Spring 5-29-2015

# The Influence of the Herfindahl-Hirschman Index and Product Complexity on Search Behaviour: A Cross-sector Study of the U.S., Germany and U.K.

Christopher P. Holland
*University of Manchester*, chris.holland@mbs.ac.uk

Julia A. Jacobs
*University of Münster*, julia.jacobs@uni-muenster.de


ISBN 978-3-00-050284-2

# THE INFLUENCE OF THE HERFINDAHL-HIRSCHMAN INDEX AND PRODUCT COMPLEXITY ON SEARCH BEHAVIOUR: A CROSS-SECTOR STUDY OF THE U.S., GERMANY AND U.K.

*Complete Research*

Holland, Christopher P., University of Manchester, Manchester, United Kingdom, chris.holland@mbs.ac.uk

Jacobs, Julia A., University of Münster, Münster, Germany, julia.jacobs@uni-muenster.de

## Abstract

*Consumer search is analysed in a cross-sector study of six markets in the US, Germany and UK using online panel data. Two constructs are used to measure the search process: the consideration set and use of price comparison engines. The consideration sets range from 2.3 to 3.1 in the US, from 2.3 to 2.6 in Germany and from 2.6 to 3.2 in the UK, regardless of the use of price comparison engines. These results are significantly smaller than expected compared to pre-Internet studies and theory predictions. However, they are consistent with the few published results that used online panel data. It is shown that the consideration set is a function of the Herfindahl-Hirschman Index. The use of price comparison engines is inversely related to product complexity. The theoretical and managerial implications of the research results are explained and the potential of using online panel data for future research into online consumer behaviour and strategy is outlined.*

*Keywords: consumer search behaviour, consideration set, online panel data, market structure, international and multi-sector research, Herfindahl-Hirschman Index, clickstream analysis*

## 1 Introduction

Consumer search is important to marketing strategy and the competitive process because it influences how consumers choose between competing suppliers. Search forms an important stage in the customer journey and has been widely researched in different marketing contexts before the widespread use of the Internet (e.g. Beatty and Smith, 1987; Bloch et al., 1986, 1989; Moorthy et al., 1997; Engel et al., 1995). The focus of this paper is to evaluate the use of the Internet for the search in six markets in the US, UK and Germany: insurance; banking; airline; telecommunications, automotive and grocery. The research explores how the product-market factors of industry concentration and product complexity are related to the consumer search process.

In online marketing the Internet has three distinct purposes: search, sales and service. Online search is primarily concerned with building awareness of alternatives and the evaluation and choice between competing products and suppliers, and has important implications for understanding the nature of the competitive process in terms of new customer acquisition, customer retention and switching behaviour. The consumer search process is therefore an important component of how markets function. Economists have recognised this and examined search strategies by modelling the cost of acquiring information and the expected benefits of additional search. Stigler (1961) demonstrated that consumers continue to search until the expected benefits are outweighed by the cost of collecting and evaluating the information. In a similar vein, Bakos (1997, 1998) and Malone et al. (1989) showed that the Internet reduces search costs dramatically and therefore leads to more extensive search patterns and the formation of electronic markets.

In this paper two measures of search are used: the consideration set and the use of price comparison engines. The consideration set is an important measurement of search and competitive intensity

because it captures the range of possible suppliers that are actively considered by consumers (Hauser and Wernerfelt, 1990). It also provides a direct method of evaluating the search process across different markets and between different countries. Price comparison engines are now an important search strategy for consumers and an important distribution channel for companies, especially in the airline and insurance markets. In competitive strategy and economy theory, industry concentration is an important factor that determines the competitive intensity of specific markets, and should therefore influence the consumer search process (Stigler 1961, Rhoades 1993). Economic theory tends to focus on the influence of industry concentration on the behaviour of suppliers but it is also reasonable to examine the relationship from the consumer perspective. Product complexity is also an important marketing variable that affects the search and buying process (Kotteaku et al. 1995, Swaminathan 2003) and its influence on the use of price comparison engines is investigated. The research analyses the nature and extent of the online consumer search process in a cross-sector study of markets in the United States, United Kingdom and Germany, using online panel data. An analysis of the literature on information search, the consideration set concept and industry concentration is given in the next section.

# 2 Literature Review

## 2.1 Consumer Search Process and the Internet

Information exchange between customers and suppliers plays a crucial role in the functioning of markets (Stigler, 1961). Customers search out information on products offered by different competitors in areas such as price, quality and service. Competitors provide information through advertising and other promotional activities. A fundamental question for economists is how much search is optimal for consumers? Stigler (1961) modelled information directly and showed that more search was required in markets with higher price variance to achieve an optimal outcome of benefits and search costs. The more general point is that the higher the level of differentiation between competitors, the higher the level of consumer search is required in order to gain a full understanding of the market.

Early research into the Internet and search by Malone et al. (1989), Bakos (1997, 1998) and Alba et al. (1997), showed that the Internet reduces search costs, provides access to a much greater breadth and depth of information, facilitates product comparison and is convenient. More recently, additional benefits have been identified such as the ability to tailor the information to specific requirements (Su, 2008; Jepsen, 2007; Hoffman and Novak, 1996, 1997). A research report by McKinsey (Bughin et al., 2011) identified further search benefits: better matching of individuals and companies, time savings, raised awareness of new products and services and other salient market information such as quality issues, increased price transparency, ability to find long-tail offerings that would otherwise be uneconomic to provide, new digital business models and entertainment. In summary, the general theoretical impact of the Internet is therefore to increase the scale and sophistication of the consumer search process, dependent on the expected benefits of additional search. An important determinant of the benefits from additional search in a specific market is the nature of competition, which is a function of industry concentration (Arndt 1977, Porter 1985, Besanko et al. 2009).

## 2.2 Industry Concentration: The Herfindahl-Hirschman Index (HHI)

The Herfindahl-Hirschman Index (HHI) is the standard measure for concentration and has been used in competition regulation (Calkins 1983, Rhoades 1993) and to provide a measure of the likely level of competitive intensity in specific markets (Besanko et al., 2009). The HHI is calculated by squaring the market shares of the major competitors and summing them (Besanko et al. 2009, Rhoades 1993, Calkins 1983). HHI is a variable that makes it possible to make an objective comparison of industry concentration between different markets. The reciprocal of HHI is termed the numbers-equivalent of competitors ‘N’, where 1/HHI = N. In order to make the relationship between industry concentration and consideration set more intuitive, the *numbers-equivalent of competitors* is used. The theoretical

logic is that a higher level of numbers-equivalent of competitors results in more variety of product offers and higher price variance, and therefore leads to a more extensive consumer search process and larger consideration set. To measure the search process, two concepts are used: the consideration set and the extent of the use of price comparison engines.

