information (see figure 1 step 2); second, there is price information (of offers in the vicinity or offers by online retailers (see figure 1 step 3); third, there is information based on user-generated content such as reviews and ratings (see figure 1 step 4). The process always starts with a search query (step 1) that returns the "results" screen. By default, the tab with the product information is displayed first. Then app users can click on price information (step 3), UGC information (step 4) and back to product information (step 2) and so on.

---

[3]   We are not allowed to disclose the name of the app and the company due to confidentiality and privacy concerns.

Insert Figure 1 about here

Because the data was generated through the smartphone app, it is possible to identify individual users. When a smartphone app is downloaded and installed, a unique anonymous user identity number (user id) is generated and transmitted with each search query emanating from that smartphone. In addition, to identify search queries that belong to one search session there is a session id. A session is defined as a series of consecutive activities (i.e., search queries or clicks) until the visitor stops or is idle for at least 20 minutes.[4] This definition of session is consistent with that used in other studies (Goh, Chu, and Soh 2009, Montgomery et al. 2004).

The behavior of the consumer on the search results screen is tracked and available to us. Examples of tracked behavior are (a) a click on the price tab to view alternative offers (including pricing) for the focal product in the vicinity, (b) a click on a product guide (a product guide is a product segment specific guide provided by the app producer, e.g., a wine guide), (c) a click on user reviews, (d) a click on test reports (i.e., third-party test reports) etc. Behavior (a) is related to product information, (b) to price information and (c) and (d) to user-generated content. Other behavior that is also available to us is information on the product category of the product that was searched (e.g., food, electronics, media, etc.), the device type (e.g., iPhone), the app's version number, the country setting (e.g., Germany), and the language.

The geographic location[5] of the user at the time of the search query is also included in the data. Modern smartphones use multiple technologies to determine the current location of the device. The most important technology is the Global Positioning System (GPS) that works with

---

[4] Please note that a session with constant ongoing activity might last longer than 20 minutes.
[5] The location is specified via longitude and latitude. Note that the location is not always available as some users deactivate location services on their smartphones or because a position could not be determined (due to technical reasons). In our analyses we only consider records with available location information (approximately 60% of total observations).

signals that are received from satellites. Hightower and Borriello (2001) report an outdoor

accuracy of one to five meters for this technology. However, the accuracy of GPS radio signals is

lower indoors (due to distortions induced by the structure) leading to lower resolution. In a recent

study on the accuracy of indoor positioning the measured root mean squared two dimensional

positioning error for state-of-the-art GPS receivers was between 5 and 10 meters for wooden

buildings and about 15 meters for large structures such as shopping malls (Kjærgaard et al. 2010).

The dataset contains all mobile customer search data of the smartphone app over a period

of 12 months (July 2011 through June 2012). There are two matching subsets of the data set: (1)

mobile customer search query data (N > 11 million): search queries and information on search

results, and (2) mobile customer click[6] behavior (N > 69 million) in the app. In most cases (72.5%),

the search queries reflect the European Article Number (EAN) that can be scanned using the

barcode on the packaging of products. In addition, manual text entry is possible and therefore some

queries contain plain text (product name, brand name, etc.). Usually the screen that yields the

results for a search query contains various different types of information. If a search query yields

a result and the corresponding product is categorized (in 70.4 % of all search queries) the product

category is tracked. This allows us to distinguish between high and low involvement, as well as

food and non-food products in our analyses.

For our analyses we merge the search data with the click data and filter on the following

criteria. Since the majority of the app users are from Germany (95.5% of all observations), we only

consider data that was generated using a smartphone in a geo-location within Germany. As we use

the two-dimensional geographical coordinates represented by longitude and latitude as defined in

WGS84 (NIMA 1984), we apply the well-established approximation method of the minimum

---

[6] Some authors refer to a click performed on a touch screen of a mobile device as "tap." For clarity reasons, we use the term "click" throughout this study.

bounding rectangle (Papadias and Theodoridis 1997) by filtering longitude and latitude values to be within an area. The minimum bounding rectangle of Germany is defined by the four extreme points in each geographic direction.

To analyze the data on a session level we determine the distance traveled during a session, the duration of a session and session fractions of categorical variables (e.g., fraction of grocery products in a session). To calculate the travel distance within a session we use the sequence of available data points within the session and calculate the distances between each (sequential) pair of points and then aggregate all these distances. The distance between two points is calculated as the orthodromic distance, which is the shortest surface distance between two points on the surface of a sphere (as opposed to the Euclidean distance which is used in two dimensional geometries). The variable session distance is used as a proxy for consumers' mobility during their search.

We also control for users' past experience with the app. We calculate the number of previous clicks made by a user within the app until the time of the session or click that we analyze. We use this number as a proxy for search experience with the barcode-scanning and product information app. Finally, we calculate the number of clicks of a user within a session as a measure of activity (i.e., search intensity). The logarithm of this variable serves as the dependent variable when we analyze search intensity on a session level. In addition to the behavioral data from the log files of the app, we generate supplementary variables. We generate dummies for time aspects (e.g., day of the week), to classify products of search queries (e.g., grocery products), and to control for the availability of information on the results screen (e.g., price information, product guide).

We then augment the behavioral data with a set of contextual data. First, we obtain data to control for demographic (e.g., population density) and economic (e.g., importance of the trade

sector) differences in different regions. Each data point is matched to the corresponding county[7] or postal region[8] based on the GPS location information. Second, we include weather as a contextual variable via the use of daily weather data in the local region of the user.[9] The data was obtained from the national weather office that maintains 78 weather stations distributed over the whole country (Deutscher Wetterdienst 2013). Using the location, we attribute the weather for that data point to the nearest weather station. Third, we create the shopping environment in terms of physical stores around each data point. We calculate the distance to the nearest store (e.g., supermarket, discounter or electronics store) in a radius of five kilometers of each data point (i.e., search location) using data on location from the nation's top ten retail companies based on revenue plus further publicly available geo-coded data.[10] In total we use data on about 40,000 physical store locations. The distance to the nearest point of interest (POI) is only considered if it is 5 km or smaller. Otherwise we assume that the presence of POI is too far to have any impact on consumer behavior. Note that we only consider stores that match the product category of the product that was searched. This variable (distance to nearest store) allows us to analyze search behavior depending on a consumer being close to a store (maybe even in a store) in contrast to locations that are not related to shopping.

Table 1 gives an overview of all variables that we use in our analyses.

Insert Table 1 about here

---

[7] The political entity that is comparable to a US county is called "Landkreis" in Germany.
[8] There are 95 postal regions in Germany, identified by the first two digits of the German postal code.
[9] Previous work has documented effect of weather on consumer behavior e.g., weather has an effects on sales Steele (1951) and on consumer spending Murray et al. (2010). In addition, the use of mobile devices may also be impacted by weather conditions.
[10] The data was collected from the companies' websites and from www.pocketnavigation.de, a forum for pocket navigation users.

Typical Search Behavior

To get a better understanding of the data-generating process and the richness of our data, we illustrate the search behavior of one typical app user for a single day selected at random. As noted above, the data that is collected through the app includes information about the device that is used (e.g., iPhone), the products that are searched for, the type of search (e.g., barcode scan vs. manual text entry) and the location where the search took place. Since the location of each activity is known, we are able to visualize the consumer search path on a map. Connecting different search locations enables us to generate path data that could provide valuable insights for marketers (e.g., Hui, Fader, and Bradlow (2009).

While one customer's behavior for one single day is unlikely to be representative of the behavior of the entire sample, the path data are useful as they give us some insights in to the usage of the app. Our chosen consumer is mainly active in a rather rural environment with medium sized towns. In this county the average discretionary income per citizen is low. The GDP per citizen is close to the average in the country. During the focal month this consumer used the app for 106 search queries on an iPhone. In figure 2 we illustrate the search behavior of one day, Saturday, June 9, 2012.

Insert Figure 2 about here

In the morning, shortly before 9 o'clock, the consumer scans a spare part for a mowing head of a grass trimmer or brush cutter ("Stihl Mähfaden 1.6 mm"; see figure 2, location 1). This search and all the subsequent ones of this consumer on that day are conducted with the barcode scanning function of the app on an iPhone. Over the course of the month 72% of all searches were performed by scan and 28% by clicks on related products in the app. Since the location shows

multiple scans on different days and because the surroundings of this location are a residential area, we assume this location is the consumer's home.

Around noon, the consumer scans another do-it-yourself product: a power drill ("Metabo 1010-Watt-Multihammer"). After that the consumer scans "The Ultimate Bourne Collection (Blu-ray)." The next search activity takes place at 5:09 p.m. in another location: the city center (see figure 2 location 2). The center is about 4.6 km (approx. 9 minutes by car or a 50 minute walk) away from the first location. Because the two relevant data points are more than three hours apart, we cannot determine the mode of transportation used by the consumer to get to the city center. However, by examining the roads and traffic patterns, walking would take almost an hour and the consumer would have to go through an industrial area and along a highway. The only form of public transportation in that area is an inconvenient regional train that would take 37 minutes to traverse this distance. So it is likely that the consumer either used a personal car or a taxi.

At 5:10 p.m. there is a scan for disposable drinking cups ("HS Becher Klar - 0,3L") and for a five-liter steel beer keg ("Bitburger Pils Fassdose pfandfrei"). A possible explanation for these kinds of needs might be the 14th European Championship for men's national football teams (UEFA Euro 2012). The group stage took place from June 8 to June 19. The German team was playing (vs. Portugal) in the evening at 8:45 p.m. of that Saturday of the consumer search. It happened to be the first game for the German national team in this tournament. Maybe the consumer prepared for a party to watch football together with friends. On his or her way home the consumer scanned a gardening product: a set of weed removal brushes ("Unkrautbürsten").

