UNIVERSITY OF HELSINKI

https://helda.helsinki.fi

Information search in the internet markets : Experience versus search goods

Basu, Shreya

2018-05-19
Elsevier B.V.
http://hdl.handle.net/10138/315029
Basu, S 2018, 'Information search in the internet markets : Experience versus search goods', Electronic Commerce Research and Applications, vol. 30, pp. 25-37. https://doi.org/10.1016/j.elerap.2018.05.004

Downloaded from Helda, University of Helsinki institutional repository. https://helda.helsinki.fi
This is an electronic reprint of the original article.
This reprint may differ from the original in pagination and typographic detail.
Please cite the original version.

Information search in the internet markets: experience versus search goods

Shreya Basu
shreya.basu@helsinki.fi

University of Helsinki

Abstract

This paper investigates optimal search paths of online shoppers for *experience* versus search goods, as they engage in continuous sequential search for product information. An optimal stopping rule is designed, based on reservation utilities where the instantaneous utility at each search is modelled as a continuous stochastic process. Furthermore, an empirical model validates the theoretical finding using browsing and purchase data from a Finnish multi-product retailer. The main finding is that, experience goods are associated with three times lower search intensities as compared to search goods. A set of hypotheses investigating varying implications of information search online for the two product groups in question, are tested. A proxy for the agents' prior information is calculated based on historic search data via novel methodology from the field of information retrieval, such as *Text frequency-Inverse document frequency*, which exhibits an estimated twelve percent increase in search for search goods, while having no effect on experience goods. Choice probabilities help identify the informativeness of search, which is shown to be inversely proportional to the intensity of search. Finally, the role of personalised recommendations is studied in the context of online search and choice, which has completely opposing effects on the two product types.

Keywords: consumer search, electronic commerce, information retrieval, personalised recommendations

1 Introduction

Nelson's seminal work on economics of information and advertising [25, 24] sheds light on how information about product quality has significant impact on consumer demand and the market structure. Based on prior information about product characteristics, he categorised goods as *search* or *experience*. For search goods, a consumer has the opportunity to inspect an option before purchasing it, whereas for experience goods, quality evaluation is made possible only through purchase and subsequent use. In other words, search goods are dominated by attributes for which full information can be known prior to purchase, while experience goods are dominated by attributes for which full information cannot be known without direct experience. Typically, information search is more expensive for experience goods. This is because sources of product related information such as, word-of-mouth, advertising, user reviews etc. are less expensive than to sample via purchase. Furthermore, Nelson postulates that buyers are more likely to follow recommendations from others for experience goods as compared to search goods. However, with the emergence of electronic commerce and subsequent popularity of the online shopping phenomenon, it has been argued by several researchers [20, 21, 3] that, consumers are now exposed to a broader spectrum of information as compared to the traditional brick-and-mortar stores, which thereby bridges the gap between search and experience goods to a major extent.

This paper identifies relevant factors affecting associated search processes, as it studies how search and experience goods can be distinguished in the context of online search for information. The key questions that are addressed in this paper are: Has the availability of information enabled by the internet markets altered search intensities for experience goods versus search goods, compared to brick-and-mortar stores? Do multi-product online retailers enforce higher search intensities as a result of facilitating increased information acquisition? Do online recommendation agents have a stronger impact for experience goods as compared to search goods, as claimed by Nelson? These questions allow the reader to not only understand the implications of sequential search online in the context of the said types of goods, but also

when a consumer stops searching for more information and makes a decision on whether to purchase a product, given her prior knowledge and information aids, such as recommendation agents. Furthermore, these questions have important implications for firms. Understanding the factors that affect information search, firms could potentially influence the extent of search. *Ex ante*, it is not evident if firms directly benefit when consumers spend more time searching for product information, however, the information obtained could either increase or decrease consumer valuation of the product. As a result, the extent of search is closely related with the optimal price and subsequently profit realisation for a firm.

Studying the same market shift and it's consequences, Klein [20] suggests that by integrating features such as expert reviews, virtual demonstrations, information about history and personality of a brand and its product, internet markets can transform the search processes associated with experience and search goods. Nakayama et al. [23] examine past studies and empirical data on SEC (Search, Experience and Credence goods) ratings to check for evidence for the argument that the web has transformed experience goods into search goods, providing "experience without ownership." From the survey data on hypothetical online purchases they find that, while some products do exhibit ES (experience to search) shifts as a result of online presence of stores and more accessible information structure, others do not, depending entirely on counter-forces of internet advertising. The classification of search and experience goods laid down in the 1970s by Nelson and ever since is based on qualitative definitions. Most of the existing literature, to my knowledge, use qualitative data or empirical methods alone to study search and experience goods. This paper aims to provide a formal framework around the search processes associated with these two types of goods, stemming from their inherent definitions. To formalise a differentiating criterion between the types, a search theoretic model is developed that pins down optimal stopping rules showing that, in equilibrium, consumers search less for experience goods as compared to search goods, i.e., the optimal number of searches prior to purchase is lower for experience goods than search goods in the internet markets.

Commercial interactions online have drastically increased in the past few years, leading to the emergence of visible product networks that explicitly link related or complementary prod-

ucts to one another. Pre-purchase information acquisition is materialised in numerous ways in order to help potential consumers make the most efficient purchase decisions. One of these several mediums happen to be recommendation agents which provides potential buyers with product recommendations based on user-specified preferences, individual shopping history, or choices made by other consumers with similar search profiles or browsing histories. Online search for product information (based on web browsing) and choice in the absence of personalised recommendations detriments the validity of any study on consumer search behaviour in the internet markets. Virtual shopping has proved to reduce search costs for consumers and helped them make informed choices as they have access to a wide range of information without incurring the associated costs of visiting the physical stores [16, 30]. However, there is mostly an enormous amount of information about a particular product from several different sources, which further leads to information explosion. The challenge which buyers commonly face in an online environment is to filter out the most relevant information that helps them make purchase decisions. Xiao and Banbasat [17] construct several propositions based on the effects of recommendation agents on consumer decision making and find that personalised recommendations generated via *collaborative filtering* or *content filtering* improves decision quality of shoppers. Chen et al.[9] collect data from Amazon on books categorising them into bestsellers, popular and less popular books with the aim of studying how recommendations impact firm revenue. Their main findings show that recommendations are mostly positively related to sales. An important issue with regard to recommendations is their perceived quality or usefulness from the consumer's point of view. Aggarwal and Vaidyanathan [2] conducted an experiment to gather information from the subjects about their reactions to recommendations which were generated based on their stated preferences. The experiment was designed for a search good (camera) and an experience good (music CD). The results of their study show that the perceived effectiveness and satisfaction from recommendations are higher for search goods as compared to experience goods. In this paper, we investigate the effect of recommendations on search intensity. The effect of the choice to follow a recommendation and the quality of a recommendation followed are studied in detail and it is shown that these two variables have very different implications when searching for different product types. The findings are aligned with

Nelson's claim that consumers follow recommendations for experience goods more strongly than search goods and that also holds true even for the internet markets. However, it must be noted that recommendations in Nelson's claim meant third-party recommendations, while this study focuses on personalised recommendations that are auto-generated. Including this information not only enriches the data and significantly improves the model, but the findings also throw light on the importance of such informational aids in studying information search. The empirical specifications in this paper takes into account two variables indicating the effects of recommendation agents on search online; the first being the choice to follow a recommendation and the second represents the quality of a recommendation, which basically determines ex-post how similar a recommendation followed, is in its attributes to the good that was bought.