## 2.3 Consideration Set Concept

Shocker et al. (1991) reviewed the literature on consideration sets and consumer decision-making and demonstrated the validity and importance of the consideration set concept. They concluded that the consideration set is an important construct for understanding and evaluating consumer behaviour and identified a further stage immediately before the purchase is made and termed this the 'choice set', although in practice, it is very difficult to clearly distinguish between the two, especially given recent evidence of the rather iterative process that consumers follow in their evaluation of the consideration set (Court et al., 2009). Brown and Wildt (1992) described the decision making process of consumers in terms of firstly having an awareness set of possible alternatives, which is then narrowed down to a smaller number of competitors that are actively considered, which is termed the consideration set. The definition used in this paper is similar to the original terminology of Howard and Sheth (1969) of 'evoked set'. As our paper focuses on the online world, a more recent definition of Holland and Mandry (2013, p. 2919) is used, who defined the consideration set as:

***"… the group of suppliers that a buyer actively considers in their decision-making before purchasing a product or service".***

## 2.4 Size and Composition of the Consideration Set

Pre-Internet, Hauser and Wernerfelt (1990) examined decision criteria for brand consideration based on economic utility and cost. In a review of the published literature to this date on the size of consideration sets, they found that the mean size of consideration sets across a variety of products was 4.05, ranging from 2.0 up to 8.1, and included categories as diverse as tea, toothpaste automobiles, coffee, beer, shampoo and medicine. A similar result of 3.98 was reported from the assessor database, a pre-test market forecasting system (see Silk and Urban, 1978). Hauser and Wernerfelt (1990) also examined the distribution of the consideration set and found it to be a lognormal distribution. That is, the majority of consumers consider two to four brands and relatively few consumers, typically less than 20%, have larger consideration sets. A similar theoretical approach was adopted by Roberts and Lattin (1991) who argued that the likelihood of considering an additional brand is related to the expected benefit and associated cost of search. All of these marketing authors, together with research from an Information Systems perspective, e.g. Malone et al. (1989), adopt essentially the same logic that was outlined by Stigler (1961). Consumers will continue to search as long as the expected marginal benefits of further search outweigh the marginal costs of search.

## 2.5 Online Search and Consideration Sets

Survey data from the EU (Loof and Seybert, 2009), the US (Jones and Fox, 2009) and China (CINIC, 2013) all demonstrate the use of the Internet for the specific purpose of searching for products and services. They also show that online search is more prevalent than online sales, which is important in multi-channel markets such as grocery, mobile phones and banking, where an important marketing function of the Internet is to support and encourage consumer search in order to encourage sales in the shop channel. Survey data provides valuable information about the scale of search but tells us nothing about the extent of the search process or the use of price comparison engines.

Peterson and Merino (2003) adopted a theoretical approach to modelling online search and put forward a series of propositions to evaluate the use of intelligent agents such as price comparison engines. Punj and Moore (2009) conducted experiments with student groups, and found that the size of the consideration set increases in a web environment, and that more alternatives lead to larger consideration sets. Whilst theoretical arguments and experiments contribute to the conceptual analysis of consumer behaviour, they are not a substitute for empirical analysis. Online panels, emerged within

the mid- 1990s (Flynn 1995). The International Organization for Standardization (2012) defined an online panel as "a sample database of potential respondents who declare that they will cooperate for future data collection if selected". Bucklin et al. (2002) and Bucklin and Sismeiro (2003, 2009) argue persuasively about the theoretical and empirical research value of online panels, in particular the potential of using clickstream data for marketing research. The use of online panel data in marketing research is in its infancy but there are a small number of important research papers related to consumer search behaviour and these are shown in Table 1.

| Authors | Study | Key Findings |
|---|---|---|
| Lohse et al. (2000) | Consumer Buying Behavior on the Internet: Findings from Panel Data | - First study to use panel data to investigate how consumers look for information<br>- Panel data is an efficient method to observe changes over time and to make forecasts for the future.<br>- Internet population looks more and more like the general US population<br>- Dynamic Internet shopping growth |
| Johnson et al. (2004) | On the Depth and Dynamics of Online Search Behaviour | - Limited online search<br>- Analysis at the level of individuals<br>- Consideration sets in the US are 1.2 for books, 1.3 for CDs and 1.8 for travel sites |
| Zhang et al. (2006) | Online Consumer Search Depth: Theories and New Findings | - Repeat study of Johnson et al. (2004) with the outcome that the consumers, at the level of households, did search more in the study of 2006 than 2004<br>- Consideration sets in the US are 2.1 for CDs, 3.3 for airline tickets and computer hardware<br>- Price of high-value products is an important variable for consumer search<br>- Internet reduces search costs<br>- Price and quality influence purchase decisions online |
| Meyer and Stobbe (2010) | Majority of bank customers in Germany do research online: Findings of a clickstream analysis | - 60% of users research financial topics online<br>- Financial research takes on average 7 ½ weeks<br>- One third of the research process include Google for financial products compared to about 13% conducted with price comparison webpages<br>- Brand names have a big impact on the extent of online research for customers<br>- Consideration set size is on average 3.8.<br>- Online presence and market share determine reach |
| Holland and Mandry (2013) | Online Search and Buying Behaviour in Consumer Markets | - Consideration sets across six market sectors in the UK, and the same sectors in the US are relatively small, regardless of the use of price comparison engines<br>- Average consideration sets are in a very narrow range between 2.40 and 2.77 (UK) and 2.13 – 2.60 (US)<br>- The consideration set is negatively correlated with consumer perceived risk |

*Table 1. Results of online panel data research literature related to consumer behaviour*

Although there are very few studies, an important common result is the relatively small size of consideration sets ranging from 1.2 for books reported by Johnson et al. (2004) to 3.8 in the German bank market (Meyer and Stobbe, 2010). Zhang et al. (2006) repeated the earlier study of Johnson et al. (2004) and showed an increase in consideration set size from the earlier study, with 2.1 for CDs and 3.3 for airline tickets and computer hardware. This may have been due to an increase in sophistication of online users and possibly changes in the significantly increased size of the online panel data set. The study by Holland and Mandry (2013) is the only study to look at a broad range of market sectors internationally. The study made an important distinction between online search activity and e-service, and the same approach is used in this paper. E-service is defined as consumers who look at one

website only, and online researchers are those who visit two or more websites. The differences in the average size of the consideration sets reported by Johnson et al. (2004), Zhang et al. (2006) and Holland and Mandry (2013) are therefore partly explained by a methodological difference.

# 3 Research Framework and Hypotheses

The research framework and hypotheses are shown in Figure 1.