The weather on that Saturday was slightly colder (13.3 °C) compared to the rest of the month (average temperature at 15.8 °C). However, it was dry (zero precipitation) and the sunshine duration was very high that day with 7 hours 48 minutes versus this month's average of 4 hours

24 minutes. There was a moderate breeze (6.9 m/s; average at 3.6 m/s) with phases of high wind and moderate gale (15.1 m/s; average at 10.8 m/s) at times. Despite the windy conditions, the overall weather was suitable to work in the garden. This might be an explanation of the scanning pattern of this consumer on this Saturday. Because of the temperature and the wind the consumer probably had to find a place indoors to watch the game in the evening.

Descriptive Statistics

Our data set comprises 80 million observations that stem from 2.5 million individual users of a product information and barcode-scanning app. Table 2 gives an overview of the entire data.

Insert Table 2 about here

A session may consist of one to many search queries. On average there are close to two search queries per session although there are many sessions with only one query. Subsequent to a search query there is a results screen displayed in the app. On this screen either zero, one or multiple clicks are possible. On average a search session contains 20.78 clicks. The mean distance that a user travels is 558 meters while using the app in multiple locations during one session. When a session contains multiple data points (search queries or clicks), we calculate the duration of this session (Session_Duration). If a session contains only one observation the duration is zero. Because some variables are only available on a session level the sample size (N) is not the same for all variables. Also, in a few cases context data were not available. Obviously, previous user experience (measured in number of clicks within the app) is available for users that were active multiple times. For all distance variables we consider a reasonable action radius of 5 km and for the maximum plausible duration of the app usage in one session the cut-off is one hour.

When looking at search queries (scans) by product category we find that food and beverages accounts for 40.0% of searches and is the top category, followed by media (8.4%), drugstore articles (7.3%), non-food products like do-it-yourself equipment and tobacco (6.3%) and electronics (4.4%). The fashion category has a very low share of all scanned products (0.4%). As a point of reference, the revenue share of offline retail in Germany across categories is food and beverages (including tobacco) at 38.4%, drugstore articles at 15.7% and fashion at 14% (BTE 2008). A possible explanation for this variance is that price comparisons are difficult for fashion items as they typically do not carry a barcode or EAN, making them hard to search about via smartphones (Daurer et al. 2013).

As noted earlier, our data are from across the entire country. Regressing app usage (search and click volume) by county on population density, number of students in a county, importance of the trade sector and average discretionary income suggests that the main driver is population density. The number of students in a county has a positive and significant impact on the number of search queries that are produced in an area in a given time period. Further, the importance of the trade sector in a county is related to the app usage. Thus, in geographies with more trade businesses product information apps are more heavily used. The average discretionary income per citizen also has a positive sign. This may be somewhat unexpected for pure price comparison apps, but in our case the app provides product information and user-generated content in addition to price. Table 3 gives an overview on drivers of product information app usage.

Insert Table 3 about here

An examination of temporal usage patterns also provides interesting results (see figure 3 for charts on temporal app usage). When we compare the magnitude of activity within the app with the popular weekdays for shopping (VuMA 2015), we find a similar pattern except for the weekend

(figure 3). The app is widely used on Sundays, even more than on some of the weekdays. This is quite surprising as most stores in Germany are closed on Sunday by law.[11] The average app usage over the course of a day deviates from the popular shopping times in Germany as well. While the most popular shopping time is from 10:00 to 12:00 in the morning (GfK 2013), the app usage peaks at 5 p.m. (figure 4). As on weekends, we find significant activity within the app outside of business hours on all days of the week. In terms of session duration, sessions are longer on Saturdays and Sundays (figure 5). However, the distance that consumers travel while using the app (search session level) is shorter on Sundays (figure 6). A possible explanation for these patterns is that is that consumers have more time on the weekends leading to longer session duration but tend to use the app predominantly from one static location (e.g., at home).

Insert Figures 3-6 about here

Figures 7 to 9 depict user search and click activity for a subsample of approx. 335,000 observations for the greater Munich area in June 2012. Figure 7 depicts the clicks by information type category across time of the day. As can be seen, most activity is during the day, but substantial activity takes place outside store opening hours that are in greater Munich from 8 a.m. to 8 p.m. Figure 8 shows this effect by day of the week (please note that with few exceptions, stores are closed on Sundays). Figure 9 depicts the overall user activity (i.e. clicks) by day vs. night (upper panels) as well as by shopping days (Monday through Saturday) vs. Sunday for clicks on the price tab only (lower panels).

Insert Figures 7-9 about here

---

[11] The only exceptions to this rule are small stores/kiosks that are attached to gas stations or for stores that are on the premises of a large train station.

Model

Our objective is to investigate the drivers of location-based information search. We therefore follow a modeling strategy of first determining the drivers of search intensity within a search session. We then examine the choices made by consumers on the types of information to process once the search results are available. Our modeling approach is descriptive in nature (Reiss 2011) and is typical of post-hoc analyses in marketing (e.g., the analysis of shopping baskets as in Manchanda, Ansari, and Gupta (1999) or the analysis of offline (Larson, Bradlow, and Fader 2005) or online consumer paths (Montgomery et al. 2004)).

We begin by specifying a fixed-effects regression model (to account for unobserved consumer heterogeneity) at the user (i) session (j) combination (ij). Specifically, we model consumers' log-transformed search intensity (i.e., the number of clicks per session) as a function of their offline travel (i.e., session distance), previous experience with the app (i.e., number of previous clicks), the opening hours (i.e., closed vs. open), product category (i.e., groceries vs. non-groceries), air temperature, available choice options (i.e., price information and UGC information), location-specific demographics (i.e., fuel price level and the importance of the trade sector), day of the week as time dummies (i.e., Monday to Sunday) as well as monthly dummy variables (i.e., January to December). Consumers' search intensity is measured as the number of clicks per session, which is a count variable. The aggregation on a user session level is necessary to be able to account for product category dependent differences that we include as the fraction of the focal product category in relation to all products that were searched in a session. We use the log of the number of clicks per session as the dependent variable in order to deal with outliers (e.g., heavily search intensive sessions from a few users). The log number of clicks per session of consumer i in session j specified as:

$$
\begin{aligned}
\log(\text{SESSION\_CLICKS}_{ij}) = {}& \alpha + \beta_1 \log(\text{SESSION\_DIST}_{ij}) + \beta_2 \text{EXPERIENCE}_{ij} \\
& + \beta_3 \text{CLOSED}_{ij} + \beta_4 \text{CATEGORY}_{ij} + \beta_5 \text{AVAILABLE\_OPTIONS}_{ij} \\
& + \beta_6 \text{ECON\_DEMOGR}_{ij} + \beta_7 \text{AIR\_TEMPERATURE}_{ij} + \beta_8 \text{DAY\_WEEK} \\
& + \beta_9 \text{MONTH} + \delta_i + \varepsilon_{ij}
\end{aligned}
$$

(1)

where δ_i is the individual-level fixed effect, α the intercept and ε_{ij} the error term.

Next, we model consumer search behavior within the app. Previous research has shown that consumers typically focus on three types of information during search – product (e.g., Daurer et al. 2013), price (e.g., Dickson and Sawyer 1990) and user-generated content (e.g., Dellarocas 2003, Hennig-Thurau and Walsh 2003). The app essentially provides information in similar buckets as the outcome of search – product information, price information, user-generated reviews, other – depending on the product that was searched (scanned). The consumer can choose to click on one or more types of information.

Given that the type of information is not constant for each search, we model the consumer decision to click on the three resulting sets of information as independent decisions. Clicking on each type of information is represented as a binary discrete choice e.g., the consumer either clicks on price information or not. Thus, we specify three discrete choice models capturing the binary choices on (1) product-, (2) price- and (3) UGC-related information. We denote these choice options as k (with $k \in \{1, 2, 3\}$). Consumer i's utility from choosing choice option k of product m is therefore given by

(2)
$$
\begin{aligned}
& U_{imtk} = v_{imtk} + e_{imtk} \\
& y_{imtk} = 1 \text{ if } U_{imtk} > 0 \\
& y_{imtk} = 0 \text{ if } U_{imtk} \leq 0
\end{aligned}
$$

where v_{imtk} is the deterministic part of i's utility and e_{imtk} is the stochastic part of i's utility (McFadden 1974). e_{imtk} contains factors that are influencing the utility U that are either non-

systematic or random. As is usual for binary choice models, we assume that consumers choose option k if their utility is larger than zero. We assume that the error term ε_{imt} is i.i.d. and follows a standard type I extreme value distribution. Thus, the probability of consumer i choosing option k when looking at product m at time t is given by:

$$(3) \qquad \Pr_{imtk} = \frac{\exp(U_{imtk})}{1 + \exp(U_{imtk})}$$

The deterministic component of consumer i's utility is a function of their mobility (session distance), previous experience, the opening hours dummy (closed vs. open), store distance (i.e., the distance to the next discounter and supermarket), category (i.e., groceries vs. non-groceries), interaction between store distance and category, location-specific demographics (e.g., population density, discretionary income per citizen and fuel price level), air temperature, day of the week dummies as well as monthly dummies. In addition, we include control variables that account for the available choice options (e.g., product, price and user-generated content). This results in the following equation:

$$(4) \qquad \begin{aligned} U_{imtk} = {} & \delta_i + \beta_1 SESSION\_DIST_{imt} + \beta_2 EXPERIENCE \\ & + \beta_3 CLOSED_{imt} + \beta_4 NEAREST\_STORE\_DIST_{imt} \\ & + \beta_5 NEAREST\_STORE\_DIST * GROCERIES_{imt} + \beta_6 GROCERIES_{imt} \\ & + \beta_7 AVAILABLE\_OPTIONS_{imt} + \beta_8 ECON\_DEMOGR_{imt} \\ & + \beta_9 AIR\_TEMPERATURE_{imt} + \beta_{10} DAY\_WEEK_{imt} + \beta_{11} MONTH_{imt} + \varepsilon_{imt} \end{aligned}$$

where the δ_i represent individual fixed-effects to control for unobserved consumer heterogeneity.