The underlying assumption of complete learning after one alternative is searched, is in line with that of search for the lowest price [11, 26] or search for the best matched alternative [28, 29, 4] in previous literature. According to the original search model developed by Stigler [27], a consumer decides on the number of searches she is going to undertake prior to searching and then choose the best match from the set of alternatives she has examined. An alternative way of understanding consumer search was proposed by McCall [22], Burdett and Judd [7] etc., namely, sequential search, wherein, a potential buyer examines the available alternatives sequentially, that is, she searches one more time if the marginal benefit from an additional search exceeds its marginal cost. In more recent times, Santos et al. test several restrictions that classical search models impose on search behaviour using browsing and purchase data from ComScore. They find that fixed sample search strategy explains observed search behaviour better than sequential search strategy, when shoppers are searching for the best price online [10]. However, in a more recent paper they explain that when consumers learn about the utility distribution by Bayesian updating, there is recall of the previously sampled options which explains why fixed sample search strategies outperforms sequential search strategies in online search behaviour [3]. In the current dataset, over 68 percent of the consumers adopt the sequential search strategy which was concluded from the 'No recall' test consistent with [10], hence we will assume that consumers engage in sequential search with no recall. [1]

[1]Under the null hypothesis of the standard sequential search model, recall of previously sampled alternatives

There is growing literature on the effects of uncertainty and learning on demand for experience goods. Erdem and Keane [14] study brand choice by estimating a Bayesian learning model on household purchases of laundry detergent, where consumers are learning about a single unobservable product attribute. Learning is identified from the behaviour of repeat buyers with different purchase and advertising exposure histories which theoretically explains the formation of brand loyalties. Ackerberg [1] builds on Erdem and Keane[14], estimating a learning model on individual level panel data on consumers' choice to purchase a new brand of yogurt which can be considered as an experience good, since consumers have no prior information about this particular brand. On the supply side, Laband [17] found based on Yellow pages data, that, sellers provide more information for experience goods than search goods, at least in terms of advertisements. In a subsequent paper, Laband[18] found that information provided by the firms is positively related to product prices, based on newspaper advertisements. This paper contributes to the literature on information search, however, differing in framework and assumptions. Unlike Ackerberg[1] or Erdem and Keane[14], in this paper, consumers learn about entire product descriptions while sampling an alternative and decide if the utility they derive from the information is greater than their reservation utility. In case that is true, they buy the product without searching any further. Furthermore, it differs from current literature in empirically looking at a space of more than one product or a controlled set of n products; the data offers hundreds of products which are categorised into two broad groups. The scope of this data is broad, as it records search behaviour accurately in face of information explosion as customers may engage in directed or undirected search.

Apart from investigating optimal search patters in the context of experience and search goods, this paper makes methodological contributions towards developing a measure for "prior knowledge". In studying search and matching models, knowledge and information acquisition prior to purchase play a vital role, and product descriptions allow us to identify the information shocks to a major extent. Novel methodology from the field of information retrieval is used to construct similarity scores between each good searched and the bought good, taking all product information into account. Using *Term frequency-Inverse document frequency* (Tf-Idf)

should not be observed unless the consumer has sampled all of the stores she is aware of.

a new proxy is defined for prior information, adding all similarity scores until the session in which a good was finally purchased. This method has several applications, especially in scoring and ascertaining relevancy. Most of the above studies have focused on a single good across several stores, however, in this paper, consumers search among different products across several categories and brands in a single multi-product store, which makes it prudent to take into account all relevant product attributes into account when developing a measure for prior knowledge. Moreover, the information sets are partitioned into prior and posterior, where the latter is precisely the current information obtained by the agent given her prior knowledge. Novel data on browsing and purchase behaviour of a relatively large set of consumers on a Scandinavian multi-product retailer between March 2014 to March 2016 is used, containing diverse set of search and experience goods. The data described in Section 2 allows us to observe how consumers typically search, percentage of repeat buyers, average search time per session, implications of controlling for the existence of recommendation agents and their impact on search etc. In the empirical part of the paper, certain products are focused on that are typically classified as search or experience goods in literature.

In Section 3, a model of optimal search is developed and it shows that the equilibrium number of searches prior to purchase is lower for experience goods as compared to search goods when consumers are searching for product information sequentially. The optimal stopping rules are based on the assumption that attribute search is continuous and the marginal utility from an additional search can be modelled as an *Itô process* [13]. The online search environment is interpreted in a way that at each search opportunity, consumers get an information shock, positive or negative, that takes them closer to identifying their 'best-matched' good. If the shock is positive then the current viewed product has similar attributes to the ideal product, while if the shock is negative, there is very little or no similarity.

2 Data

Click stream data records entire search paths of agents and is an extremely powerful source of information on consumer behaviour online [6, 8]. It not only monitors purchases, but also tracks consumer actions, such as their search behaviour over time. Using click stream data involves several challenges such as understanding and structuring the online shopping environment with the associated data preprocessing. Figure 1 looks at a typical example of clickstream data which shows the entire customer journey online, from the point she clicks on the first good to the point a purchase is made. At ever search (click), the products are viewed along with their descriptions and the customer updates her preference.

Figure 1: Clickstream data: Example

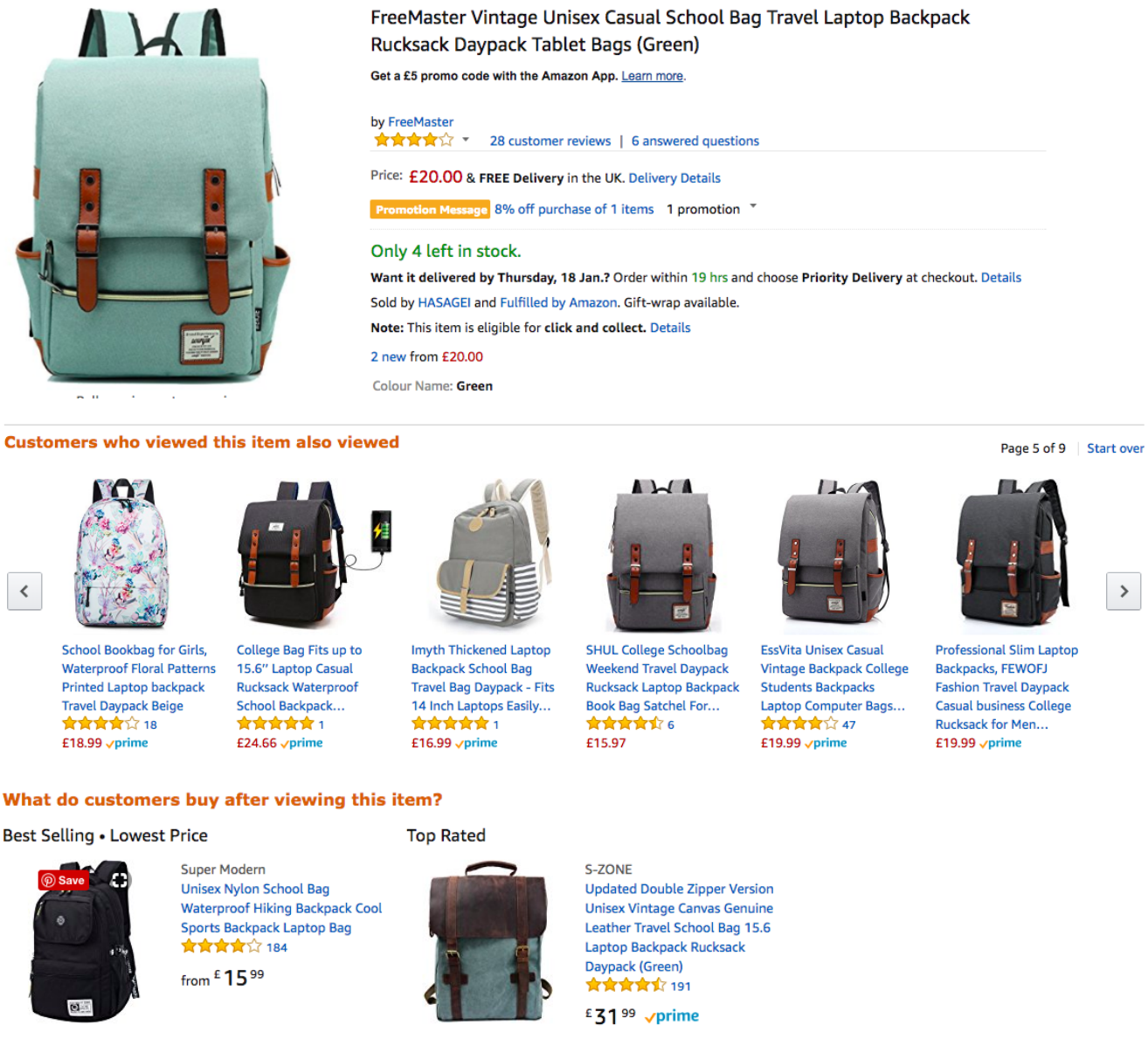


The dataset is constructed by monitoring purchases of anonymised Scandinavian shoppers online between November 2015 and March 2016. The data comes from a leading Scandinavian multi-product retail store that offers sufficiently many types of search and experience goods[2]. The main idea is to examine search patterns of the representative agent across search and experience goods, hence browsing and purchase data is collected only for a single large store, in order to avoid heterogeneity of preferences across stores offering substitutes. Only browsing sessions of users who have eventually made a purchase from the store have been considered. This is because often consumers visit online stores to obtain product related information and then make purchases from physical stores. The results in section 5 are representative of search paths when the consumer definitely buys at the end. Purchases during the period between November 2015 to March, 2016 are observed along with the related browsing behaviour of the unique users, since their first arrival at the store. This means that although all the non-purchase sessions were excluded as data points, but the whole search history was used to construct some of the control variables explained in section 4. Including entire user search histories, data was collected from

[2]The data source cannot be revealed due to non-disclosure agreements.