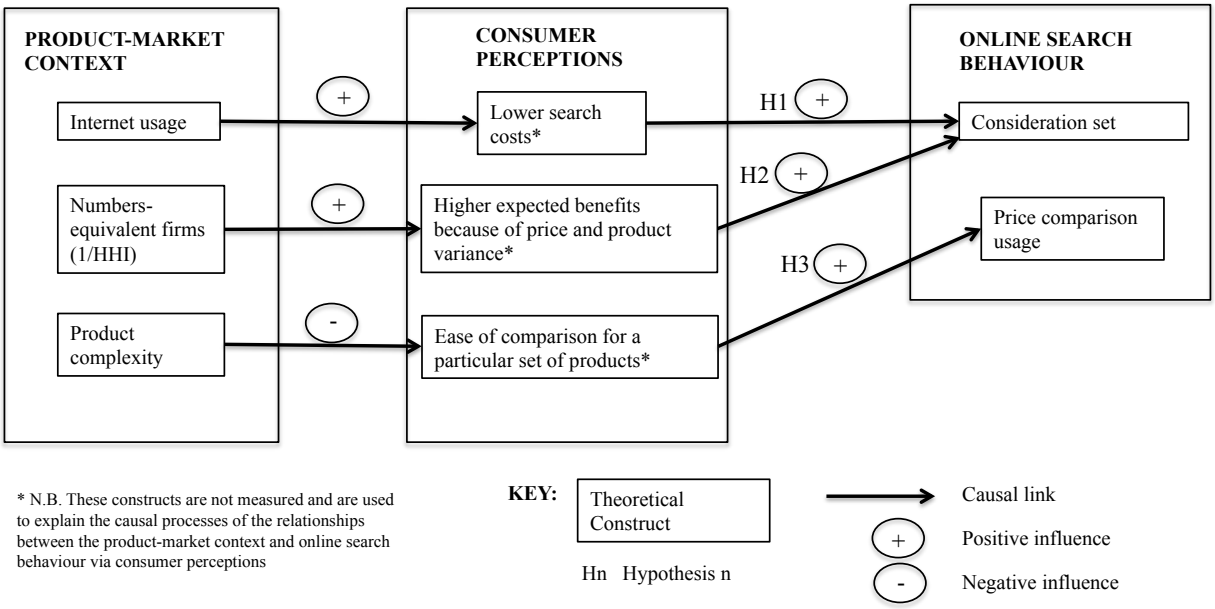


*Figure 1. Research Framework*

The logic of the model is that product-market context influences consumer perceptions that result in observable search behaviour. The nature of the relationships and the hypotheses are described below.

> *Hypothesis 1: The online search process measured by the consideration set construct will be more extensive than pre-Internet search.*

The Internet reduces search costs significantly and may be more convenient to use than other channels for many consumers (Malone et al., 1989; Bakos, 1997; The Economist, 2000; Zhang et al., 2006; Bughin et al., 2011). One would therefore expect more extensive search patterns than pre-Internet because the expected benefits from additional search are more likely to be greater than the relatively small search costs, based on the economic theory of Stigler (1961).

> *Hypothesis 2: The consideration set is a function of the numbers-equivalent of competitors, which is defined as 1/HHI.*

From economic and marketing theory, market structure influences the consideration set because it affects the available choice set, it determines the nature of advertising and communication (Cooper and Inoue, 1996; Urban et al., 1984) and determines to an extent the nature of competition (Porter 1985, Besanko et al. 2007). In a highly concentrated market, consumers would have a good overview of potential suppliers, and the range of choices would be smaller. In a less concentrated market, one would expect consumers to have less knowledge about all of the competitors, and would have a wider set of choices. One would therefore expect more search in a less concentrated market than in a more concentrated market. The relationship between HHI and the consideration set concept is theoretically important because it relates economic theory of market concentration directly with consumer search behaviour and provides a novel perspective of the competitive process.

> *Hypothesis 3: The use of price comparison engines is related to the ease of comparison between product choices for consumers.*

Malone et al. (1989) identified the theoretical limitation of the complexity of product descriptions on the development of electronic markets, but the relationship between price comparison engines and product complexity has not been tested on a large scale.

# 4 Research Methodology

Industry concentration is measured by the Herfindahl-Hirschman Index (HHI), which is used to derive the numbers-equivalent of competitors. Product complexity is measured using an ordinal scale. The theoretical construct consideration set is based on the marketing definitions from Hauser and Wernerfelt (1990) and Holland and Mandry (2013). The use of price comparison engines is defined as the level of price comparison usage measured as a percentage of total online activity within each market in terms of the number of online users. An explanation of the data sources is given below.

## 4.1 Industry Concentration

The US, UK and German online markets were chosen because of their scale and sophistication. The US has an Internet population of over 200 million and Germany is the largest and most developed online market in Europe with approximately 60 million online users, closely followed by the advanced UK online market. Six economically important consumer market sectors in each of the countries are insurance, banking, telecommunications, automotive, grocery and airline. The composition of each market includes the top competitors defined by market share representing about 90% of the major competitors. The Herfindahl-Hirschman Index was calculated from market share data taken from authoritative sources comprised of company reports, Government organisations, market research companies and industry standard reports. See Table 2 for details of the German and US market. UK data were similarly collected from company reports, industry bodies and trade organisations.

| **Industry** | **Market Share Data Sources** | |
|---|---|---|
| | **Germany** | **US** |
| Insurance | Direct from company reports. Allianz, 2013; AXA, 2013; Huk-Coburg, 2013; Zurich, 2013; Cosmodirekt, 2013; Ergo, 2013; R+V Versicherung, 2013; Signal Iduna, 2013; Versicherungskammer Bayern, 2013; Debeka, 2013. | US Federal Insurance Office, 2013. |
| Banking | The Banker, 2012, Top 1,000 banks annual report. | Forbes, 2012. |
| Tele-communications | Wireless Intelligence, 2012. | Fierce Wireless, 2013. |
| Automotive | Verband der Automobilindustrie, 2013. | The Wall Street Journal, 2014. |
| Grocery | Lebensmittelzeitung, 2013. | Progressive Grocer, 2013. |
| Airline | Lufthansa, 2013; Airberlin, 2013; Ryanair 2013; Germanwings, 2013; EasyJet, 2013; Turkish Airlines, 2013; Wizz Air, 2013; KLM, 2013; Austrian Airlines, 2013. | Allegiant Air, 2013; Transtats, 2013. |

*Table 2. Sources used for market share data*[1]

## 4.2 Online Panel Data

Online Panels emerged in the mid-1990s (Flynn 1995) and comScore has established itself as an industry leader with approximately two million members worldwide. Online panel data from ComScore is used to provide cross-sector data based on a panel of approximately one million in the US, and of the order of magnitude of 100,000 in Germany, and the UK (ComScore, 2009). Online panel data provides clickstream data that is compiled into a series of standard reports by the company (Bucklin et al., 2002; Bucklin and Sismeiro, 2009). It has several advantages over survey data because

[1] UK, German and US market share data sources are available on request.

it provides extremely detailed, accurate and comprehensive longitudinal data that is based on actual behaviour rather than accounts of historical behaviour or statements of future intent (Göritz et al. 2002). It therefore has features of both intensive and extensive research methodologies as described by Sayer (1993), because the automatic, electronic collection of data makes is possible to track surfing patterns across multiple websites. This cross-competitor behaviour was also a feature of pre-Internet panels (Goodhardt and Ehrenberg, 1967; Goodhardt et al., 1984), and was used to analyse marketing problems such as brand preference, loyalty and buying behaviour. Online panels have been successfully used in research for the past 20 years in a variety of fields, including medical, market and social research (Callegaro et al. 2014). It is now becoming more established in consumer behaviour research (see Table 1) and also within industry, e.g. see Meyer and Stobbe (2010) that was sponsored by Deutsche Bank, GfK and Google. The methodological approach is explained in more detail by Bucklin et al. (2002), Johnson et al. (2004), Zhang et al. (2006), Bucklin and Sismeiro (2009), Napoli et al. (2014) and Holland et al. (2015).