The results of the Hausman (1978) specification tests provide support for using models based on fixed effects (session level: $\chi^2 = 20{,}112.70$, prob $> \chi^2$: 0.0000; individual level: $\chi^2 = 518.11$ (1), 5,442.37 (2), 1,008.74 (3), prob $> \chi^2$: 0.0000).

Results and Managerial Implications

Intensity of Search

In this section, we present the results of the fixed effects regression based on consumers' search intensity. The first row of the results in table 4 indicates that the session distance is significant and positively correlated with the logarithm of the number of clicks per session (i.e., consumers' search intensity). We expect a 0.16 percent increase in search intensity when session distance increases by 1 kilometer[12]. Geographically mobile consumers are thus more active when it comes to mobile search. This result is consistent with previous literature (Ghose and Han 2011, p. 1683). Moreover, the results show that usage experience is significantly and negatively correlated to the search intensity. Therefore, past experience is a negative predictor for future usage. This might be explained by the fact that experience leads to a more target-oriented and more efficient search app usage, resulting in lower search intensity. Furthermore, learning effects might cause consumers to decrease their usage intensity over time after multiple sessions (Bucklin and Sismeiro 2003).

Perhaps the most surprising result is that consumers' search intensity is about 7 percent[13] higher when the stores are closed. This is perhaps the first documentation of the fact that mobile search is not all geared towards the purchase decision but does represent a broader search pattern (this is akin to the notion of "ongoing search" proposed in Bloch et al. (1986)). Thus, practices such as "showrooming" (Luger 2013) may be less prevalent than expected. These results are also consistent with other results in digital domains (e.g., Manchanda et al. (2006) show that online advertising affects purchase behavior of current users after a temporal gap and typically not via an immediate click-through). We speculate that we find this because of two types of factors. First, it

---

[12] $1.01^{\beta_1}=1.01^{0.16162}=1.00161$

[13] $\exp(0.06744)=1.06977$

is possible that a lot of search occurs around the time of product consumption (e.g., to learn about a product's nutrition facts or to prepare for the next purchase) and not so much at the point of sale. Another explanation might be that consumers spend time on the weekend to view or to produce user-generated content (e.g., product reviews).

The category-specific variable grocery variable also has a significantly positive effect. A larger share of groceries increases consumers' search intensity by almost 19 percent[14] compared to non-groceries. The weather-specific variable, air temperature, is positively related to consumers' search intensity. A similar effect for sunshine has been documented in previous research (Hirshleifer and Shumway 2003).

Insert Table 4 about here

Our other variables also deliver some interesting insights. For instance, we control for the available information options. Both price information and UGC information are positively related to the search intensity. This finding suggests that the availability of additional information options (next to product information that is almost always available) also leads to increased search intensity. In addition, explanatory power can also be attributed to the demographic variables population density, discretionary income per citizen and fuel price level. Population density and the discretionary income have a positive impact on the search intensity. Fuel price level instead has a negative impact. While this finding is counter-intuitive it might be explained by the fact that as fuel prices increase, consumers shift to one-stop shopping formats and reduce the monthly number of shopping trips more than they reduce purchase volume (Ma et al. 2011). Consumers'

---

[14] exp(0.17393)=1.18997

search intensity differs between the days of the week. Consumers are most active on Mondays compared to the other days of the week (the reference is Monday).

Choice of Information

In this section, we present and discuss the results of the discrete choice models with fixed effects. As noted above, consumers' information choice behavior is analyzed on the individual level. Table 5 presents the estimation results of the three discrete choice models. Table 6 shows the marginal effects based on the estimation results.

Insert Table 5 and Table 6 about here

Recall that product information is almost always available at the end of a search query. Therefore, we do not include R_Prod as explanatory variable. Regarding the models on price- and UGC-information, we only consider observations where the respective information is actually available. The estimates show that the session distance has a significantly positive impact on consumers' choice to search for product, price and UGC-related information. Increasing the distance by one unit increases the choice probability towards searching for product, price and UGC-related information by about one percent. This result is qualitatively similar to the result obtained at the session level. Comparably, consumers' previous experience is negatively related to the tendency to choose product, price and UGC-related information. The impact of the opening hours on consumers' choice to search differs by the choice type. Similar to the results on the session level, product and UGC choices are positively related to the opening hours, meaning that consumers are more likely to choose product and UGC-related information when the stores are closed. If stores are closed, the likelihood for searching for product-related information increases by 0.76 percent, for UGC-related information by 1.8 percent, while the likelihood for price-related

information decreases by 0.42 percent. This can again be explained by the ongoing (post-purchase) information search. However, information on prices is negatively affected by the opening hours. This finding indicates that price information might be more important during the opening hours and thus around the time of purchase.

The distance to the nearest store has a positive effect on consumers' search behavior regarding product, price and UGC-related information, suggesting that consumers do not necessarily search at the point of sale. However, when consumers search for groceries, the distance to the nearest store has a negative and significant effect on the choice of product and UGC information. This means that the closer the proximity between consumers' location and the next store the higher the probability that product and UGC information are chosen when it comes to searching for groceries. This finding is similar to the one in Ghose et al. (2013), who find that consumers prefer offers close to their home location. The category-specific grocery variable is significantly negative related to consumers' choice to search for product-, price- and UGC-related information. Our results show that the likelihood of searching for the three different information types decreases by between 2.5 and 6.8 percent if users are searching for groceries.

The population density as well the discretionary income per citizen has a significant negative impact on the likelihood to choose product, price and UGC-related information. In areas where consumers are more affluent, these particular types of product information seem to be less important. This is particularly relevant for price information. Similar to the session level, the fuel price is significant but has another sign compared to the population density and the discretionary income. However, the fuel price level leads to a higher choice probability regarding product, price and UGC information. Weather also has an impact on consumers' search behavior. In contrast to the results at the session level, the sign of the coefficient for air temperature indicates that it is

significantly negative related to all three information types. This is correlationally consistent with previous research showing that lower temperatures are associated with negative affective states (Bell and Baron 1977).

We also control for the information options available, the day of the week as well as the months. The availability of choice and search options (i.e., price, UGC, guide, report, other) is negatively related to all three information types, given the focal information is available (i.e., product, price or UGC). This means that the propensity to choose a certain type of information decreases if further information is available.[15] For instance, price information is less likely to be clicked on when product quality information is available. This is related to recent analytical research that proposes that for search goods especially lower quality firms are better off to display more quality related information as opposed to price information (Anderson and Renault 2013, p. 69). In addition, the impact of the day of the week varies between our three models. Compared to the reference Monday, Tuesdays and Wednesdays are positively related to product-, price- and UGC-related information. Apart from these days, the sign and direction of all other days of the week differ between the three different samples but are largely significant.

Managerial Implications

The results from our analyses have several implications for practitioners. First, there is a main effect of mobility. Geographically mobile users are more active in their search. Thus, firms need to make sure that they have a mobile presence in order to capitalize on decisions consumers make via this medium. Besides leading to purchases directly from the mobile platform, this could also result in increased offline purchasing (Pauwels et al. 2011). Second, our results indicate that product information and barcode scanning apps are used (spatially) outside of stores or shopping

---

[15] The only exception is "Other" in the price information sample.

areas (e.g., at home) to a large extent. Furthermore, we detect a considerable amount of search activity (temporally) outside of store opening hours. Thus, a significant amount of mobile search occurs away from the point and moment of purchase. This suggests that companies may be overstating the impact of mobile search at the point of purchase (Knight 2014; Luger 2013). Specifically, they may be missing out a large part of the market (the ongoing searches) by focusing their resources towards providing answers for searches at the point of purchase. Third, more experienced users of mobile search exhibit a lower intensity of search, independent from the type of information. Thus, this suggests a simple targeting strategy of focusing on current users by providing them rich information about the products as well as access to user-generated content, potentially leading to higher customer retention. Fourth, the finding demonstrating that clicks on price information are lower in the presence of other information should encourage firms to provide a lot of product information along with pointers to user-generated content. This could become a valuable tactic for lowering the price sensitivity of consumers in such competitive markets.

Conclusions, Limitations and Future Research

Our research adds to the small but growing set of studies focused on the mobile Internet. In this study, our focus is on consumer search behavior using behavioral data with location-based information. We illustrate the interplay between mobile online search and offline travel behavior and we show when and how consumers use a product information and barcode-scanning app. Our findings are novel and may be summarized as follows. First, mobile search is not bound to opening hours of physical stores (e.g., comparable search volume on Sundays and weekdays) indicating that consumers also search in many situations other than shopping (e.g., while consuming the product). Second, search intensity is influenced by geographic travel (mobility), the availability of

specific types of product information and contextual factors (e.g., economic surroundings, competition, and weather). Third, we show that consumers search using different types of information. This information choice depends on various factors including context, product category, user experience and availability of (other) information. A specific novel finding here is that access to more types of information, especially product related information, reduces search on price information, suggesting that information content can lower price sensitivity.

Overall, we contribute to the literature on multiple dimensions. We use actual behavioral data coupled with location data to describe consumer mobile search patterns. We provide evidence that context – location, time, availability of information choices etc. – matters for both search intensity and information choice. Finally, we show that the availability of different information choices affects the propensity to search on different dimensions in real settings.