March 12, 2014 to March 23, 2016. Furthermore, the dataset includes information when an agent clicks on a *recommended* product. Recommendation agents track each customer's search behaviour and shows her a set of relevant products, based on similar search profiles of other customers. Figure 2 gives an example of how recommendations generated via recommender systems typically appear on online retail platforms. For each product viewed, appears one or more panels below with suggestions with statements such as **"Customers who viewed this item also viewed"** or **"What do customers buy after viewing this item"**.The effect of clicking on each such recommended good and of its similarity with the bought good on optimal search for information, is studied at length in the empirical model.

Figure 2: Online recommendations: Example



For accuracy, we look at the goods purchased and consider when in the browsing sessions those particular goods were carted immediately before purchase. This is because once a good is carted, a potential buyer may browse through other goods that he would like to buy, for example, a consumer carts good A in period t and is also interested in buying good B afterwards, such that he searches for B from t+1 through to t+4, and then carts B in period t+5 and finally purchases both goods A and B together in the following period. In this case, from period t+1 to t+4, the consumer does not search for any information related to good A, hence considering carted events rather than purchase events as data points is likely to yield more accurate results. The dataset contains duration of visit, number of brands and products, price of goods and related product descriptions of each good viewed. Table 1 gives a description of the data sample. [3]

Table 1: Descriptive statistics of the multi-product store sample

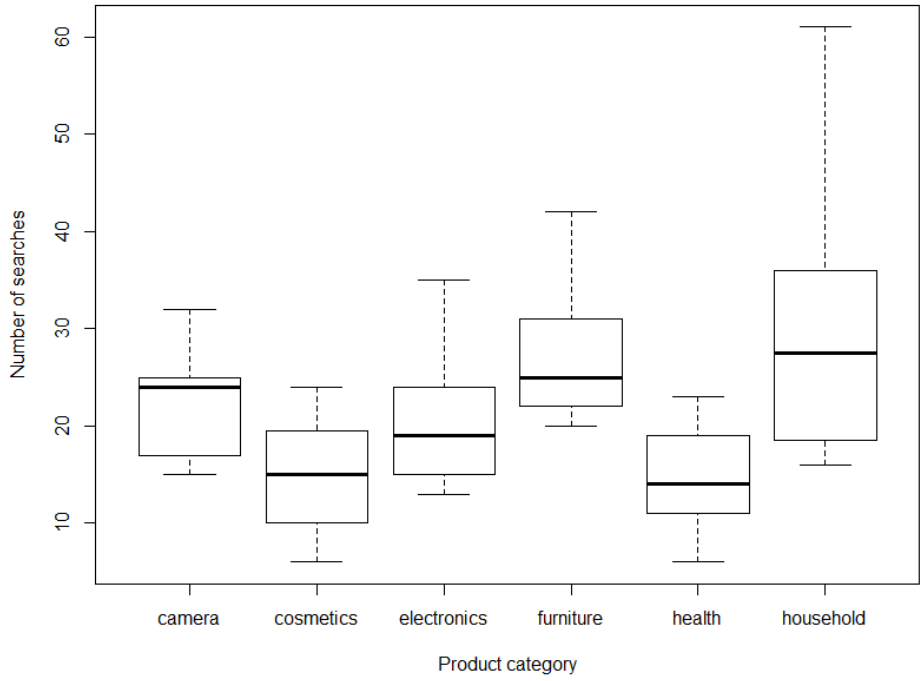
Search patterns	Mean	St. Dev.
Number of goods searched prior to purchase	18.34	16.63
Search time prior to purchase(in minutes)	26.48	26.35
Number of goods searched when no purchase is made	4.81	7.59
Search time when no purchase is made(in minutes)	6.87	14.50
Total umber of goods during November 2015 - March 2016		
Number of unique goods bought	10000	
Number of buyers	15000	
Percentage of repeat buyers	40.5	
Number of unique goods in each category in the sample		
Camera	97	
Cosmetics	278	
Electronics (Mobile phones, Laptops etc.)	521	
Furniture	452	
health	1795	
Household appliances	866	
Total	4009	

As is observed Table 1, approximately 40 percent of the shoppers in this particular sample are repeat buyers. However, for the baseline model, every purchase is considered unique. This means that if an agent buys two products on two different occasions, they are treated as two

[3] 'Number of goods bought' and 'Number of buyers' in Table 1 represent approximate figures since the data sample is proprietary.

unique search paths. The difference in average time spent on search when it leads to a purchase and when it does not, is stark. This motivates the exclusion of non-purchase search paths when analysing differences in consumer behaviour towards search and experience goods. The total number of goods and consumers in the period of five months is much larger than what is used in the dataset for this paper. This is due to several restrictions that had to be imposed for modelling purposes and are mentioned in course of the paper, where necessary. The final sample consists of a total of 4009 observations, each of which is a consumer-product pair, that is, each point a consumer is on a unique product page.

Figure 3: Number of searches before purchase across different products for a super-store over a period of 5 months



In the empirical section of the paper, many of the variables of interest are derived or calculated from the data, however figures 3 and 4 exhibit some other key features that are more directly observable, namely, the extent of search and prices, respectively. It can be observed that, cosmetics and healthcare products are searched the least among all goods in the sample, the number of searches ranging between 8 and 20, and the median being approximately

around 12-14. Cosmetics and healthcare products are classic examples of experience goods and consumers tend to behave similarly while searching for these goods. On the other hand, consumers engage in longer periods of searches for cameras, furniture and other electronics, for example, the maximum number of searches prior to purchase for a household appliance was 60, the median number of searches being approximately around 30. This implicit clustering behaviour of consumers' extent of search across the two categories of goods can be explained simply from their inherent definitions. As the experience good shopper is able to evaluate her match quality with a particular good only after buying it, there is not much incentive for her to search extensively prior to purchase, while search good shoppers have much higher marginal benefits of search.

Figure 4: Prices for different products for a super-store over a period of 5 months

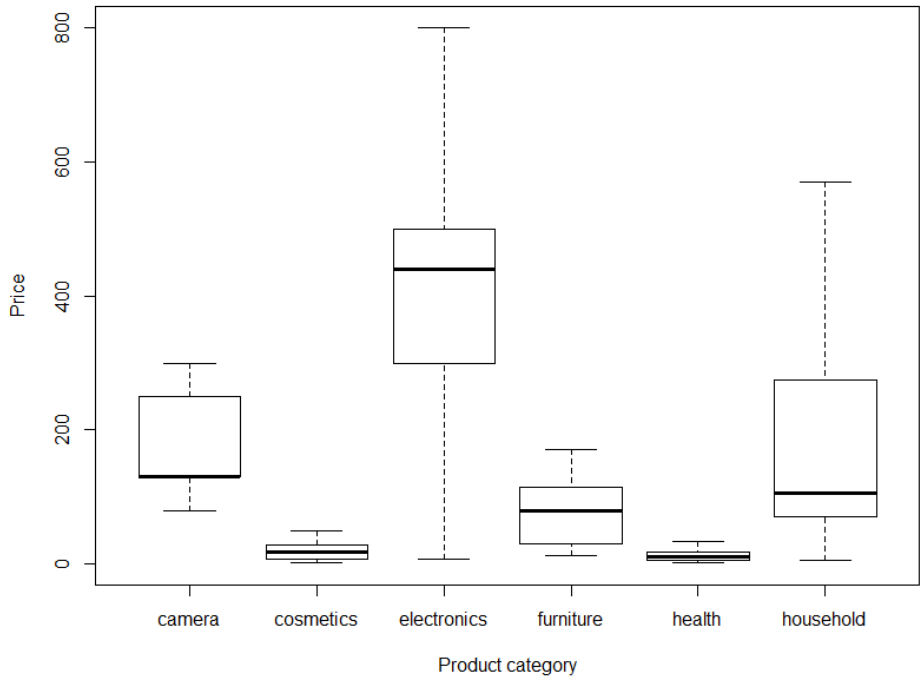


Figure 4 shows that all those under the broad umbrella of electrical and electronic goods have a broad range of prices, the upper limit being rather high. For example, in this sample, the most expensive personal electronic good (laptop/mobile) bought was 800 euros, camera was approximately 300 euros and household appliance, around 600 euros. However, the median

prices for each of these three goods vary quite a bit; while, the median price of personal electronics is closer to the third quartile, it is between the first and second quartile for cameras and household electronics. The experience goods on the other hand, namely, cosmetics and healthcare products are relatively a lot less expensive, the upper limit being approximately around 100 euros.

3 Model

Search and experience goods are categorised based on highly subjective characteristics and it is entirely dependent on individuals as to how much information they consider sufficient before making purchase decisions and the search costs they are willing to incur to obtain that information. Based on several observations, some goods are typically characterised as search and some as experience goods in literature. Electronics, furniture and household appliances are classic examples of search goods,while cosmetics and health products are typical examples of experience goods [15, 19, 23, 25].