## 4.3 Definition and Measurements of Online Behavioural Constructs

The consideration set is defined as the range of companies that a consumer searches by visiting two or more websites within a market sector[2]. The studies conducted by Zhang et al. (2006) and Johnson et al. (2004) included consumers that visited one website only. The rationale for defining researchers as those looking at two or more websites is that consumers who only visit one website are likely to be conducting some kind of e-service activity, e.g. online banking, topping up a mobile phone account, or placing a grocery order. It is therefore potentially misleading to include these e-service users as 'searchers' in the context of measuring a consideration set. Two standard comScore reports were used to measure the consideration set and the use of price comparison engines, the 'Key Measure' and 'Audience Duplication' reports. The audience duplication report provides data on cross-visiting between competitors. It is therefore possible to calculate the number of searchers within a market, and the number of different competitors visited by the searchers. The consideration set within a particular market is measured by calculating the average of the number of different competitor websites visited by all searchers.

The use of price comparison engines is defined as the proportion of activity that consumers allocate to price comparison engines as a percentage of visitors to competitor websites. The measurement is consistent across sectors and therefore gives a comparable measure. Price comparison engines were identified within each sector using secondary data, the knowledge of the researchers who are very familiar with the US, German and UK markets, comScore listings of price comparison engines, and triangulated with online searches. The size of the price comparison engines is measured by unique visitors and is given in the 'Key Measures' report for price comparison engines and competitors for each market. The researchers are confident that all of the major price comparison engines were captured, and if smaller ones were omitted, then their impact is very small because of the semi-logarithmic distribution of online users to websites.

The ease of comparison score is a simple ordinal scale of high, medium and low and is inversely related to product complexity, which is based on the measurement of Swaminathan (2003). This approach has also been taken by other researchers to define broad based measures of variables such as risk, involvement, price, time and differentiation (Beatty and Smith 1987, Moorthy et al. 1997, Swaminathan 2003, Su 2008, Parra and Ruiz 2009).

[2] By looking at one webpage, as e.g. www.lufthansa.com, we consider this to be representative for a brand. However, it cannot be ensured that within the airline sector for example, other brands might show up as an alternative option especially if the airline is part of an airline network.

# 5 RESULTS

## 5.1 Consideration Sets

Results for the numbers-equivalent of competitors, average consideration sets and use of price comparison engines in Germany (DE), the United States (US) and the United Kingdom (UK) are presented in Table 3.

| Market Sector | Numbers-Equivalent of Competitors | | | Average Consideration Set | | | Price Comparison Usage | | Ease of Comparison |
|---|---|---|---|---|---|---|---|---|---|
| | DE | US | UK | DE | US | UK | DE | US | |
| **Insurance** | 2.95 | 6.68 | 11.27 | 2.31 | 2.45 | 3.16 | 10% | 24% | H |
| **Banking** | 5.08 | 5.49 | 6.08 | 2.34 | 2.65 | 2.75 | 19% | 10% | M |
| **Airlines** | 6.37 | 6.36 | 6.74 | 2.44 | 2.73 | 2.65 | 75% | 51% | H |
| **Tele-communications** | 5.59 | 3.65 | 9.39 | 2.51 | 2.40 | 2.84 | 4% | 6% | L |
| **Automotive** | 6.40 | 7.23 | 14.35 | 2.53 | 3.10 | 3.05 | 9% | 51% | H |
| **Grocery** | 7.35 | 4.58 | 6.95 | 2.60 | 2.28 | 2.63 | 4% | <1% | L |

*Table 3. Results database for 2012*

The first important result to note is that all of the consideration sets in the US, UK and Germany are relatively small and fall within a very narrow band of between 2.28 and 3.10 in the US, between 2.31 and 2.60 in Germany and between 2.63 and 3.16 in the UK. The average consideration sets are 2.60 (US), 2.85(UK) and 2.46 (Germany). This means that most consumers consider just two or three competitors and only a very small percentage conduct an extensive search process of four or more competitors. This is true regardless of the use of price comparison engines. In an authoritative and widely cited review of the literature on reported sizes of consideration sets, Hauser and Wernerfelt (1990) reported two sets of data. The first was published literature and the second was from a commercial source called the 'Assessor database'. The comparison between the pre-Internet results reported by Hauser and Wernerfelt (1990) and our results are shown in Table 4.

| | Our study results (A) | Assessor database (B) | Results of 9 studies (C) |
|---|---|---|---|
| Mean | 2.63 | 3.98 | 4.05 |
| S.D. | 0.26 | 1.29 | 1.52 |
| Welch-Test | | (A vs. B): 0.00** | (A vs. C): 0.00** |

*Table 4. Statistical Welch-Test results[3] (**p<.001)*

The assessor database (B) was based on 23 categories and these ranged from 2.2 to 6.9 with a mean of 3.98. The results of nine studies (C) in the literature reported consideration sets ranging from 2.0 to 8.1 with a mean of 4.05. The results of a Welch-Test to compare data sets A and B, and A and C, show that our results are significantly lower than data sets B and C with p<.001. The pre-Internet studies focused on fast moving consumer goods, which is different to the sectors presented here, although the automotive sector was included in data set C. Whilst it is not possible to control for differences in sectors and methodology to measure the consideration set construct between pre and post Internet studies, the differences are considerable, and prima facie one would expect more extensive online search patterns from theory predictions. *Hypothesis 1 is therefore rejected.*

[3] Datasets from our study can be seen in Table 3, assessor database and the results of the 9 studies can be found in Hauser and Wernerfelt (1990).

Johnson et al. (2004) reported even smaller consideration sets but they included consumers who only visited one website and the difference in results is explained by a methodological difference. Meyer and Stobbe (2010) reported a consideration set size of 3.8, which is higher than the banking sector in the US (2.65), UK (2.75) and also Germany (2.34). However, Meyer and Stobbe (2010) defined a very broad market under the general heading 'financial services'. This included a much larger range of online companies than was used in this study, which partly explains their higher consideration set. In this study, a market was defined using a tighter definition based on direct competitors, i.e. a served market (Caves and Porter, 1978; Porter, 1983). The results are similar to an international study based on 2011 data of the US and UK markets by Holland and Mandry (2013).