Our analysis has some limitations, primarily driven by our data. First, the data come from one national market. Second, we only have data on what happens when the consumer uses the app. So actions such as search queries on a computer while at home or via a mobile browser do not appear in our data. Third, we do not have access to demographics of individual consumers. Fourth, our data do not go all the way to final purchase. Finally, since we conduct a post-hoc analysis in a non-experimental setting, statements about causal relationships are difficult.

In general, the availability of location data coupled with behavioral data opens up myriad research opportunities for managers and researchers. While we focus on product search, these types of data can be used to answer questions around assortment choice at the point of purchase, the positioning and location of products within a physical store and store choice. We hope that our study spurs future research into this important area.

References

Anderson, Simon P. and Régis Renault (2013), "The Advertising Mix for a Search Good," *Management Science*, 59 (1), 69–83.

Bell, Paul A. and Robert A. Baron (1977), "Aggression and Ambient Temperature: The Facilitating and Inhibiting Effects of Hot and Cold Environments," *Bulletin of the Psychonomic Society*, 9 (6), 443–445.

Bloch, Peter H., Daniel L. Sherrell, and Nancy M. Ridgway (1986), "Consumer Search: An Extended Framework," *Journal of Consumer Research*, 13 (1), 119–126.

BTE (2008), "Annual Statistics (Statistik-Report Textileinzelhandel Deutschland)," Cologne.

Bucklin, Randolph E. and Catarina Sismeiro (2003), "A Model of Web Site Browsing Behavior Estimated on Clickstream Data," *Journal of Marketing Research*, 40 (3), 249–267.

Chen, Guanling and David Kotz (2000), "A Survey of Context-Aware Mobile Computing Research," Time, College, Dartmouth, Hanover, NH.

Daurer, Stephan, Dominik Molitor, and Martin Spann (2012), "Measuring Individual Search Costs on the Mobile Internet," in *Proceedings of the 20th European Conference on Information Systems (ECIS)*.

——, ——, ——, and Puneet Manchanda (2013), "The Impact of Smartphones, Barcode Scanning, and Location-Based Services on Consumers' Search Behavior," in *Proceedings of the 34th International Conference on Information Systems (ICIS)*.

Deutscher Wetterdienst (2013), "Klimadaten der deutschen Messstationen. Tageswerte," (accessed August 28, 2013), [available at http://www.dwd.de/bvbw/appmanager/bvbw/dwdwwwDesktop?_nfpb=true&_pageLabel=_dwdwww_klima_umwelt_klimadaten_deutschland&activePage=&_nfls=false].

eMarketer (2013), "Mobile Expands Its Share of Worldwide Digital Ad Spend," (accessed December 19, 2013), [available at http://www.emarketer.com/Article/Mobile-Expands-Its-Share-of-Worldwide-Digital-Ad-Spend/1010170].

GfK (2013), "Distribution of purchases during the week by time (Germany)," Focus 41/2013, 2013-10-07, p. 82; accessed through Statista.com ID:274841 (accessed March 1, 2015), [available at http://de.statista.com/statistik/daten/studie/274841/umfrage/beliebteste-einkaufszeiten-in-deutschland].

Ghose, Anindya, Avi Goldfarb, and Sang P. Han (2013), "How Is the Mobile Internet Different? Search Costs and Local Activities," *Information Systems Research*, 24 (3), 613–631.

——— and Sang P. Han (2011), "An Empirical Analysis of User Content Generation and Usage Behavior on the Mobile Internet," *Management Science*, 57 (9), 1671–1691.

———, ———, and Kaiquan Xu (2013), "Mobile Commerce in the New Tablet Economy," in *Proceedings of the 34th International Conference on Information Systems (ICIS)*, 1–18.

Goh, Khim Y., Hunhong Chu, and Winnie Soh (2009), "Mobile Advertising. An Empirical Study of Advertising Response and Search Behavior," in *Proceedings of the 30th International Conference of Information Systems (ICIS)*, 1–17.

Hausman, Jerry A. (1978), "Specification Tests in Econometrics," *Econometrica*, 46 (6), 1251–1271.

Hightower, Jeffrey and Gaetano Borriello (2001), "Location Systems for Ubiquitous Computing," *IEEE Computer*, 34 (8), 57–66.

Hinze, Annika M., Carole Chang, and Dave M. Nichols (2010), "Contextual Queries Express Mobile Information Needs," in *Proceedings of the 12th International Conference on Human Computer Interaction with Mobile Devices and Services*, 327.

Hirshleifer, David and Tyler Shumway (2003), "Good Day Sunshine: Stock Returns and the Weather," *Journal of Finance*, 58 (3), 1009–1032.

Hui, Sam K., Peter S. Fader, and Eric T. Bradlow (2009), "Path Data in Marketing. An Integrative Framework and Prospectus for Model Building," *Marketing Science*, 28 (2), 320–335.

Ilarri, Sergio, Eduardo Mena, and Arantza Illarramendi (2010), "Location-Dependent Query Processing. Where We Are and Where We Are Heading," *ACM Computing Surveys*, 42 (3), 12:1.

IntelliAd (2014), "Distribution of showrooming in Germany," accessed through Statista.com ID:376218.

Kjærgaard, Mikkel B., Henrik Blunck, Torben Godsk, and Thomas Toftkjær (2010), "Indoor Positioning Using GPS Revisited," in *Pervasive Computing*. *Lecture Notes in Computer Science*. Berlin: Springer, 38–56.

Knight, Kristina (2014), "Placed: Showrooming now mainstream in auto sector. BizReport, Mobile Marketing," (accessed April 22, 2015), [available at http://www.bizreport.com/2014/01/placed-showrooming-now-mainstream-in-auto-sector.html].

Larson, Jeffrey S., Eric T. Bradlow, and Peter S. Fader (2005), "An Exploratory Look at Supermarket Shopping Paths," *International Journal of Research in Marketing*, 22 (4), 395–414.

Liu, Chengyi, Pei-Luen P. Rau, and Fei Gao (2010), "Mobile Information Search for Location-based Information," *Computers in Industry*, 61 (4), 364–371.

Luger, Josh (2013), "How The Mobile "Showrooming" Threat Is Changing The Retail Industry,"
(accessed April 22, 2015), [available at http://www.businessinsider.com.au/mobile-
showrooming-is-changing-retail-2013-5].

Luo, Xueming, Michelle Andrews, Zheng Fang, and Chee W. Phang (2014), "Mobile
Targeting," *Management Science*, 60 (7), 1738–1756.

Ma, Yu, Kusum L. Ailawadi, Dinesh K. Gauri, and Dhruv Grewal (2011), "An empirical
investigation of the impact of gasoline prices on grocery shopping behavior," *Journal of
Marketing*, 75 (2), 18–35.

Manchanda, Puneet, Asim Ansari, and Sunil Gupta (1999), "The "Shopping Basket". A Model
for Multicategory Purchase Incidence Decisions," *Marketing Science*, 18 (2), 95–114.

———, Jean-Pierre Dubé, Khim Y. Goh, and Pradeep K. Chintagunta (2006), "The Effect of
Banner Advertising on Internet Purchasing," *Journal of Marketing Research*, 43 (1), 98–108.

McFadden, Daniel (1974), "Conditional Logit Analysis of Qualitative Choice Behavior," in
Frontiers in Econometrics, P. Zarembka, ed. New York, 105–142.

Montgomery, Alan L., Shibo Li, Kannan Srinivasan, and John C. Liechty (2004), "Modeling
Online Browsing and Path Analysis Using Clickstream Data," *Marketing Science*, 23 (4), 579–
595.

Murray, Kyle B., Fabrizio Di Muro, Adam Finn, and Peter Popkowski Leszczyc (2010), "The
Effect of Weather on Consumer Spending," *Journal of Retailing and Consumer Services*, 17
(6), 512–520.

NIMA (1984), "National Imagery and Mapping Agency (NIMA) Technical Report 8350.2.
World Geodetic System 1984: Its Definition and Relationships with Local Geodetic Systems,"
US Department of Defense, St. Louis, MO, USA.

Papadias, Dimitris and Yannis Theodoridis (1997), "Spatial Relations, Minimum Bounding Rectangles, and Spatial Data Structures," *International Journal of Geographical Information Science*, 11 (2), 111–138.

Pauwels, Koen, Peter S. Leeflang, Marije L. Teerling, and K. E. Huizingh (2011), "Does Online Information Drive Offline Revenues?," *Journal of Retailing*, 87 (1), 1–17.

PwC (2013), "Worldwide Fixed Broadband and Mobile Internet Penetration from 2008 to 2017," Accessed through Statista (accessed January 5, 2014), [available at http://www.statista.com/statistics/280430/worldwide-fixed-broadband-and-mobile-internet-penetration/].

Reiss, Peter C. (2011), "Structural Workshop Paper - Descriptive, Structural, and Experimental Empirical Methods in Marketing Research," *Marketing Science*, 30 (6), 950–964.

Shugan, Steven M. (2004), "The Impact of Advancing Technology on Marketing and Academic Research," *Marketing Science*, 23 (4), 469–475.

Steele, A. T. (1951), "Weather's Effect on the Sales of a Department Store," *Journal of Marketing*, 15 (4), 436–443.

TNS infratest (2013), "Mobile Life. Mobile im Handel - Showrooming," (accessed March 1, 2015), [available at http://www.tns-infratest.com/presse/presseinformation.asp?prID=3205].

VuMA (2015), "Customers in Germany by day of week (groceries)," accessed through Statista.com ID:297624 (accessed March 1, 2015), [available at http://de.statista.com/statistik/daten/studie/297624/umfrage/umfrage-unter-kaufland-kunden-zu-einkaufstagen-von-lebensmitteln].