Consider a set up where consumers are gathering information about their best matched product, sequentially at the online platform of a multi-product retailer, with no recall. They consider multiple attributes of a product before a purchase a made, such as, colour, size, brand, prices, technical specifications etc. Each product has its own set of attributes and at each search (when a consumer clicks on a product and views it), she learns of these attributes. The realisation of each attribute either increases of decreases the total expected utility. Every product viewed gives incremental information about the potential attributes of her best matched product, assuming she does not know the exact location of the good she wants to buy. Let an agent derive an instantaneous utility, u_t from the information obtained at each search event, t. At every search t, $u(t)$ increases or decreases by q. *Ex ante*, the consumer does not know the value of q at any t, she pays a search cost of time and learns the realisation of q. If we imagine $u(t)$ as a *random walk* which increases or decreases at every t, and if the product attributes and time intervals are extremely small valued, then in the limit $u(t)$ is a *brownian motion*. As there

are many alternatives to sample from and the process of browsing through products online is quick, the search process is assumed to be continuous [5].[4]

Optimal stopping rules are defined for search and experience goods in a search theoretic model framework, consistent with Weitzman [28], where the optimal sequential search procedure is to start searching at the highest reservation utility and to terminate search whenever the maximum sampled utility exceeds the reservation utilities of all remaining unsampled alternatives. The consumer's utility function can be written as follows, given, $u(t)$ is the instantaneous utility at search event t :

$$U = E \int_0^\infty e^{-\rho t} u(t) dt$$

where ρ is the discount rate. u is considered here to be a state variable that changes stochastically with the number of searches, and hence, the change in utility from an additional search, du is modelled as an Itô process, a continuous time stochastic process, such that,

$$du = \mu dt + \sigma dz$$

An Itô process is a generalised brownian motion where parameters μ denote the drift rate and σ^2 denotes the instantaneous rate of variance or diffusion rate. The drift rate that exhibits the average rate of growth in the long run, must be positive in the current set-up. This is because with each marginal search, a potential buyer obtains information about her best-matched good, so it must be the case that utility from getting additional information is growing positively on an average, otherwise there exists no incentive to search over time. In the economic sense, σ may be interpreted as indicating the informativeness of search, such that a high σ would imply less informative search and vice versa. In a way it captures how the agent processes relevant information from the entire search path, as she updates her reservation utility after each search.

The central assumption of the model is that, informativeness of search which is represented

[4]Branco et al. study optimal search for product information, where consumers search across multiple attributes of one product. With an infinite number of attributes, each providing an infinitesimal amount of information, one can assume that the expected valuation follows a Brownian motion while searching. They assume attributes to be sufficiently small-valued, hence search to be continuous, such that marginal utility from attribute search gets infinitely smaller as the number of searches go to infinity. They model the u process as a *brownian motion* in the limit.

as the inverse of σ, is lower for experience goods as compared to search goods. This is because, by definition, potential buyers are better informed about the attributes of search goods, hence it is reasonable to assume that they will search for the most important attributes (ones they are unaware of) first. Attribute search takes place in decreasing order of importance for search goods, so the change in utility from each marginal search is changing consistently. However, for experience goods, the first time potential buyers have very little to no information about their best matched product, hence their search for information is undirected and the change in utility is likely to be inconsistent across periods. Apart from basing the assumptions on qualitative definitions of these product classes, we look into the data to verify any existing search patterns. Figure 5 shows two representative agents' search paths, one that buys an example search good (electronics) and one that buys and example experience good (clothing) at the end. The first row observes the customer journey of the latter, from the first search event until a purchase was made, and the send row observes the former. Evidently, the second customer engages in directed search (low variance), while the first customer browses through markedly different products from the one that she buys at the end. Performing several exposed checks in the current data set, combined with the intuitive classification of the product types, we assume that experience goods are related to lower search variances as compared to search goods.

Figure 5: Search paths: experience versus search good



If a consumer stops search at any point without buying the good, her instantaneous pay-off is zero. The stopping rule is to terminate search and buy the good, whenever utility derived from the good being 'sampled' is greater than or equal to a cut-off level, $u^\star$. In case utility is strictly less than $u^\star$, it is optimal for an agent to continue her search.

$$u^\star = \begin{cases} u \geq u^\star & \implies \text{stop search and buy} \\ u < u^\star & \implies \text{continue search} \end{cases}$$

Let $u' = u + du$ and $t' = t + dt$, where dt is a very small change in time. Then the value function can be written as,

$$V(u,t) = \max[udt + \frac{1}{1+\rho dt} EV(u',t')]$$

This is the Bellman equation that defines the consumer's optimal decision rule, given that the utility derived from having searched for t periods is udt and the present value of searching in future is $EV(u',t')$ times the discount rate, ρ. In the continuation region, the expected utility reduces to

$$\rho V(u)dt = udt + E(dV) \tag{1}$$

Now, since du follows an Itô process, we can apply the Itô's lemma,

$$dV = \left[\frac{\partial V}{\partial t} + \frac{\partial V}{\partial u}\mu + \frac{1}{2}\frac{\partial^2 V}{\partial u^2}\sigma^2\right]dt + \frac{\partial V}{\partial u}\sigma dz$$

$$\implies E(dV) = \left[\frac{\partial V}{\partial t} + \frac{\partial V}{\partial u}\mu + \frac{1}{2}\frac{\partial^2 V}{\partial u^2}\sigma^2\right]dt$$

since dz is a Wiener process. Plugging the above into (1), the reduced equation is obtained,

$$-\rho V + \mu V' + \frac{\sigma^2}{2}V'' = 0$$

whose solution can be written as,

$$-\rho e^{ru} + \mu r e^{ru} + \frac{\sigma^2}{2}r^2 e^{ru} = 0$$

$$r = \frac{-\mu \pm \sqrt{\mu^2 - 2\sigma^2\rho}}{\sigma^2}$$

The consumer's objective function can be expressed in the following way:

$$U = E \int_0^\infty e^{-\rho t} u(t) dt = \int_0^\infty e^{-\rho t} [\mu t] dt = \frac{\mu}{\rho^2}$$

Then, the particular solution obtained is of the following form,

$$V(u) = \frac{\mu}{\rho^2}$$

such that the general solution is:

$$V(u) = \frac{1}{\rho}(u + \frac{\mu}{\rho}) + c_1 e^{r_1 u} + c_2 e^{r_2 u}$$

At every t, the consumer must decide whether to search one more time or to stop search and purchase the good. $u^\star$ pins down the optimal value of the control variable, when shoppers derive utility from searching for product related information. In order to find a solution for the optimal $u^\star$, a set of boundary conditions are required. The following depicts the *value matching* condition which ensures continuity of the value function, $V(u)$. In economic terms, it implies that utility derived from search is large enough for the consumer to be indifferent between searching once more and buying the good. $u^\star$ is the upper bound such that when u reaches $u^\star$, the consumer buys the good with certainty, getting utility $u = u^\star$.

$$V(u^\star) = u^\star \tag{2}$$

The following depicts the *smooth pasting* condition which is the partial derivative of $V(u^\star)$ which ensures there are no kinks at the boundaries. The condition is implied by the fact that the consumer maximises $V(u)$ for all u [5] [12].

$$V'(u^\star) = 1 \tag{3}$$

Substituting the boundary conditions into the general solution yields the following,

$$V(u^\star) = \frac{\mu}{\rho^2} + c_2 e^{r_2 u} = u^\star$$

$$V'(u^\star) = r_2 c_2 e^{r_2 u} = 1$$

(since, $\lim_{u\to\infty} V'(u) = \frac{1}{\rho}$, then from the boundary conditions, $c_1 = 0$). The optimal stopping rule is therefore represented by the following expression,

$$u^\star = \frac{\mu}{\rho^2} - \frac{\sigma^2}{\mu + \sqrt{\mu^2 + 2\sigma^2\rho}} \tag{4}$$