The result that consideration sets in the US, UK and Germany are relatively small, and lower than pre-Internet consideration sets is different to theoretical predictions but are consistent with other online panel research. One explanation is that price comparison engines are being used as a substitute for online search with individual suppliers. In this study, this could be true in the airline industry where online price comparison use accounts for 51% of online activity in the US and 75% in Germany. To a much smaller extent, price comparison accounts for 24% for insurance in the US, and for 19% in the German bank market. However, the use of price comparison in eight out of the 12 market sectors studied is 10% or less of total activity within each market, which is relatively small. It is therefore reasonable to conclude that consideration sets are small, regardless of the use of price comparison websites. In order to test the stability of the measurement of the consideration set, the analysis was repeated for 2013 and 2014 in the German market and the results are shown in Table 5.

| | **Average Consideration Set in Germany** | | |
|---|---|---|---|
| **Market Sector** | **2012** | **2013** | **2014** |
| **Insurance** | 2.31 | 2.27 | 2.39 |
| **Banking** | 2.34 | 2.32 | 2.44 |
| **Airline** | 2.44 | 2.45 | 2.51 |
| **Telecommunications** | 2.51 | 2.56 | 2.66 |
| **Automotive** | 2.53 | 2.59 | 2.81 |
| **Grocery** | 2.60 | 2.58 | 2.58 |

*Table 5. Average consideration sets from 2012 to 2014 in Germany*

There is stability in the measurement and in the same way that market shares normally change relatively slowly in mature markets, it is reasonable to expect consideration sets do the same.

## 5.2 Industry Concentration and Consideration Sets

The relationship between the average consideration set and industry concentration expressed as numbers-equivalent of competitors are shown in Figure 2.

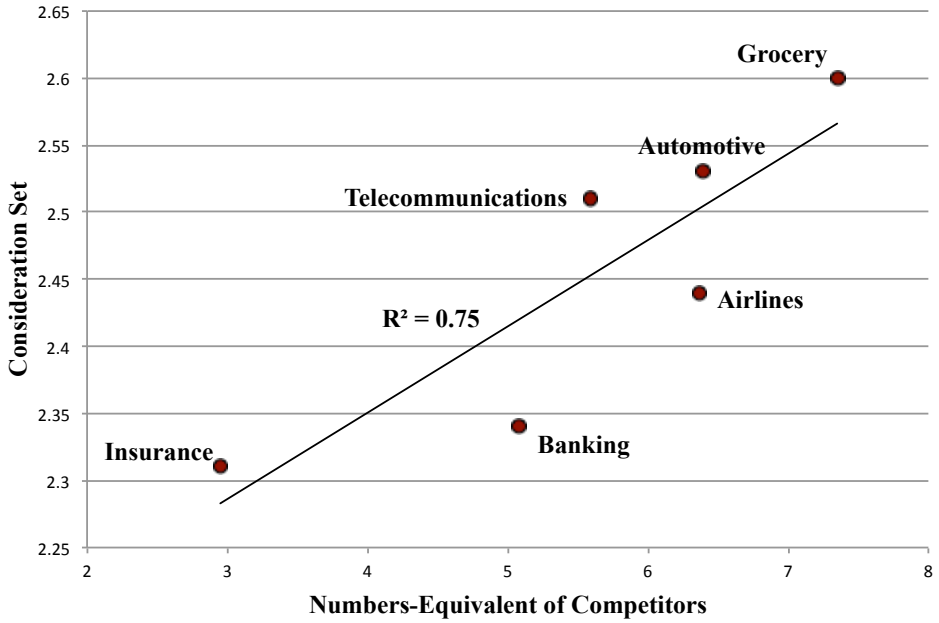


*Figure 2. A cross-sector plot of consideration set against numbers-equivalent competitors in the German online market, 2012.*

The strength of the correlation between consideration set and the numbers-equivalent competitors is 0.75 (DE), 0.53 (US) and 0.72 (UK). *Hypothesis 2 is therefore accepted*.

Some brief comments on the German market are given to provide context for the results. The insurance market is extremely concentrated and insurance intermediaries still constitute a strong distribution channel. It is therefore likely that customers conduct a significant part of their research through their insurance agent. Banking is below the line and therefore has a lower consideration set than the predicted outcome. An important factor here is that the industry concentration was calculated using national bank data whereas consumers make decisions on a localized basis. The prevalence of regional banks in the Germany therefore means that the national calculation overstates the number of banks that are actually accessible on a regional basis. Price comparison engines account for the majority of online activity in the airline market. It is likely that consumers follow a two-stage process of firstly using online price comparison and then using airline websites. In the telecommunications market, consumers are conducting more online research than is indicated by the theoretical model. A possible explanation is that mobile phones are a high-value and high-involvement purchase in a dynamic marketplace. Consumers therefore gain significant benefits from additional search about new technologies, tariffs and handsets. The automotive market is a high-value, high-involvement purchase and there is intense competition. Cars are highly configurable and the market trend is towards personalised purchasing, where consumers can specify an array of options. The online customisation process is likely to be a factor that explains the relatively high consideration set. The grocery market has six significant sized competitors with market shares of between 8% and 23%. It is characterised by price promotions that change on a frequent basis. Although the Internet is not used for online sales and home delivery, it is important for customers to gain price information about offers, which change frequently. A generalization of the results is shown in Figure 3.