TABLE 1

Variable description of behavioral and contextual data[16]

Variable	Description
Session_Clicks	Number of clicks in the app within a session
Session_Distance	Distance traveled during a session (in meters)
Experience	Experience of a user with the app before the focal search session measured in clicks (in number of clicks generated by the user in the app in the past)
Closed	Store opening hours dummy (1 if the law of the state requires stores to be closed at the time of the observation)
Groceries	Dummy for product category (1 if groceries)
Groceries_Fraction	Fraction of grocery products scanned in a session relative to all scans in the same session
Monday, Tuesday, …	Dummies for the day of the week
R_Prod, R_Price, R_UGC, R_Guide, R_Report, R_Other	Search result dummies (1 if {information on product characteristics; price information; user-generated content; product guide; neutral test report by a third party; other information} was available)
R_Price_Fraction	Fraction of products with available price information in a session
R_UGC_Fraction	Fraction of products with available UGC information in a session
GDP_per_Citizen	Nominal gross domestic product (GDP) per citizen in a county (in €)
Importance of trade sector	Ratio of gross value added of the trade sector and the overall gross value added in a county
Income_per_Citizen	Average discretionary income per citizen per county (in €)
N. of students/county	Number of students in higher education in a county
Pop_Density	Population density in the postal region (in inhabitants per square kilometer)
Fuel_Price_Level	Proxy variable for regional difference in price levels: Difference from average fuel price (product "Super") in the postal region (in € / 1,000)
Cloud_Amount	Cloud amount (in 1/8 of sky cloud coverage)
Rel_Humidity	Relative humidity (in %)
Air_Temp	Air temperature (in degree Celsius)
Wind	Maximum wind speed (in meters per second)
Peak_Wind	Peak wind velocity (max. wind in meters per second)
Rain	Precipitation height (rain in millimeters)
Sunshine	Sunshine duration (in hours)
Nearest_Store	Distance to the nearest store in a radius of 5 km of the search location (in meters)

Please note: In the subsequent models, the dummies for the day of the week are combined in the vector DAY_WEEK. The result dummies go into vector AVAILABLE_OPTIONS and the product categories (e.g., Groceries) are represented by the vector CATEGORY.

---

[16] Economic data and demographic data were obtained from www.destatis.de, weather data from the national weather office. Fuel price information was provided by mehr-tanken.de which is a price comparison website for gas stations.

TABLE 2

Descriptive statistics

Variables	n	Mean	S.D.	Min	Median	Max
Log_Session_Clicks	11,548,637	1.17	1.01	0.00	1.10	6.90
Log_Session_Dist	11,548,637	1.25	2.46	0.00	0.00	13.53
Experience	11,548,637	10.73	46.88	0.00	1.00	48,279.00
Closed	69,145,465	.18	.39	.00	.00	1.00
Near_Store	52,729,087	1,414.42	1,207.55	0.00	1,053.00	5,000.00
Near_store*Groc	52,729,087	759.14	1,111.71	0.00	201.00	5,000.00
Groceries	69,145,465	.52	.50	.00	1.00	1.00
Grocery_Fraction	11,548,637	0.43	0.49	0.00	0.00	1.00
R_Price	69,145,465	.72	.45	.00	1.00	1.00
R_Price_Fraction	11,548,637	.63	.48	.00	1.00	1.00
R_UGC	69,145,465	.72	.45	.00	1.00	1.00
R_UGC_Graction	11,548,637	.63	.48	.00	1.00	1.00
R_Guide	69,145,465	.62	.49	.00	1.00	1.00
R_Report	69,145,465	.19	.39	.00	.00	1.00
R_Other	69,145,465	.01	.09	.00	.00	1.00
Pop_Density	69,115,507	1,213.64	1,190.16	37.00	766.00	4,468.00
Fuel_Price_Level	69,048,308	-1.34	6.54	-14.43	-1.24	28.57
Income_per_Citizen	69,115,507	19,137.56	2,636.79	13,895.00	18,975.00	31,020.00
Air_Temp	67,302,429	14.18	8.33	-24.60	14.20	35.00
Monday	69,145,465	0.13	0.34	0.00	0.00	1.00
Tuesday	69,145,465	.13	.33	.00	.00	1.00
Wednesday	69,145,465	.13	.33	.00	.00	1.00
Thursday	69,145,465	.13	.34	.00	.00	1.00
Friday	69,145,465	.15	.36	.00	.00	1.00
Saturday	69,145,465	.18	.38	.00	.00	1.00
Sunday	69,145,465	.15	.36	.00	.00	1.00
Month_Jan	69,145,465	.10	.30	.00	.00	1.00
Month_Feb	69,145,465	.07	.26	.00	.00	1.00
Month_Mar	69,145,465	.07	.25	.00	.00	1.00
Month_Apr	69,145,465	.10	.30	.00	.00	1.00
Month_May	69,145,465	.16	.36	.00	.00	1.00
Month_Jun	69,145,465	.20	.40	.00	.00	1.00
Month_Jul	69,145,465	.04	.18	.00	.00	1.00
Month_Aug	69,145,465	.04	.19	.00	.00	1.00
Month_Sep	69,145,465	.04	.20	.00	.00	1.00
Month_Oct	69,145,465	.05	.22	.00	.00	1.00
Month_Nov	69,145,465	.06	.23	.00	.00	1.00
Month_Dec	69,145,465	.08	.27	.00	.00	1.00

TABLE 3

Search volume by county

Dep. Variables:	Log of search queries		Log of number of clicks	
	Par. Est. Sig.	*SE*	*Par. Est. Sig.*	*SE*
Log (population in county)	.47672 ***	(.05088)	.49373 ***	(.05456)
Number of students in the county	.00002 ***	$(2.88E^{-06})$	.00001 ***	$(3.09E^{-06})$
importance of trade sector	$6.11E^{-08}$ **	$(2.57E^{-08})$	$7.83E^{-08}$ ***	$(2.76E^{-08})$
Income_per_Citizen	.00008 ***	(.00001)	.00008 ***	(.00001)
Constant	3.33656 ***	(.62868)	4.12884 ***	(.67418)
Observations	378		378	
F-Test	141.85		135.42	
Prob > F	.0000		.0000	
R^2	.6034		.5922	
R^2 (adjusted)	.5991		.5878	

Model: OLS regression

Significance levels: *** $p < .01$; ** $p < .05$; * $p < .1$

TABLE 4

Estimation results on search intensity

Variables	Search intensity	
	Par. Est. Sig.	*SE*
Log Session_Dist	.16162 ***	(.00011)
Experience	-.00049 ***	(.00001)
Closed	.06744 ***	(.00169)
R_Price_Fraction	.11663 ***	(.00060)
R_UGC_Fraction	.11172 **	(.00068)
Groceries_Fraction	.17393 ***	(.00065)
Pop_Density	.00002 ***	(4.90E^{-07})
Fuel_Price_Level	-.00169 ***	(.00009)
Income_per_Citizen	4.83E^{-06} ***	(2.42E^{-07})
Air_Temp	.00457 ***	(.00006)
Tuesday	-.03057 ***	(.00103)
Wednesday	-.00590 ***	(.00104)
Thursday	-.00500 ***	(.00104)
Friday	-.00508 ***	(.00100)
Saturday	-.00424 ***	(.00097)
Sunday	-.02897 ***	(.00193)
Intercept	.50199 ***	(.00479)
Observations	11,215,493	
F-Test	122,752.19	
Prob > F	.0000	
R^2 (overall)	.3503	

Note: monthly dummies included in estimation

Dependent variable: Log (Session_Clicks)

Model: Fixed-effects regression

Significance levels: *** $p < .01$; ** $p < .05$; * $p < .1$

TABLE 5

Estimation results on information choice behavior[17]

Variables	Product_Info			Price_Info			UGC_Info		
	Par.Est.	*Sig.*	*SE*	*Par.Est.*	*Sig.*	*SE*	*Par.Est.*	*Sig.*	*SE*
Log Session_Dist	.04752	***	(.00035)	.06660	***	(.00022)	.06018	***	(.00027)
Experience	-.00049	***	(.00003)	-.00135	***	(.00002)	-.00043	***	(.00002)
Closed	.03889	***	(.00622)	-.02290	***	(.00396)	.09992	***	(.00459)
Near_Store	.00003	***	$(1.11E^{-06})$	.00001	***	$(7.26E^{-07})$	.00002	***	$(9.84E^{-07})$
Near_Store *Groc	$-3.17E^{-06}$	***	$(1.59E^{-06})$	.00001	***	$(1.00E^{-06})$	-.00001	***	$(1.25E^{-06})$
Groceries	-.12864	***	(.00329)	-.36010	***	(.00200)	-.27646	***	(.00240)
R_Price	-.36086	***	(.00203)				-.44027	***	(.00165)
R_UGC	-.06043	***	(.00232)	-.19310	***	(.00152)			
R_Guide	-.04793	***	(.00238)	-.01206	***	(.00144)	-.08860	***	(.00181)
R_Report	-.23175	***	(.00269)	-.13287	***	(.00148)	-.18427	***	(.00186)
R_Other	-.09872	***	(.01203)	.03254	***	(.00672)	-.03142	***	(.00792)
Pop_Density	-.00005	***	$(1.87E^{-06})$	-.00004	***	(.00000)	-.00005	***	$(1.53E^{-06})$
Fuel_Price_Level	.00178	***	(.00039)	.00201	***	(.00024)	.00128	***	(.00031)
Income_per_Citiz.	-.00001	***	$(9.54E^{-07})$	-.00001	***	$(5.92E^{-07})$	-.00001	***	$(7.69E^{-07})$
Air_Temp	-.00574	***	(.00024)	-.00849	***	(.00015)	-.00723	***	(.00018)
Tuesday	.01040	***	(.00387)	.03505	***	(.00239)	.03640	***	(.00296)
Wednesday	.00892	***	(.00391)	.05001	***	(.00241)	.03269	***	(.00300)
Thursday	-.00798	***	(.00391)	.02108	***	(.00241)	.02173	***	(.00299)
Friday	-.02742	***	(.00379)	.00674	***	(.00234)	-.00961	***	(.00292)
Saturday	-.00679	***	(.00365)	.02311	***	(.00227)	-.01882	***	(.00283)
Sunday	.02450	***	(.00709)	.04330	***	(.00450)	-.02782	***	(.00527)
Observations	31.7 million			34.4 million			29.2 million		
Log Likelihood	-5,423,194.1			-12,144,349			-8,27,3978.2		
LR-Test	147,006.12			696,281.05			394,428.51		
Prob > Chi²	.0000			.0000			.0000		

Note: monthly dummies included in estimation

Dependent variables: Click on product information, price information or UGC

Models: Discrete choice models with fixed effects

Significance levels: *** $p < .01$; ** $p < .05$; * $p < .1$

---

[17] Please note that each model is based on a different sample – the sampling criterion is on the dependent variable i.e., we include the observation only if the respective type of information is available.