$u^\star$ depicts the threshold value or the stopping rule itself; when $u > u^\star$, it is optimal to stop search and buy the good and when $u \leq u^\star$, search continues for a more suitable alternative. Now, consider two separate Itô processes for the two type of goods in question, such that the across t, for search goods is, ${\sigma_s}^2$ and that of experience goods is, ${\sigma_e}^2$

$$du = \mu dt + \sigma_s dz$$

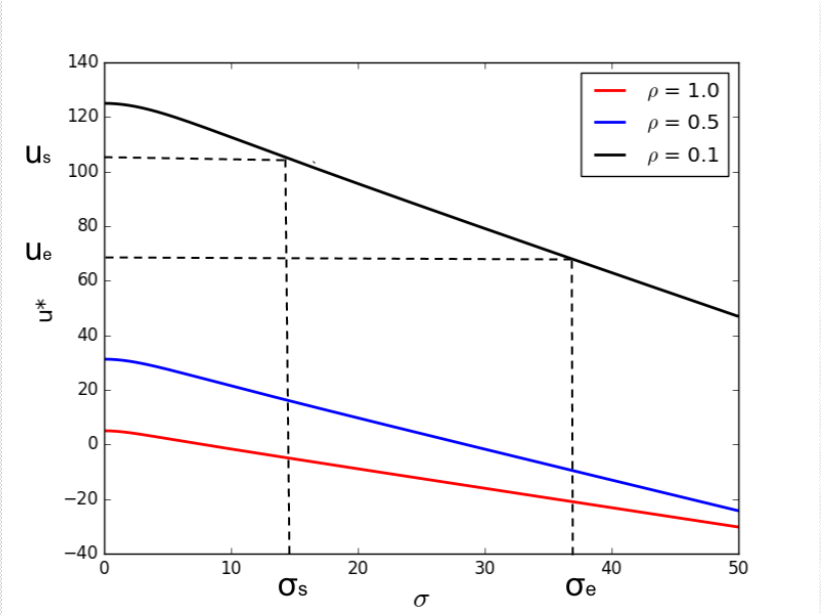
$$du = \mu dt + \sigma_e dz$$

such that ${\sigma_e}^2 \geq {\sigma_s}^2$. Equation (4) along with this condition immediately gives the following proposition.

Proposition 1. *The optimal level of utility to stop search and buy the good being sampled is higher for search goods as compared to experience goods, i.e.,* $u_e^\star < {u_s}^\star$*, as long as* ${\sigma_e}^2 \geq {\sigma_s}^2$ *holds true.*

This implies that consumers have higher search intensities for search goods as compared to experience goods which is explained by the difference in informativeness of each additional search. Therefore, the continuation range is much larger for former type and as a result, the time to buy is also higher on average compared to the latter type. Figure 3 exhibits changes in the path of $u^\star$ as σ changes for varying levels of ρ. It is evident that increasing the marginal

Figure 6: Optimal search paths for varying discount rates



benefits of search, by making it more informative, in turn decreases the level of utility at which the consumer is indifferent between buying the good and searching one more time. So, the main finding is that, the extent of search for experience goods is less than search goods, simply because, search for experience goods may be less informative as consumers arrive to the store with little to no prior information and can only evaluate match quality through consumption.

4 Empirical Analysis

Section 3 models consumer search behaviour across two product types and formally shows that the extent of search, prior to purchase, varies between search and experience goods when a consumer engages in information search online. In line with *Proposition* 1, an empirical model is developed that aims to validate or disprove this finding, which further motivates the choice of the dependent variable, *Number of searches* in the model. The objective is to investigate thoroughly the factors that affect search, and how they vary across the two product groups, which leads to the following key hypothesis:

Hypothesis 1. *Search intensities for experience and search goods are the same when consumers engage in online search for information*

The above is in line with several studies in the past (mentioned in section 1) that claim that the internet markets or online retailers allow customers access to unconstrained information that enables them to search for experience and search goods equally intensively, such that it bridges the inherent gap between these two product types. Considering the key assumption that pins down our primary finding, it is evident that this hypothesis hinges largely on the role of informativeness of search in determining the optimal number of searches. In order to identify the variance of information shocks, let us start by defining the probability of agent n buying a good at search i, by P_{ni}. Let the utility derived by agent n from search i be denoted by, U_{ni}. Then, agent n chooses good i over j if $U_{ni} > U_{nj}$ and the probability is given by,

$$P_{ni} = Prob(U_{ni} > U_{nj}, \forall j \neq i)$$

U_{ni} is decomposed as $U_{ni} = V_{ni} + \epsilon_{ni}$, where, ϵ_{ni} captures all the factors that affect utility but are not observable by the researcher. It is assumed that ϵ_{ni} is i.i.d type- I extreme value distributed, which represents idiosyncratic taste of consumer n for good i,. The part of utility observed by the researcher or the representative utility is contained in V_{ni}, such that,

$$V_{ni} = \beta price_i + \theta H_{ni} + \lambda S_{nii^*}$$

where, H_i is a vector of environmental factors which includes dummies for holiday season or ongoing dicounts/sale periods and S_{nii^*} represents the similarity score between each good being sampled and the the bought good.

$$P_{ni} = Prob(\epsilon_{nj} - \epsilon_{ni} < V_{ni} - V_{nj} \forall j \neq i)$$

$$\implies P_{ni} = \frac{e^V_{ni}}{\sum_j e^V_{ni}}$$

So, agent n buys good i with probability P_{ni} when the decision to buy depends on information derived from attribute search. The difference in choice probabilities across search, $P_{n(i+1)} - P_{ni}$ identifies the marginal utility of every additional search opportunity. Furthermore, the variance of $P_{(i+1)} - P_i$ for each customer n describes how informative search has been in identifying

one's optimal match, thereby affecting the optimal number of searches prior to purchase. The informativeness of search is calculated by taking the inverse of search variance for the full journey of each customer. Different choice models can be applied to evaluate choice probabilities depending on the specifications of the density of unobserved factors. Assuming that unobserved factors are uncorrelated over choices and for the sake of consistency with existing literature, a logit model is used to estimate the probability of purchase at every search event for each customer. The weights of the control variables obtained from a standard logistic regression are used to calculate P for customer n for the good being sampled at search event i. As prices and similarity scores (key control variables), among other factors, vary across products, so do their weights, which contributes towards varying choice probabilities. The variance of these marginal probabilities ($P_{n(i+1)} - P_{ni}$) signals the informativeness of a customer's journey of information search. [5]

Additionally, the model aims to investigate if being more knowledgeable of product characteristics, owing to ease of information search online, makes search more efficient. In other words, the model aims to test if there is an inverse relationship between prior information and search intensity, which leads to the following proposition:

Hypothesis 2. *Increased product knowledge (through prior search) leads to increased search intensity, consistently across all product types*

One of the most prominent factors that classifies goods as search or experience, is clearly the amount of prior information, hence a reasonable assumption is that prior information impacts time to buy. In order to calculate prior information from the data, first we observe the browsing sessions of each unique buyer from the time of her first arrival to the store, until she carts the good (or goods) that is finally purchased. In each search until purchase of the said good, there is transfer of information, only some of which is relevant to the buyer. So, in order to understand how information acquisition influences purchase decisions, search events that could potentially inform consumers about the particular good that she has purchased, must be filtered. The idea

[5]The sample size was restricted only to those agents who engaged in at least four searches in order to get a reasonable measure for variance.

is to collect all such search events that help a consumer make his final choice. To this end, a measure of similarity between products that have been viewed before the final purchase and the bought good is defined, based on the related descriptions. The method used to calculate the similarity scores are described in detail in the next section.

4.1 Calculating information based similarity scores

Each unique product has a description of the set of attributes associated to it in the data. These attributes typically describe the category or the type of every product, along with its key features. However, entity-related search on web data is non-trivial, in the sense that the attributes are described in different ways for different products and collecting the most important ones out of them is not straight-forward. The objective is to find similar products based on their intrinsic attributes, hence its essential to find commonality between every word of the attribute vectors (combination of several keywords). For example, an agent buys camera B in the period, t. In period t-2, say, she views camera A and in period t-4, she views a mobile phone. Now suppose, product descriptions for each of the three goods are described as in Table 2.

Table 2: Product attribute example

Product	**Attribute**
Camera B	'camera B with new features'
Camera A	'camera A'
Mobile phone	'mobile phone with new features '

It is evident that similarity measures constructed based on the keywords in the attribute vectors is rather misleading. This is because, sampling camera A before purchasing camera B must give a buyer more information about her final purchase, rather then sampling a mobile phone. However, comparing the keywords in each of the vectors shows that, based on common keywords,

$$similarity\ score_{\{Camera\ A,\ Camera\ B\}} = 1$$

$$similarity\ score_{\{mobile\ phone,\ Camera\ B\}} = 3$$

Therefore, constructing similarity measures from the product descriptions, as is, yields highly inaccurate similarity measures between goods and hence constructing a statistic for prior information from these measures can lead to distorted results. In order to mine the relevant product attributes from the entire vocabulary of attribute elements, several methods from the field of text analysis and information retrieval can be used. In information retrieval, the task is to find important or relevant words from a large body of text or information set. *Vector space models* are most commonly used in computational linguistics and cognitive science, in the context of information retrieval.