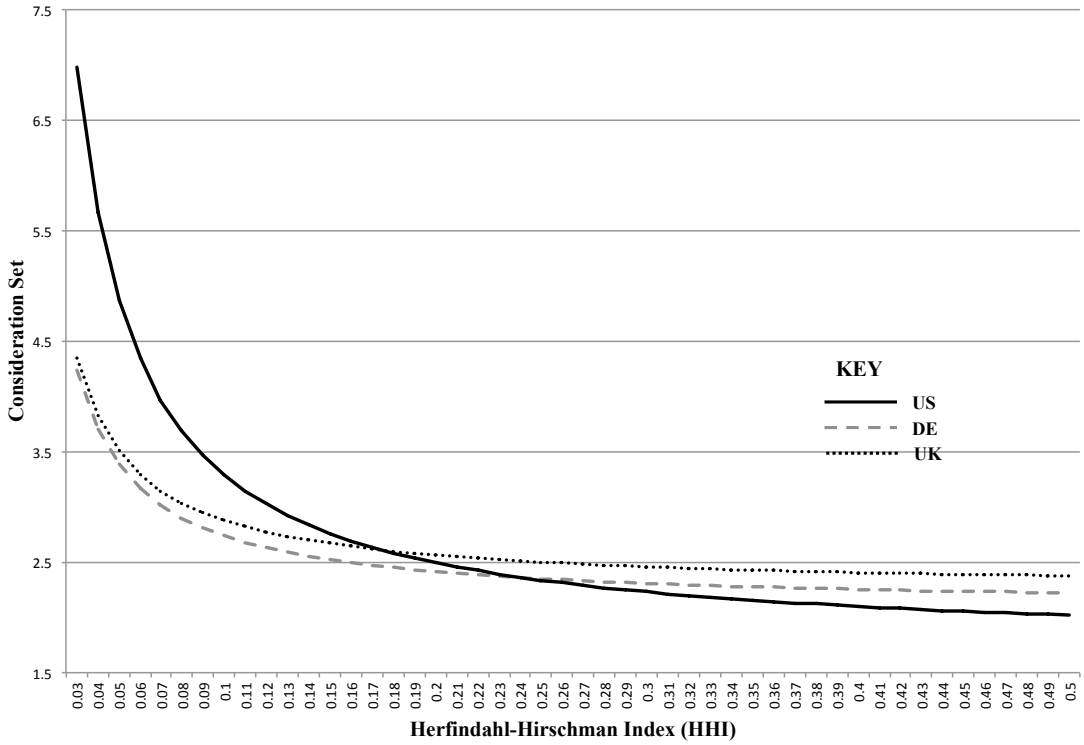


*Figure 3. Consideration set curves defined by HHI for the US, Germany and UK*

The consideration set curves follow a similar pattern for the three national markets. The nature of the relationship between the Herfindahl-Hirschman Index and consideration set is a power law, i.e. it can be expressed as:

$$\text{Consideration Set (CS)} = a\left(\frac{1}{HHI}\right) + k$$

These results are important because they empirically demonstrate that industry concentration limits the extent of the consumer search process based on a cross-sectional analysis of 18 markets in the US, UK and Germany. The markets in the sample have HHIs in the range 0.07 to 0.34. At HHI = 0.2, which represents a numbers-equivalent of five competitors, the consideration sets are almost the same with a coefficient of variation of just 4%, which is significant because many mature markets have a stable market structure of 4-6 major competitors. 15 companies are within the HHI range 0.14 – 0.34, where the coefficient of variation is within the range 4-9%, i.e. relatively close to each other. These initial results suggest that the UK and Germany are similar to each other across the whole range and that all of the markets are similar for the HHI range 0.12 – 0.38. For HHI less than 0.12, US consumers have a higher propensity to search. For concentrated markets, with HHI > 0.38, the US starts to show lower consideration sets compared to the UK and Germany.

From Table 3, the use of price comparison engines does appear to be related to the ease of comparison, which is inversely related to product complexity, and is in agreement with Malone et al. (1989). This result supports the proposition that it is easier to compare simpler products using standardized methods of online comparison. *Hypothesis 3 is therefore accepted*. Further research needs to be conducted into the interaction effects between price comparison usage and primary or direct search with individual competitors' websites (e.g. see Holland et al. 2015 for a discussion of this topic).

# 6 CONCLUSIONS

## 6.1 Consideration Set and Industry Concentration

The consumer search process is relatively narrow, regardless of the use of price comparison engines. The results are consistent with the few published studies on consideration sets that used online panel data (notably Johnson et al., 2004; Zhang et al., 2006; Meyer and Stobbe, 2010) when one takes into account differences in the methodological approaches. This is an empirical contribution to the literature on online search patterns and is one of only a handful of papers that uses online panel data to measure the size of consideration sets.

These results are based on a very large, international database of 18 markets, and run counter to prevailing thinking in economics and marketing, notably influential papers by Bakos (1997, 1998), The Economist (2000), Malone et al. (1989) and Hoffman and Novak (1996). The consideration sets are significantly smaller than the size of consideration sets for consumer products pre-Internet reported in a survey of the literature by Hauser and Wernerfelt (1990). The hypothesis that the Internet increases the size of the consideration set was therefore rejected. It is therefore necessary to consider other theoretical explanations of narrow search patterns:

(1) Consumers feel that they are already familiar with the market, i.e. they have pre-existing knowledge about the products and suppliers (Schmidt and Spreng, 1996); (2) Consumers are conducting research in other channels such as television advertising, agents, dealerships, shops and magazines (Cheema and Papatla, 2010; Klein and Ford, 2003); (3) The perception of consumers is that there are relatively small differences between competing offers in terms of price and functionality. This would reduce the value of the expected benefits, and is consistent with economic theory on highly concentrated markets where the level of rivalry between firms is low (Besanko et al., 2009); (4) Information overload arising from difficulties of structuring and comparing information from competing suppliers increases search costs significantly and therefore discourages extensive search patterns (Browne et al., 2007); (5) Consumers make an online impulse purchase (Madhavaram and Laverie, 2004); (6) Consumers are using research gained through recommendations from their social network, both online and offline, which could reduce the active consideration of suppliers and therefore reduce the consideration set (Brown et al., 2007); (7) Consumers are becoming more sceptical about the value and credibility of online information, particularly in complex product markets (Grant et al., 2007).

The principal theoretical contribution of this research is to link economic theory of industry concentration using the Herfindahl-Hirschman Index (HHI) with consumer behaviour and in particular the consideration set. Economic and strategy theory (e.g. Arndt 1977, Porter 1985, Rhoades 1993, Besanko et al. 2009) is focused primarily on the relationship between industry concentration and the competitive behaviour of competitors within a particular market. The argument made here is that a logical development of the ideas of Stigler (1961) is that the extent of the consumer search process measured by the consideration set should be inversely related to HHI because of increased variety, product differentiation and price variance. By using online panel data it has been possible to explore this relationship using a standard measure of consideration set size across six major market categories in three countries. The proposed consideration set curves model in Figure 3 is a generalization of the results and one that can be tested further using different configurations of markets and countries. One obvious idea is to conduct single sector studies on an international basis, which would test out the concept and remove to a large extent the market context differences, i.e. explore the model for specific markets such as banking, telecommunications and grocery in a range of countries.

## 6.2 Price Comparison

The level of usage of price comparison engines is directly related to the ease of comparison of products between competing suppliers, and the empirical results support the importance of the constructs product complexity and ease of comparison, based on the theoretical model of Malone et al. (1989). The use of price comparison engines in the US and Germany is comparable in all of the

sectors except the automotive market. Specific national market factors regarding pricing and distribution are likely to account for this difference, specifically the fact that in Germany manufacturers exercise much tighter control over consumer prices compared to the US where dealers operate with a high degree of autonomy.

### 6.3 Managerial Implications

The relatively small consideration sets reported here and in earlier research (e.g. Johnson et al. 2004, Zhang et al. 2006) mean that it is vital for companies to actively promote themselves in order to be included in the online search process. This means that online marketing such as natural and paid search strategies, online banner adverts, affiliate marketing and email campaigns are all important promotional activities. Companies should therefore monitor their competitors' online marketing strategies in order to maintain parity with them or risk being neglected from the consumer search process. Small consideration sets also indicate the importance of retaining online visitors, i.e. encouraging repeat visits, thereby leading to an increased probability of subsequent purchase (Bhatnagar and Ghose, 2004). The marketing variables that are of particular importance here are website usability, i.e. design, information content and ease of navigation, and the value of the offer. The differences in search behaviour between buyers also suggests that search is an important segmentation variable and that online retailers should take this into account in the design of websites.