TABLE 6

Marginal effect estimates[18]

| Variables	Product_Info			Price_Info			UGC_Info		
	dy/dx	Sig.	SE	dy/dx	Sig.	SE	dy/dx	Sig.	SE
Log Session_Dist	.00922	***	(.00011)	.01236	***	(.00009)	.01090	***	(.00010)
Experience	-.00009	***	(.00000)	-.00025	***	(.00000)	-.00008	***	(.00000)
Closed	.00759	***	(.00122)	-.00423	***	(.00073)	.01839	***	(.00087)
Near_Store	.00001	***	(.00000)	$1.79E^{-06}$	***	(.00000)	$3.31E^{-06}$	***	(.00000)
Near_Store *Groc	$-6.16E^{-07}$	**	(.00000)	$1.15E^{-06}$	***	(.00000)	$-9.58E^{-07}$	***	(.00000)
Groceries	-.02501	***	(.00067)	-.06773	***	(.00054)	-.05136	***	(.00058)
R_Price	-.07274	***	(.00072)				-.08526	***	(.00068)
R_UGC	-.01180	***	(.00046)	-.03687	***	(.00035)			
R_Guide	-.00933	***	(.00047)	-.00224	***	(.00027)	-.01623	***	(.00035)
R_Report	-.04334	***	(.00064)	-.02416	***	(.00030)	-.03241	***	(.00042)
R_Other	-.01870	***	(.00223)	.00609	***	(.00127)	-.00564	***	(.00141)
Pop_Density	-.00001	***	(.00000)	-.00001	***	(.00000)	-.00001	***	(.00000)
Fuel_Price_Level	.00035	***	(.00008)	.00037	***	(.00004)	.00023	***	(.00006)
Income_per_Citiz.	$-2.24E^{-06}$	***	(.00000)	$-2.33E^{-06}$	***	(.00000)	$-1.47E^{-06}$	***	(.00000)
Air_Temp	-.00111	***	(.00005)	-.00157	***	(.00003)	-.00131	***	(.00003)
Tuesday	.00202	***	(.00076)	.00654	***	(.00045)	.00664	***	(.00055)
Wednesday	.00173	**	(.00076)	.00936	***	(.00046)	.00596	***	(.00056)
Thursday	-.00155	**	(.00076)	.00393	***	(.00045)	.00395	***	(.00055)
Friday	-.00529	***	(.00072)	.00125	***	(.00044)	-.00174	***	(.00053)
Saturday	-.00132		(.00070)	.00430	***	(.00043)	-.00340	***	(.00051)
Sunday	.00477	***	(.00139)	.00809	***	(.00085)	-.00501	***	(.00094)
Observations	31.7 million			34.4 million			29.2 million		

Note: monthly dummies included in estimation

Dependent variables: Click on product information, price information or UGC

Models: Discrete choice models with fixed effects

Significance levels: *** p < .01; ** p < .05; * p < .1

---

[18] Please note that each post-estimate model is based on a different sample – the sampling criterion is on the dependent variable i.e., we include the observation only if the respective type of information is available.

FIGURE 1

Illustration of the data generation process

1. search (e.g., scan)
2. product information
3. price comparison
4. reviews (UGC)
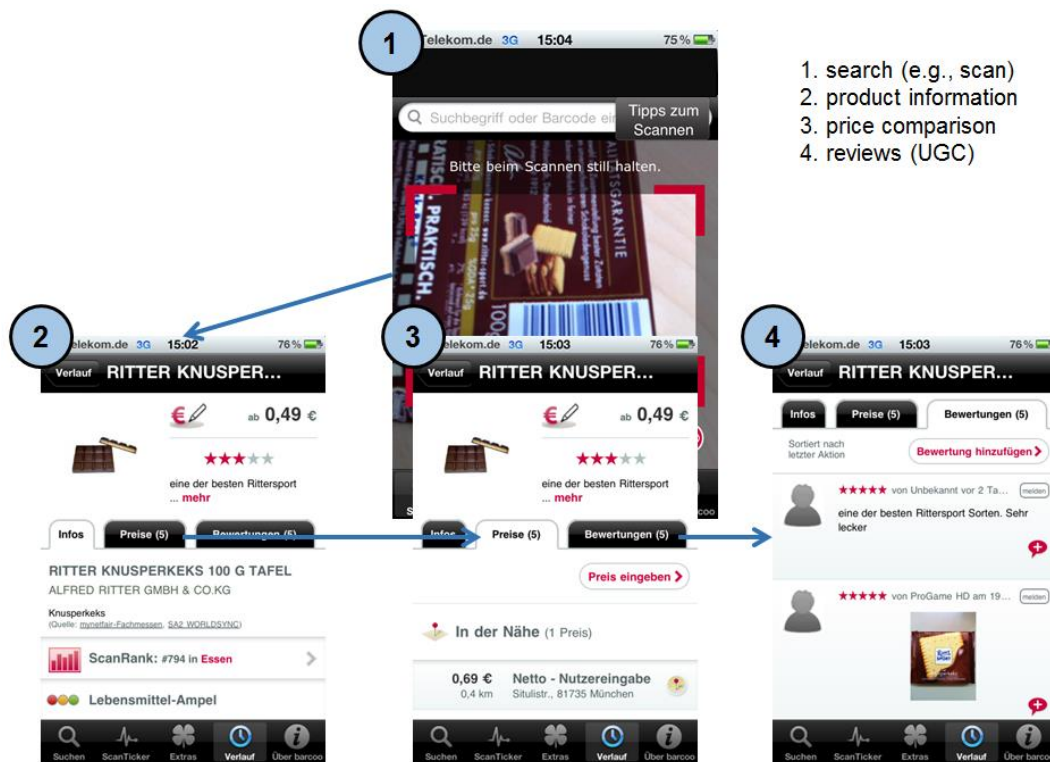
FIGURE 2

Explorative analysis of the search path of a typical consumer
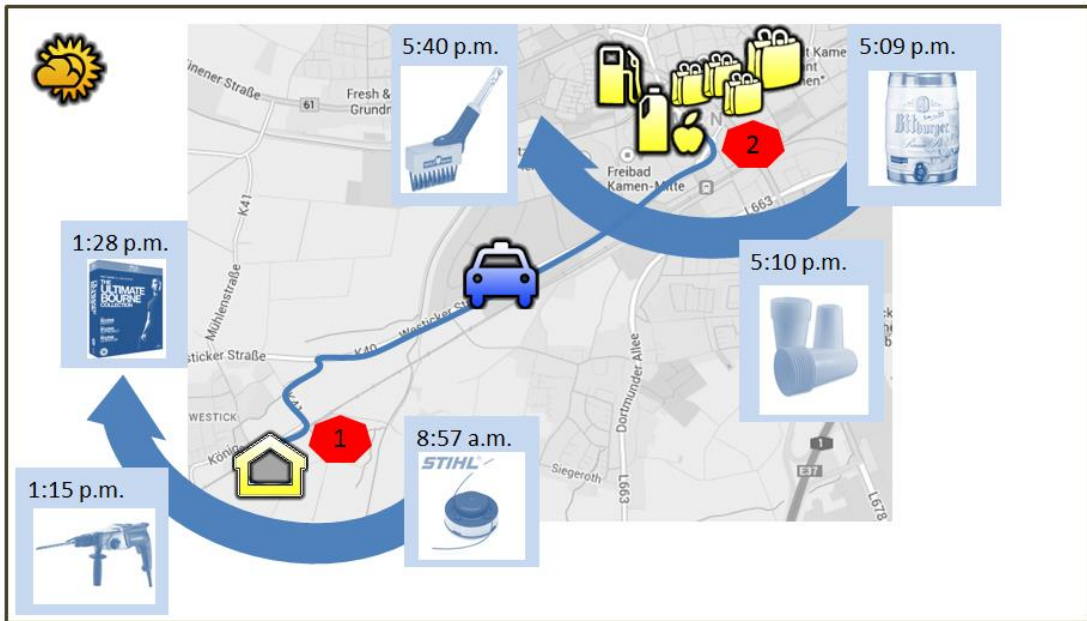
FIGURE 3