A product description (document) is represented as a vector of it attributes (terms). The term frequency inverse document frequency (tf-idf) model is the most classical vector space model, wherein tf-idf is a measure used to evaluate how important a word/term is in a document amongst the entire collection of words/terms. It is often used as a weighting factor, as the importance of a word increases proportionally to the number of times it appears in the document, but is offset by the frequency of the word in the entire collection. Let t_y be the term frequency or the raw frequency of the term t in a document y, and y'_t be the document frequency or the number of documents in which term t appears. The inverse document frequency means how important a term is in a given collection, that is, how common it is across all the documents. Finally, let n be the total number of documents (product descriptions) in the data. Then the term frequency inverse document frequency weight, ω is calculated by,

$$\omega = t_y \times (y'_t)^{-1}$$

where,

$$(y'_t)^{-1} = log \frac{n}{y'_t}$$

Once the relevant words from the descriptions of the products have been retrieved, a measure of similarity between the attributes must be constructed. The *cosine similarity* between two vectors or two documents in the vector space is a measure that calculates the cosine of the angle between them. This is a comparison metric between two documents on a normalised space as it

not only takes the magnitude of each word count (tf-idf) of each document into consideration, but also the angle between the documents. Let us consider vectors a and b, then the cosine of these two vectors is given by the Euclidean dot product:

$$a \cdot b = || \ a \ || || \ b \ || \ cos \ \theta$$

$$S(a, b) = \frac{a \cdot b}{|| \ a \ || || \ b \ ||}$$

The above is the metric for cosine similarity, (S(a,b)) shows how correlated two goods are based on their product descriptions. The cosine similarity of each product in the search history with the bought good is then calculated; expectedly, similar products have higher degrees of cosine similarity compared to products differing in attributes. Observing similarity scores between products in several random samples of goods, only those products are collected from the browsing history that have a similarity coefficient of 0.4 or above. This seems to be a reasonable threshold, as browsing through products that have markedly different attributes compared to the purchased good, do not essentially give potential buyers relevant information. Hence, towards the aim of constructing a variable for prior information related to the bought product, only those goods are considered that are evidently similar to the final purchase. For each purchase, the entire search history is observed in order to construct similarity scores for each sampled good with respect to the final purchase. Finally, summing over all the similarity scores greater than 0.4 until the purchase session finally gives us a proxy for the level of prior information of the agents.

Online search behaviour cannot be fully captured without understanding the effects of personalised recommendations on consumer choices in the event such a feature exists on an e-commerce platform. It is important to study the outcomes when a consumer's purchase decision has been in some way influenced by the recommendation agents compared to when it has not. The effect although is difficult to isolate, as recommendations might prolong the search process by exposing more of the product catalogue on one hand, on the other hand, it may suggest a potential buyer a well-matched product that one buys immediately, thereby reducing search time and effort. In order to capture this particular effect, a categorical variable, indicating the presence

of a recommendation in the search path is introduced, which takes the value 1 if a purchase has been induced by a click on a recommended product during the final browsing session (that has ended in a purchase), or the value 0 if the buyer has rejected all recommended goods and sequentially searched until he has found his best-matched product. Recommendations aid the search and purchasing process for online shoppers, hence it is likely to have an effect on the time to buy, however according to prevalent claims in literature, experience good shoppers are more affected by these recommendations comparatively. The firm's incentives to recommend must be noted here: not only does a firm increase the likelihood of buy to making it an almost sure event by recommending the best matched product, but also keeps a potential buyer engaged in searching more by recommending a product that is matched just sufficiently well, such that a shopper believes that he is in his continuation region, or at least moving towards it. Therefore arises the need to study the effects of recommendations generated online and whether they can be viewed as a source of information from the perspective of potential buyers.

Furthermore, the 'quality' of a recommendation or its degree of match with each consumer's preferences, influences her search path and in turn often decides whether or not she is in the continuation region. In order to account for this effect, a variable is constructed that sufficiently indicates the quality of the recommendation clicked on at any point during the final browsing session that induces the final purchase. This is done in the same way as the attribute-based similarity scores are calculated for the prior information. This means that if in the final session, an agent clicks on a recommended good, the attributes of that particular good is matched against the set of attributes of the bought good and depending on the match, a measure for the quality of the recommendation is arrived at. This motivates the final hypothesis:

Hypothesis 3. *Recommendations agents impact search intensities associated with experience goods more strongly than search goods*

The following equation models the relationship between number of dependent variable, $Y =$ *search intensity* and the vector of independent variables, $X = [P_j, I, K, (P_j * E), (I * E), (K * E)]$, where, P_j represents price for alternative j, I represents informativeness of search which is calculated from the choice probabilities, $I = (\sigma^2)^{-1}$, and K represents prior information that

agent n possesses when she samples alternative j,

$$\mathbf{Y} = \mathbf{X}\beta + \epsilon \tag{5}$$

The indicator variable, E takes value 1 if a good is of type, experience and 0 if it is of type, search. Furthermore, interaction terms are added to observe the difference in the effect of the controls on number of searches across the two groups of consumer goods.

5 Results

The model is estimated using three classic examples of search goods (personal electronics such as mobile phones, laptops etc., furniture and household appliances) and two classic examples of experience goods (health and beauty products) for n shoppers, who have in total purchased 4009 products over a period of five months.

Table 3 (1) shows OLS estimates of search intensity on several important constructs across search and experience goods, while in (2), variables representing the choice to follow a recommendation and its quality are included in the baseline model. It is evident that including the effects of recommendation agents significantly improves the model, which shows the importance of these shopping aids in search and purchasing behaviour online. An increase in price has a significant positive impact on search intensity for search goods which implies, that more expensive the goods are, the more thoroughly they are searched for. On the other hand, an increase in prices for experience goods reduces search, although nominally.

An increase in prior information has significant impact on the number of searches prior to purchase for search goods, but no impact on the extent of search for experience goods. As product knowledge increases, search intensity increases by approximately 12 percent for search goods. This result allows us to reject *Hypothesis 2*: increased product knowledge that is facilitated by the unconstrained availability of information online, enables higher search intensity for search goods, but has no impact on experience good. This is because, in case of

search goods, consumers direct their search towards important product attributes they might be unaware of, because they are able to identify the order of importance of these attributes. As more information only exposes more of the available options, they are likely to search more. As a result, more relevant information keeps customers engaged in search which increases the likelihood of a purchase per customer. This also sheds light on the firm's incentive to employ recommender systems on their retail platforms: by providing relevant product information, firms are able to increase customer engagement and improve chances of a successful transaction, at least for search goods. On the other hand, information search does not enable consumers determine match quality for experience goods. So, even for an increase in prior product knowledge, consumers have no effect on their search path, as prolonging their search does not enable them to learn about their optimal match. Browsing through more product characteristics simply does not pose as an incentive to search longer as it does not guarantee improved match quality for experience goods.

The coefficient for variance of search is highly significant and is inversely proportional to the amount of search, shoppers engage in. The inverse of variance represents informativeness of search, in identifying an agent's optimal match. This implicitly affects the marginal benefits of searching for product information. Uninformative search reduces the extent of search across both types of goods, which is consistent with the optimal stopping rule in section 3. However, findings from the empirical model suggest that, the impact is much larger in magnitude for experience goods. The extent of search associated with experience goods decreases three times more than search goods for an unit increase in variance. Given the assumption that experience goods are associated with higher search variance, even if search was equally as informative for both product types, a typical consumer would optimally search at least three times less for experience goods than search goods, prior to purchase. This result allows us to reject *Hypothesis 1* and deviate from prevalent claims in literature that the internet markets negates the difference between search and experience goods. As experience good shoppers have little to no information about true match quality at the beginning of search, the variance of the information shocks they face over time is rather high. This implies that their search is not informative of the optimal match and they have less incentive to continue searching. Due to limited knowledge, they

may browse through goods that may have no similar attributes to the purchased good at all. This further reinforces the assumption in Section 3 that the variance of the information shocks from an additional search is higher for experience goods than search goods. In other words, a typical experience good shopper does not know of the most important attributes of his preferred product, hence has less incentive to search more, as opposed to search good shoppers, who are well informed of their preferences, hence search efficiently when they choose to buy.

Inclusion of explanatory variables that capture the effects of personalised recommendations on the extent of online search for information, improves the model considerably, which sheds light on the importance of recommendation agents on consumer search in the internet markets. The quality of a recommendation that has been followed (clicked on) by a shopper is not significantly different from zero for search goods, which implies that recommending a good with a higher match quality does not have any implications on search intensity as far as search goods are concerned. However, the quality of a recommendation followed is highly significant for experience goods and improving the quality reduces search by 15 percent. Firms have high incentives to design recommendation agents in a way that increases efficiency in how consumers shop online. What this result essentially means is that the lack of product information can be replaced by recommendations that inform potential buyers of their optimal match. Although engaging potential buyers for longer search periods increases the likelihood of purchase, in case of experience goods, where the incentives to search are less, an effective way of ensuring high conversion rates is to improve the quality of recommendations.