The consideration set curves in Figure 3 have important implications for the regulation of competition because they demonstrate that online search is related to HHI, the standard measure of industry concentration. Given that the online channel now accounts for the majority of search behaviour in many consumer markets, the conceptual framework developed in this paper can be used to gauge, model and evaluate the level of competition in a specific market based directly on consumer behaviour. In an evaluation of a possible merger or acquisition it would be possible to model the likely effects of an increase in industry concentration on consumer behaviour, and make accurate and direct comparisons with similar international markets.

### 6.4 Future Research

The variance of the size of consideration sets is highly correlated with industry concentration in the US, UK and Germany. Is this relationship also true in other national markets? If consumers only consider between two and three competitors, why is this the case? Why do more consumers not conduct an extensive search process? If consumers are using economic heuristics of benefits and costs, then consumers clearly do not perceive that the benefits of additional search outweigh the costs of conducting an extensive search process. A variety of possible explanations have been proposed and further research is needed to investigate their individual role and their collective influence on consumer search. There is significant potential for using online panel data in other economic and marketing research and other applications of this methodology need to be evaluated.

## References

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

Arndt, J. (1977). "Exploring Relationships between Market Structure and Performance in Retailing." In: *Strategy + Structure = Performance: The Strategic Planning Imperative.* Ed. by H.B. Thorelli. Indiana University Press: Bloomington & London, pp. 237 – 246

Bakos, J. Y. (1997). "Reducing buyer search costs: implications for electronic marketplaces." *Management Science* 43 (12), 1676 – 1692.

Bakos, J. Y. (1998). "The emerging role of electronic marketplaces on the Internet." *Communications of the ACM* 41 (8), 35 – 42.

Beatty, S. E. and Smith, S. M. (1987). "External search effort: An investigation across several product categories." *Journal of Consumer Research* 14 (1), 83 – 95.

Besanko, D., Dranove, D., Shanley, M., and Schaefer, S. (2009). *Economics of strategy*. John Wiley & Sons.

Bhatnagar, A. and Ghose, S. (2004). "Online information search termination patterns across product categories and consumer demographics." *Journal of Retailing* 80 (3), 221 – 228.

Bloch, P. H., Ridgway, N. M., and Sherrell, D. L. (1989). "Extending the concept of shopping: An investigation of browsing activity." *Journal of the Academy of Marketing Science* 17 (1), 13 – 21.

Bloch, P. H., Sherrell, D. L., and Ridgway, N. M. (1986). "Consumer search: an extended framework." *Journal of Consumer Research* 13 (1), 119 – 126.

Brown, J., Broderick, A. J., and Lee, N. (2007). "Word of mouth communication within online communities: Conceptualizing the online social network." *Journal of Interactive Marketing* 21 (3), 2 – 20.

Brown, J. J., and Wildt, A. R. (1992). "Consideration set measurement." *Journal of the Academy of Marketing Science* 20 (3), 235 – 243.

Browne, G. J., Pitts, M. G., and Wetherbe, J. C. (2007). "Cognitive stopping rules for terminating information search in online tasks." *MIS Quarterly* 31 (1), 89 – 104.

Bucklin, R. E., Lattin, J. M., Ansari, A., Gupta, S., Bell, D., Coupey, E., Little, J. D. C., Mela, C., Montgomery, A., and Steckel, J. (2002). "Choice and the Internet: From clickstream to research stream." *Marketing Letters* 13 (3), 245 – 258.

Bucklin, R. E., and Sismeiro, C. (2003). "A model of web site browsing behavior estimated on clickstream data." *Journal of Marketing Research* 40 (3), 249 – 267.

Bucklin, R. E., and Sismeiro, C. (2009). "Click here for Internet insight: Advances in clickstream data analysis in marketing." *Journal of Interactive Marketing* 23 (1), 35 – 48.

Bughin, J., Corb, L., Manyika, J., Nottebohm, O., Chui, M., de Muller Barbat, B., and Said, R. July (2011). *The impact of Internet technologies: Search.* McKinsey&Company, High Tech Practice.

Calkins, S. (1983). "The New Merger Guidelines and the Herfindahl-Hirschman Index." *California Law Review* 71(2), 402 – 429.

Callegaro, M., Baker, R., Bethlehem, J., Göritz, A.S., Krosnick, J.A. and Lavrakas, P.J. (2014). "Online Panel Research: History, concepts, applications and a look at the future." In: *Online Panel Research: A Data Quality Perspective.* 1st Edition. Ed. by M. Callegaro, R. Baker, J. Bethlehem, A.S. Göritz, J.A. Krosnick and P.J. Lavrakas. Chichester, UK: John Wiley & Sons, pp. 1 – 22.

Caves, R. E., and Porter, M. E. (1978). "Market structure, oligopoly, and stability of market shares." *The Journal of Industrial Economics* 26 (4), 289 – 313.

Cheema, A. and Papatla, P. (2010). "Relative importance of online versus offline information for Internet purchases: Product category and Internet experience effects." *Journal of Business Research* 63 (9), 979 – 985.

CINIC (2013). *Statistical Report on Internet Development in China*. Report, China Internet Network Information Center. URL: http://www1.cnnic.cn/IDR/ReportDownloads/201302/P020130221391269963814.pdf (visited on 24/11/2014).

ComScore (2009). *Media Metrix User Guide*. ComScore: Reston VA.

Cooper, L. G., and Inoue, A. (1996). "Building market structures from consumer preferences." *Journal of Marketing Research* 33 (3), 293 – 306.

Court, D., Elzinga, D., Mulder, S., and Vetvik, O. J. (2009). "The consumer decision journey." *McKinsey Quarterly* 3: 96 – 107.

Engel, J. F., Blackwell, R. D. and Miniard, P. W. (1995). *Consumer Behaviour*. 8th Edition. The Dryden Press, Harcourt Brace College Publishers: Orlando, FL.

Flynn, L. (1995). *In search of Nielsens for the Internet*. Newspaper article, May 29th, *New York Times*. URL: from: http://www.nytimes.com/1995/05/29/business/in-search-of-nielsens-for-the-internet.html?pagewanted=all&src=pm (visited on 03/03/2015).

Göritz, A.S., Reinhold, N. and Batinic, B. (2002). "Online panels." In: *Online Social Sciences*. Ed. by B. Batinic, U.-D. Reips and M. Bosnjak. Seattle: WA: Hogrefe & Huber, pp. 27 – 47.

Goodhardt, G. J. and Ehrenberg, A. S. C. (1967). "Conditional trend analysis: A breakdown by initial purchasing level." *Journal of Marketing Research* 4 (2), 155 – 161.

Goodhardt, G. J., Ehrenberg, A. S. C. and Chatfield, C. (1984). "The Dirichlet: a comprehensive model of buying behaviour." *Journal of the Royal Statistical Society,* Series A (General) 147 (5), 621 – 655.