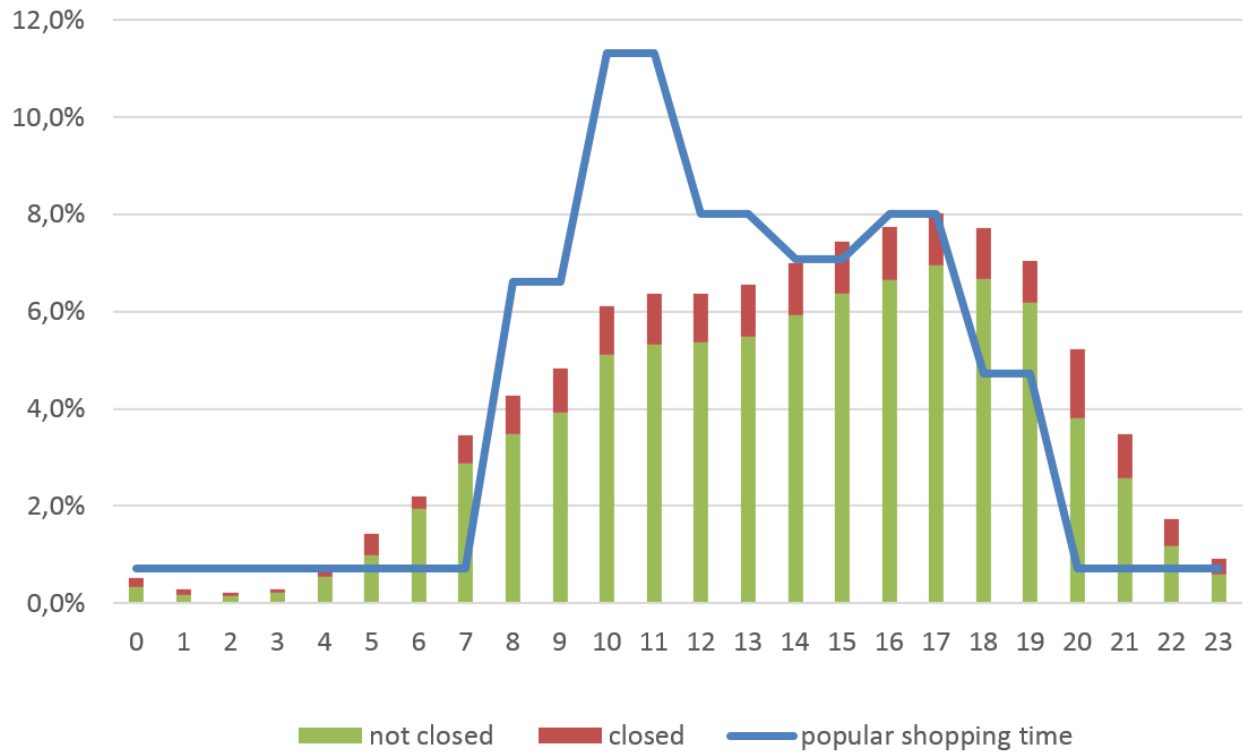**Temporal usage patterns of users by day of the week**

Percentages of app usage over weekdays and popular shopping days in Germany (VuMa 2015); Note: closed = closing time by law of the respective state.
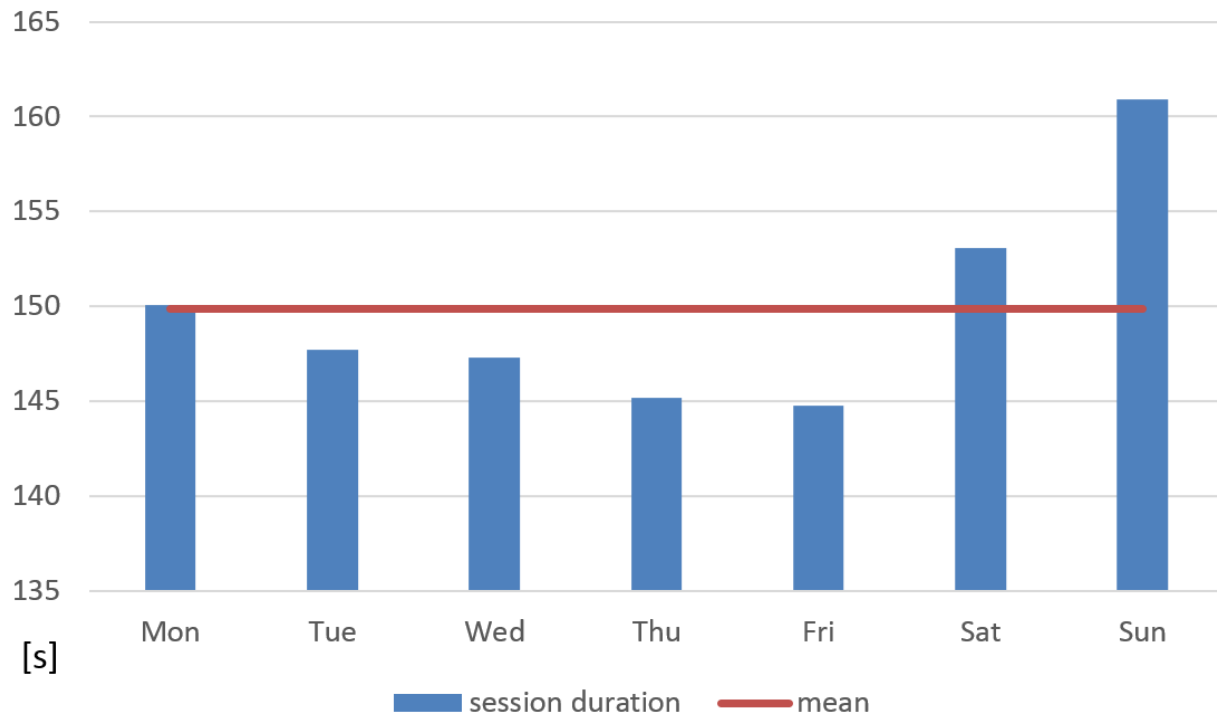
FIGURE 4

Temporal usage patterns of users by time of the day

App usage by time of day and popular shopping times in Germany (GFK 2013). Note: closed = closing time by law of the respective state. Closed portion during usual business hours represents activity on Sundays.

FIGURE 5

Search session duration by day of the week

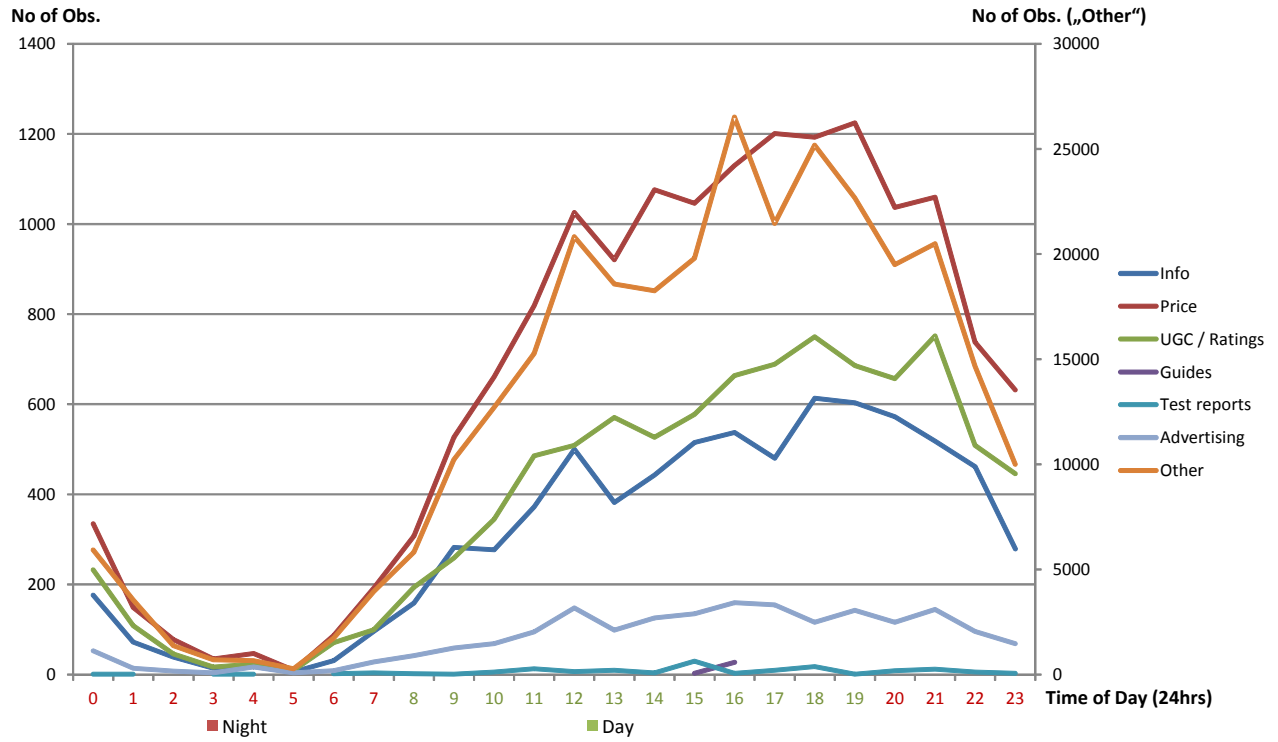Mean session durations in seconds; only sessions: 0 < t < 3600 s.