On the other hand, the choice of following a recommendation significantly affects both search and experience goods; when consumers choose to follow a certain recommendation, it increases the optimal number of searches by an estimated 16.8 percent for search goods and decreases the number of searches almost by 9 percent for experience goods. The former is indeed a surprising result, since recommendations are expected to inform shoppers of their best-matched product, thereby reducing search. However, if the recommendations are such that the consumer is being informed of close substitutes to her expected match, she is likely to investigate more options before deciding on a particular good. This is indeed true for search

goods, as investigating substitutes enables a buyer to determine her optimal match. Online recommendation agents have significantly opposing effects on search for the two products groups, impacting search goods more that experience goods, which rejects *Hypothesis 3*. In previous literature, recommendations have only been perceived, studied and analysed as decision aids that make the shopper more efficient, or reduce her search time. Contrarily, this result provides reason to investigate the behavioural mechanisms that firms use to employ in designing the recommendation algorithms. Keeping shoppers engaged on a website is beneficial for firms, although, it is often the case that when a potential buyer has spent considerable amount of time viewing a particular (type of) product, the recommendations thereafter are unrelated. This is contrary to the principle of recommendation agents per se, but firms may design them in such a way only to maximise the likelihood of the purchase at every search, due to lack of other relevant visible alternatives. For example, while searching for a coffee machine, if a buyer is shown recommendations of coffee filters, it is suggestive of the fact that there exists no better options than what she is viewing at present, which in turn leads her to confirm the purchase without getting distracted by other options. The challenge for firms is to design the cut-off point between when recommendations are useful sources of product related information and when they might distract the buyer by prolonging search.

Table 3: Estimates of search

	Dependent variable: Number of searches	
	(1)	(2)
Price	0.021***	0.018***
	(0.003)	(0.002)
Prior information	0.118***	0.131***
	(0.012)	(0.009)
Variance of search	−1.753***	−1.940***
	(0.087)	(0.095)
Recommendation quality		−0.101
		(0.007)
Recommendation (1=used:;0=not used)		0.168***
		(0.028)
Price*E	−0.032***	−0.102***
	(0.005)	(0.016)
Prior information*E	−0.192	−0.116
	(0.026)	(0.014)
Variance:E	−4.092***	−4.470***
	(0.633)	(0.324)
Recommendation quality*E		−0.151***
		(0.012)
Recommendation*E		−0.254***
		(0.038)
Constant	2.778***	2.696***
	(0.024)	(0.027)
Observations	4,009	4,009
Adjusted R^2	0.363	0.404

Note:
*p<0.1; **p<0.05; ***p<0.01
(1) exhibits the estimates of search;
(2) exhibits the estimates of search with recommendations
E=1 for experience goods; 0 for search goods

Table 4: Summary of results

Hypothesis	**Result**
1. Search intensities for experience and search goods are the same when consumers engage in online search for information	Reject
2. Increased product knowledge (through prior search) leads to increased search intensity, consistently across all product types	Reject
3. Recommendations agents impact search intensities associated with experience goods more strongly than search goods	Reject

6 Discussion

Search intensities associated with experience goods is at least three times lower than search goods

This paper investigates the extent of online search that shoppers engage in as they search for product information for specific type of consumer goods. They obtain information by searching through webpages featuring different goods and their associated descriptions. By using detailed data on browsing and purchase behaviour of shoppers, as well as product descriptions of every good searched, optimal search paths of potential buyers across search and experience goods are studied. One of the primary findings is that search good shoppers search at least three times more as compared to experience good shoppers in equilibrium when they engage in continuous sequential search. This result further validates the finding in Proposition 1 empirically. The theoretical model in Section 3 aims to formally pin down a criterion that discriminates between experience and search goods in the context of information search in the internet markets. Optimal stopping rules that are derived from imposing boundary conditions allows the reader to understand search for these types of products in question beyond only qualitative definitions that are currently the benchmark. The framework of the model is chosen to reflect the nuances of attribute search online and suitable assumptions are made by internalising definitions of the two product groups, further vetted by the data used in this study. The empirical results reject Hypothesis 1 which claims that both types of goods have equivalent search intensities in the

internet markets due to ease of information search, as we find that experience goods are searched for at least three times more as compared search goods. The theoretical and the empirical result combined contributes to the primary finding of this study. The paper starts with the popular claim from literature, that search good shoppers search more in brick-and-mortar stores than experience good shoppers and the counter-claims, that with the advent of e-commerce, both types of shoppers will search equally, as online search allows "experience without ownership", thereby bridging the gap between these two product groups. The theoretical model shows formally that the latter claim does not hold true, which is further validated by the empirical result later in the paper.

Combining the theoretical and empirical result, this study provides a formal criterion that distinguishes between search and experience goods, besides the amount of product knowledge of consumers prior to their arrival at the store, as postulated by Nelson. Furthermore, the extent of search depends highly on how informative search is, in the sense that it enables potential buyers identify their optimal match. Predicting the likelihood of a purchase at every click or pageview remains one of the more challenging problems in literature, however by looking at purchase likelihood in retrospect we are able to identify the variance of search or it's informativeness. It is interesting to observe that higher the variance of search, lower is the intensity across both product groups. Given that experience good shoppers systematically exhibit high search variance, their search paths are three times longer than search good shoppers.

Online recommendation agents have opposing effects on search for the two products groups, impacting search goods significantly more that experience goods

Personalised recommendations are shown to affect online search and purchase decisions significantly as it takes into account user intent based on their search histories. The supply side value of deploying recommendation agents on an online shop has been studied in length as it creates massive cross-sell and up-sell opportunities for firms. On the demand side, however, there are fewer studies pertaining to this subject. The consensus is that online recommendations reduce search costs ,as shoppers can find information without actual search. This study contrarily points out that, how recommendations affect search intensities largely

depends on shoppers being able to determine match quality for every product they sample. While, a recommendation reduces search for an experience good, it increases that for a search good. It is observed that online recommendations play a major role in consumers' purchase decisions. Following recommendations do not seem to affect the extent of search for search goods, however, in the occasion that a search good shopper does follow a recommendation, its quality matters: closer the match of the recommended good to her true preference, more she searches. This makes intuitive sense: having the opportunity to improve match quality, a typical shopper will search longer as he gets more information about the available assortment. On the other hand, recommendations agents reduce search time for experience good shoppers.

Novel methodology has been used to quantify prior consumer knowledge in the internet markets, which positively affects only search goods

In understanding online shopping behaviour, a measure has been developed in this study for prior product knowledge. The underlying assumption is that, browsing histories on a particular store sufficiently influences current purchase decisions, hence a metric is developed indicating the varying levels of product knowledge consumers begin their search with, in the session whose purchases are being considered. Vector space models such as Tf-idf is used to measure the relevancy of each good sampled to the good bought, which allows us to calculate prior information directly from search. Partitioning the information sets accurately to discriminate current and prior knowledge, enables successful mapping of search intensity to product knowledge. This leads to one of the key findings that prior knowledge only impacts search goods in increasing the extent of search by close to 12 percent, while having no impact on experience goods.

The results may be generalised only to an extent primarily due to two *limiting factors*. Firstly, the data set does not contain any demographic or behavioural information about the shoppers. It is obvious that consumer search and purchase behaviour should vary across factors such as age, sex, location, psychological factors etc. The focus of this study is extensively on product characteristics as most of the explanatory variables are constructed based on those. Inclusion of consumer characteristics in the analysis would improve our understanding of the behaviour

of online shoppers. Secondly, the metric for prior information may not be a true measure of the prior knowledge. Prior information in this paper, is calculated from the point an unique agent visits this particular store in question, for the first time. It may be restrictive to assume that if the agents have bought a good from this store, they obtain information only from this particular store. This assumption may be violated in reality as there may be other sources of information, including the physical counterparts of the online store. However, the goal of this paper is primarily to demonstrate the role that information search plays in purchase decisions online, and that these decisions are not simply based on the price distributions. The idea is to focus on the direction of the effects as well as the magnitude in order to get an idea of how the discussed factors determine the extent of search based on the type of good.

Acknowledgements

This work was funded by a personal grant to the author from the Yrjo Jahnsson Foundation.

References

[1] Daniel A. Ackerberg. Advertising, learning, and consumer choice in experience good markets: an empirical examination*. *International Economic Review*, 44(3):1007–1040, 2003.

[2] Praveen Aggarwal and Rajiv Vaidyanathan. Perceived effectiveness of recommendation agent routines: search vs. experience goods. *International Journal of Internet Marketing and Advertising*, 2(1/2), 2005.