Grant, R., Clarke, R. J. and Kyriazis, E. (2007). "A review of factors affecting online consumer search behaviour from an information value perspective." *Journal of Marketing Management* 23 (5-6), 519 – 533.

Hauser, J. R. and Wernerfelt, B. (1990). "An evaluation cost model of consideration sets." *Journal of Consumer Research* 16 (4), 393 – 408.

Hoffman, D. L. and Novak, T. P. (1996). "Marketing in hypermedia computer-mediated environments: conceptual foundations." *Journal of Marketing* 60 (3), 50 – 68.

Hoffman, D. L. and Novak, T. P. (1997). "A new marketing paradigm for electronic commerce." *The Information Society* 13 (1), 43 – 54.

Holland, C. P. and Mandry, G. D. (2013). "Online Search and Buying Behaviour in Consumer Markets," *46th Hawaii International Conference on System Sciences* 2918 – 2927. Maui, Hawaii: IEEE.

Holland, C.P., Jacobs, J.A. and Klein, S. (2015). "An Analysis of Consumer Search Patterns in the German Airline Market using Panel Data." In: *Information and Communication Technologies in Tourism 2015*, Ed. by I. Tussyadiah and A. Inversini. Springer International Publishing, pp. 87-100.

Howard, J.A. and Sheth, J.N. (1969). *The Theory of Buyer Behavior*. John Wiley & Sons: New York.

International Organization for Standardization (2012). *ISO 20252 Market, opinion and social research: Vocabulary and service requirements*. 2nd Edition. Geneva: ISO.

Jepsen, A. L. (2007). "Factors affecting consumer use of the Internet for information search." *Journal of Interactive Marketing* 21 (3), 21 – 34.

Johnson, E. J., Moe, W. W., Fader, P. S., Bellman, S. and Lohse, G. L. (2004). "On the depth and dynamics of online search behavior." *Management Science* 50 (3), 299 – 308.

Jones, S. and Fox, S. (2009). *Generations online in 2009*. Pew Research Center.

Klein, L. R. and Ford, G. T. (2003). "Consumer search for information in the digital age: an empirical study of prepurchase search for automobiles." *Journal of Interactive Marketing* 17 (3), 29 – 49.

Kotteaku, A.G., Laios, L.G. and Moschuris, S.J. (1995). "The Influence of Product Complexity on the Purchasing Structure." *International Journal of Management Science* 23(1), 27 – 39.

Lohse, G. L., Bellman, S. and Johnson, E. J. (2000). "Consumer buying behavior on the Internet: Findings from panel data." *Journal of interactive Marketing* 14 (1), 15 – 29.

Loof, A. and Seybert, H. (2009). *Internet Usage in 2009 – Households and Individuals*. Eurostat publication, Data in focus, 46.

Madhavaram, S. R. and Laverie, D. A. (2004). "Exploring impulse purchasing on the internet." *Advances in Consumer Research* 31 (1), 59 – 66.

Malone, T. W., Yates, J. and Benjamin, R. I. (1989). "The logic of electronic markets." *Harvard Business Review* 67 (3), 166 – 172.

Meyer and Stobbe, T. (2010). "Majority of bank customers in Germany do research online: Findings of a clickstream analysis." *Digital economy and structural change* 79, 1 – 32.

Moorthy, S., Ratchford, B. T. and Talukdar, D. (1997). "Consumer information search revisited: Theory and empirical analysis." *Journal of Consumer Research* 23 (4), 263 – 277.

Napoli, P.M., Lavrakas, P.J. and Callegaro, M. (2014). "Internet and mobile ratings panels." In: *Online Panel Research: A Data Quality Perspective.* 1st Edition. Ed. by M. Callegaro, R. Baker, J. Bethlehem, A.S. Göritz, J.A. Krosnick and P.J. Lavrakas. Chichester, UK: John Wiley & Sons, pp. 387 – 407.

Parra, J. F. and Ruiz, S. (2009). "Consideration sets in online shopping environments: the effects of search tool and information load." *Electronic Commerce Research and Applications* 8(5), 252 – 262.

Peterson, R. A. and Merino, M. C. (2003). "Consumer information search behavior and the Internet." *Psychology and Marketing* 20 (2), 99 – 121.
Porter, M. E. (1983). "Industrial organization and the evolution of concepts for strategic planning: the new learning." *Managerial and Decision Economics* 4 (3), 172 – 180.
Porter, M.E. (1985). *Competitive Advantage: Creating and Sustaining Superior Performance*, Free Press: New York.
Punj, G. and Moore, R. (2009). "Information search and consideration set formation in a web-based store environment." *Journal of Business Research* 62 (6), 644 – 650.
Rhoades, S.A. (1993). "The Herfindahl-Hirschman index." *Federal Reserve Bulletin,* issue Mar, 188-189.
Roberts, J. H. and Lattin, J. M. (1991). "Development and testing of a model of consideration set composition." *Journal of Marketing Research* 28 (4), 429 – 440.
Sayer, A. (1993). *Method in Social Science: A Realist Approach*. London and New York: Routledge.
Schmidt, J. B. and Spreng, R. A. (1996). "A proposed model of external consumer information search." *Journal of the Academy of Marketing Science* 24 (3), 246 – 256.
Shocker, A. D., Ben-Akiva, M., Boccara, B. and Nedungadi, P. (1991). "Consideration set influences on consumer decision-making and choice: Issues, models, and suggestions." *Marketing letters* 2 (3), 181 – 197.
Silk, A. J. and Urban, G. L. (1978). "Pre-test-market evaluation of new packaged goods: A model and measurement methodology." *Journal of Marketing Research* 15 (2), 171 – 191.
Stigler, G. J. (1961). "The economics of information." *The Journal of Political Economy* 69 (3), 213 – 225.
Su, B.-C. (2008). "Characteristics of consumer search on-line: How much do we search?" *International journal of electronic commerce* 13 (1), 109 – 129.
Swaminathan, V. (2003). "The Impact of Recommendation Agents on Consumer Evaluation and Choice: The Moderating Role of Category Risk, Product Complexity, and Consumer Knowledge." *Journal of Consumer Psychology* 13(1&2), 93 – 101.
The Economist (2000). "A thinkers' guide." Magazine Article, March 30th, *The Economist* 355 (8164), 64 – 66. URL: http://www.economist.com/node/298218 (visited 11/12/2014).
Urban, G. L., Johnson, P. L. and Hauser, J. R. (1984). "Testing competitive market structures." *Marketing Science* 3 (2), 83 – 112.
Zhang, J. J., Fang, X. and Liu Sheng, O. R. (2006). "Online consumer search depth: Theories and new findings." *Journal of Management Information Systems* 23 (3), 71 – 95.

## Acknowledgements

This research is supported by the Fonds National de la Recherche, Luxembourg (7842603). The authors would like to acknowledge ComScore for providing the online panel data for this paper. Please see http://www.comscore.com/About-comScore for further information about ComScore. The authors have conducted the analysis and interpretation of the data, and any errors in the paper are the sole responsibility of the authors.