FIGURE 6

Distance traveled per search session by day of the week
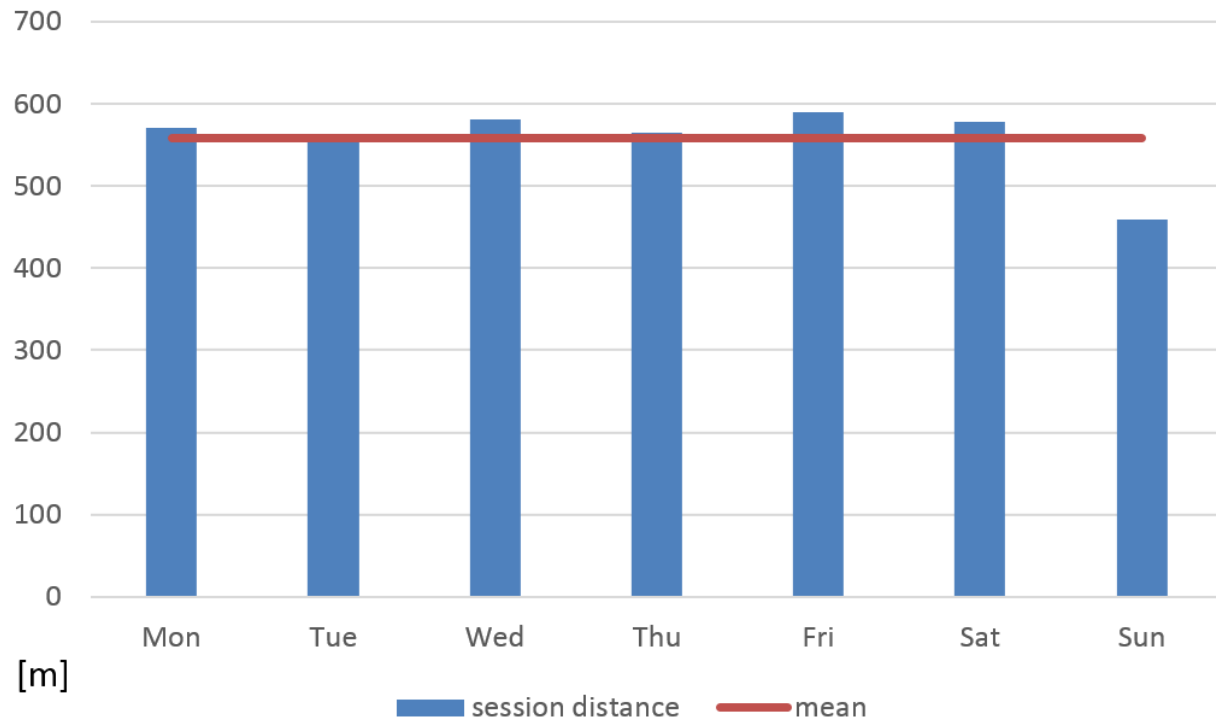
Mean travel distances during a session in meters; only sessions: 0 < d < 5000 m.
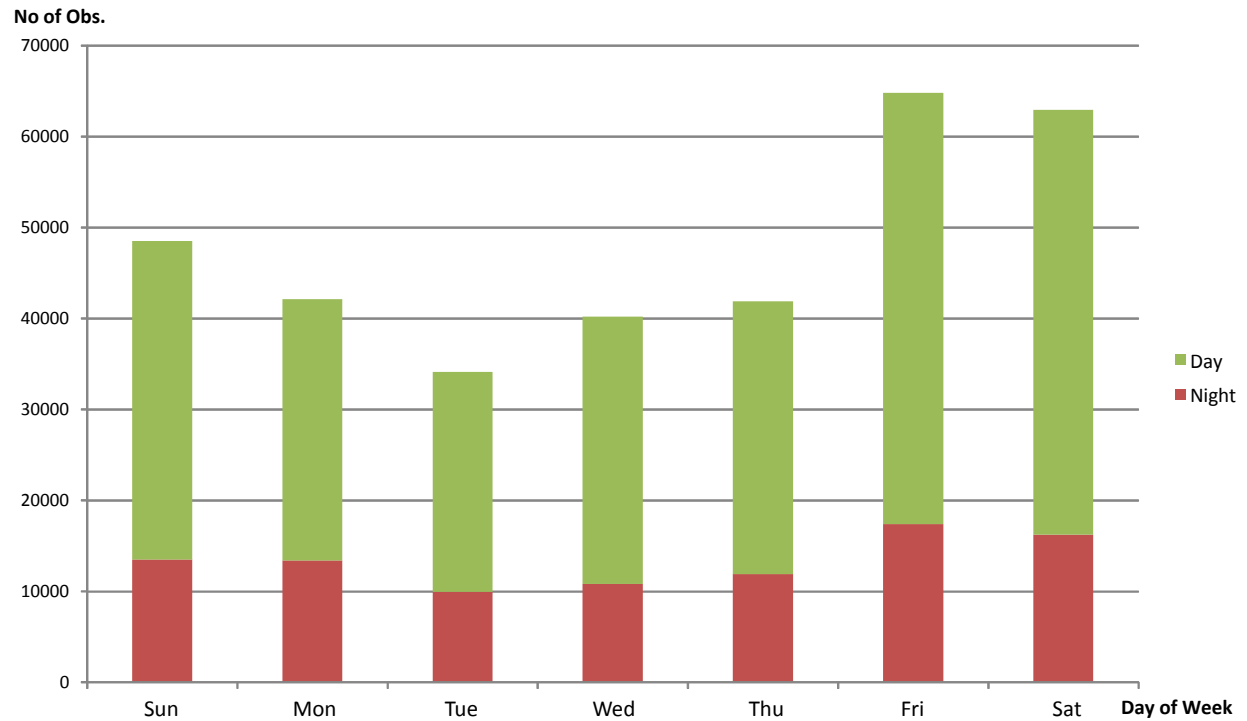
FIGURE 7

Clicks by information type category

Subsample of greater Munich area in June 2012. Right axis: no. of obs. for click type category "other". Left axis: all other click type categories.
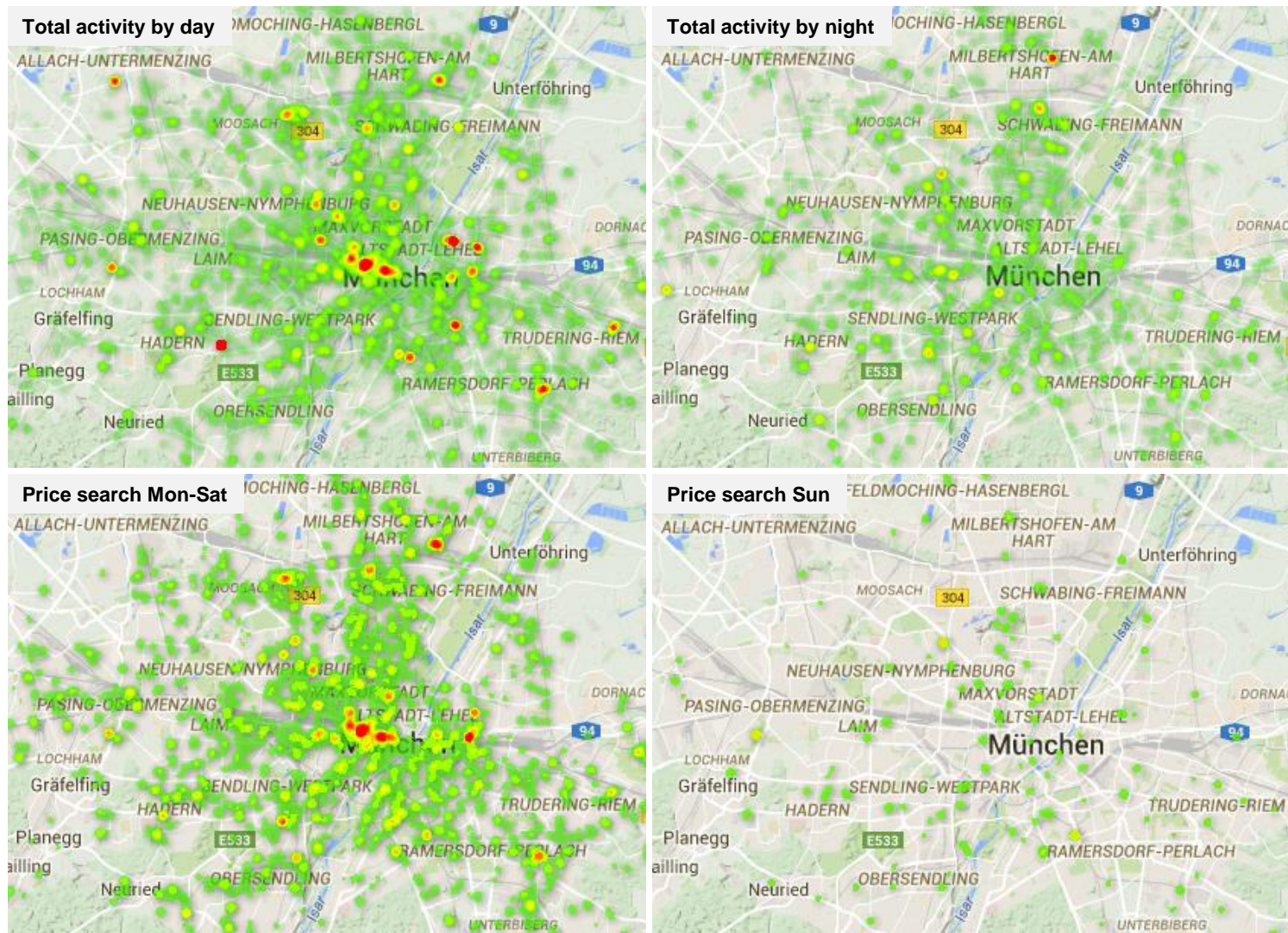
FIGURE 8

Clicks by day of week and day vs. night

Subsample of greater Munich area in June 2012. Store opening hours Monday through Saturday from 8a.m. to 8p.m. Sundays closed.

FIGURE 9

Heatmaps of user activity (clicks)

Subsample of greater Munich area in June 2012. Overall clicks (upper panels). Clicks for price information only (lower panels).