[3] Joseph Alba, John Lynch, Barlton Weitz, Chris Janiszewski, Richard Lutz, Alan Sawyer, and Stacy Wood. Interactive home shopping: Consumer, retailer, and manufacturer incentives to participate in electronic marketplaces. *Journal of Marketing*, 61(3):38–53, 1997.

[4] Mark Armstrong, John Vickers, and Jidong Zhou. Prominence and consumer search. *The RAND Journal of Economics*, 40(2):209–233, 2009.

[5] Fernando Branco, Monic Sun, and J. Miguel Villas-Boas. Optimal search for product information. *Management Science*, 58(11):2037–2056, 2012.

[6] Randolph E. Bucklin and Catarina Sismeiro. A model of web site browsing behaviour estimated on clickstream data. *Journal of Marketing Research*, 40(3):249–267, 2003.

[7] Kennet Burdett and Kenneth L. Judd. Equilibrium price dispersion. *Econometrica*, 51(4):955–969, 1983.

[8] Patrali Chatterjee, Donna L. Hoffman, and Thomas P. Novak. Modeling the clickstream: Implications for web-based advertising efforts. *Marketing Science*, 22(4):520–541, 2003.

[9] Pei-Yu Chen, Shin-yi Wu, and Jungsun Yoon. The impact of online recommendations and consumer feedback on sales. *ICIS 2004 Proceedings*, (58):520–541, 2004.

[10] Babur De Los Santos, Ali Hortaçsu, and Matthijs R. Wildenbeest. Testing models of consumer search using data on web browsing and purchasing behavior. *American Economic Review*, 102(6):2955–80, 2012.

[11] Peter A Diamond. A model of price adjustment. *Journal of Economic Theory*, 3(2):156 – 168, 1971.

[12] Avinash K Dixit and Robert S Pindyck. Investment under uncertainty. *(Princeton University Press, Princeton, NJ*, 1993.

[13] Avinash K. Dixit and Robert S. Pindyck. *Investment Under Inverstment*. Princeton University Press, Princeton, New Jersey, 1994.

[14] Tulin Erdem and Michael Keane. Decision-making under uncertainty: Capturing dynamic brand choice processes in turbulent consumer goods markets. *Marketing Science*, 15(1):1 – 20, 1996.

[15] Peng Huang, Nicholas H. Lurie, and Sabyasachi Mitra. Searching for experience on the web: An empirical examination of consumer behavior for search and experience goods. *Journal of Marketing*, 73:55–69, 2009.

[16] Jun B. Kim, Paulo Albuquerque, and Bart J. Bronnenberg. Online demand under limited consumer search. *Marketing Science*, 29(6):1001–1023, 2010.

[17] David N Laband. Advertising as Information: An Empirical Note. *The Review of Economics and Statistics*, 68(3):517–21, August 1986.

[18] David N Laband. An objective measure of search versus experience goods. *Economic Inquiry*, 29(3):497–509, 1991.

[19] Arthur S. Leahy. Search and experience goods : evidence from the 1960's and 70's. *The Journal of Applied Business Research*, 21(1):45–51, 2005.

[20] Klein Lisa. Evaluating the potential of interactive media through a new lens: Search versus experience goods. *Journal of Business Research*, 41(3):195–203, 1998.

[21] John G. Lynch Jr. and Dan Ariely. Wine online: Search costs affect competition on price, quality, and distribution. *Marketing Science*, 19(1):83–103, 2000.

[22] John J. McCall. Economics of information and job search. *The Quarterly Journal of Economics*, 84(1):113–126, 1970.

[23] Makoto Nakayama, Norma Sutcliffe, and Yun Wan. Has the web transformed experience goods into search goods. *Electronic Markets*, 20(3):251–262, 2010.

[24] Philip Nelson. Advertising as information. *Journal of Political Economy*, 82(4):729–54, 1974.

[25] Phillp Nelson. Information and consumer behaviour. *Journal of Political Economy*, 78(2):311–329, 1970.

[26] II Stahl, Dale O. Oligopolistic Pricing with Sequential Consumer Search. *American Economic Review*, 79(4):700–712, 1989.

[27] George Stigler. The economics of information. *Journal of Political Economy*, 69, 1961.

[28] Martin Weitzman. Optimal search for the best alternative. *Econometrica*, May, 1979.

[29] Asher Wolinsky. True monopolistic competition as a result of imperfect information. *The Quarterly Journal of Economics*, 101(3):493–511, 1986.

[30] Bo Xiao and Izak Benbasat. E-commerce product recommendation agents: Use, characteristics, and impact. *MIS Q.*, 31(1):137–209, March 2007.

7 Appendix

Figure 7: Posterior information across different products over the sample period

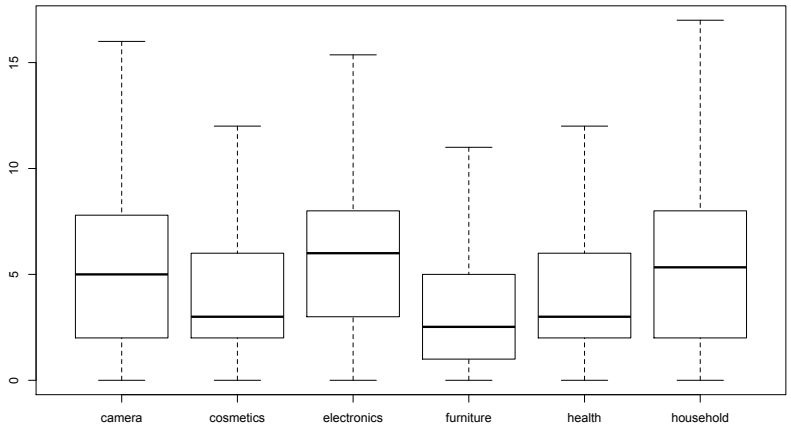
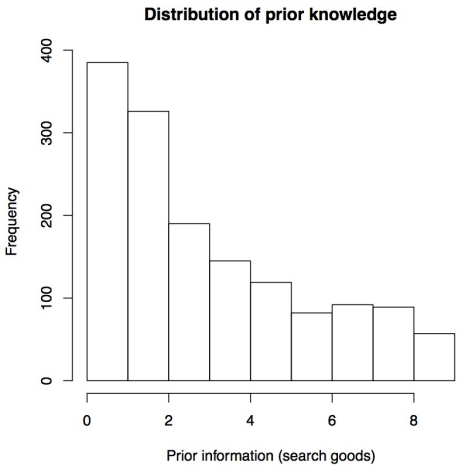


Figure 8: Distribution of prior knowledge



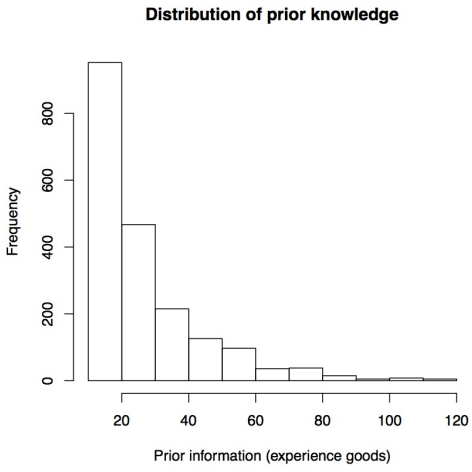


Table 5: Estimates of the number of searches prior to purchase for product categories: OLS

	Dependent variable: Number of searches					
	(Camera)	(Cosmetics)	(Furniture)	(Healthcare)	(Household appliances)	(Personal electronics)
Price	−0.021	−0.021***	0.053*	−0.015**	0.005***	0.031***
	(0.004)	(0.001)	(0.001)	(0.004)	(0.001)	(0.003)
Prior information	0.124**	−0.207	0.184***	−0.162	−0.096*	0.213***
	(0.017)	(0.013)	(0.012)	(0.015)	(0.009)	(0.014)
Variance	−1.162***	−4.334***	0.245***	−3.038***	−2.006***	−2.727***
	(0.149)	(0.903)	(1.633)	(0.849)	(0.337)	(0.410)
Recommendation quality	−0.134	−0.168***	0.024	−0.026**	−0.076	−0.014
	(0.045)	(0.020)	(0.011)	(0.016)	(0.012)	(0.011)
Recommendation (1=used:;0=not used)	0.274***	−0.342***	0.147***	− 0.341***	0.238***	0.230***
	(0.016)	(0.084)	(0.049)	(0.059)	(0.027)	(0.042)
Constant	2.571***	3.122***	2.425***	3.079***	3.174***	2.640***
	(0.134)	(0.089)	(0.050)	(0.060)	(0.026)	(0.040)
Observations	97	278	452	1,795	866	521
Adjusted R^2	0.341	0.356	0.315	0.306	0.384	0.312

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