Ph. D. Dissertation in Economics

Modeling Consumers' Information Search with Learning in the Pre-purchase Stage

학습을 고려한 소비자의 구매 이전 정보탐색 모형 개발

February 2016

Graduate School of Seoul National University

Technology Management, Economics, and Policy Program

Sesil Lim

Modeling Consumers' Information Search with Learning in the Pre-purchase Stage

지도교수 이 종 수

이 논문을 경제학박사학위 논문으로 제출함

2016 년 2 월

서울대학교 대학원

협동과정 기술경영경제정책 전공

임 세 실

임세실의 경제학박사학위 논문을 인준함

2015 년 12 월

위 원 장 _____김 연 배_____(인)

부위원장 _____이 종 수_____(인)

위　　원 _____김 용 대_____(인)

위　　원 _____조 영 상_____(인)

위　　원 _____박 유 리_____(인)

Abstract

Modeling Consumers' Information Search with Learning in the Pre-purchase Stage

Sesil Lim
Technology Management, Economics and Policy Program
The Graduate School
Seoul National University

With the growing dependence on the Internet to search for information, consumers search for online reviews of products and update their utility repeatedly until their product uncertainties are sufficiently reduced. However, previous literature considers consumers' behavioral process within a limited context and fails to describe it explicitly through structural estimation. This study aims to fill this gap by proposing a structural model that could explain consumers' sequential information search behavior including learning from the acquired information. A dynamic discrete choice model is proposed to describe consumers' information search behavior. In addition, the Bayesian learning mechanism is applied to consumers' process of updating their utility functions. The research framework proposed by this study enables the prediction of consumers' product preferences using data related to their information search behavior.

The structural estimation is conducted based on the Bayesian inference scheme. To reduce the computational burden of solving the dynamic programming, this study applies one of the modified Metropolis-Hastings algorithms for Markov Chain Monte Carlo (MCMC) sampling, called the IJC algorithm (Imai, Jain, & Ching, 2009). Lastly, using the data obtained from choice experiments, the empirical validity of the proposed model is examined. The empirical results of the proposed model provide more practical implications than the standard discrete choice model by providing the information of the initial impression of uncertainty in utility perceived by consumers, such as perception bias or scale of uncertainty.

Keywords: Information search; sequential search; consumer learning; dynamic discrete choice model; Bayesian estimation

Student Number: 2010-21082

Contents

List of Tables

List of Figures

Chapter 1. Introduction

1.1 Research Background

Consumers' information search behavior has been considered a critical, early stage precursor to consumers' decisions to buy products (Detlor, Sproule, & Gupta, 2003; Shim, Eastlick, Lotz, & Washington, 2001). Prior to making a purchase decision, consumers search for product information in order to reduce their uncertainty about the utility obtained by purchasing the product. This is particularly relevant for high-technology products, where understanding consumers' information search behavior is important to develop effective marketing strategies, as consumers typically face high levels of uncertainty due to complex and rapidly evolving technology (Mohr, Sengupta, & Slater, 2010).

The recent development of the Internet has changed consumers' information search behavior. Owing to its benefits, such as enabling access to a large amount of information at lower cost than traditional information sources, the dependence on the Internet to search for information has grown (Klein & Ford, 2003). According to several reports, over 80% of consumers search for product information via the Internet before they make a purchase (Holloway, 2012; Kaye, 2014; Morrison, 2014). Furthermore, the Internet allows consumers to learn about product quality, which is normally difficult to ascertain prior to purchase, from the experience of other consumers as well as from product attributes, which are relatively easy to observe including price, appearance, or

specification. In other words, as the Internet allows consumers to share their product usage experience without the constraints of time and space, consumers learn of the quality of the product prior to purchase by obtaining product reviews made by other consumers (Bei, Chen, & Widdows, 2004). In addition, consumers search for product information multiple times rather than just once, given the low cost of search on the Internet.

This behavioral change caused by the Internet has drawn interest from researchers and practitioners. Most researchers focus either on examining the differences in consumers' information search behavior across information sources, especially the Internet versus traditional information sources, or on analyzing the effect of consumer or product characteristics on online information search behavior (e.g., Bei et al., 2004; Chang, Lee, & Huang, 2012; Huang, Lurie, & Mitra, 2009; Klein, 1998; Klein & Ford, 2003; Lee, Kim, & Chan-Olmsted, 2011).

On the other hand, some recent studies have made an effort to specify explicitly consumers' behavioral process to search or process product information by constructing structural models (Chorus & Timmermans, 2008; De los Santos, Hortaçsu, & Wildenbeest, 2012; Erdem & Keane, 1996; Zhao, Yang, Narayan, & Zhao, 2013). According to Erdem and Keane (1996), the structural modeling approach allows researchers to understand consumer choice behavior more deeply and construct a more accurate, predictive model of consumer behavior. However, existing literature considers consumers' decision-making process in a bounded context. In addition, while consumers search online product reviews and update their utility repeatedly until their uncertainty is

sufficiently reduced, there have only been a few attempts to describe it theoretically without conducting structural estimations with empirical studies (Branco, Sun, & Villas-Boas, 2012; Lelis & Howes, 2008).

This study aims to fill this gap using the structural modeling approach. The model proposed in this study specifies an explicit structure to explain consumers' information search behavior including learning from acquired information. Furthermore, the proposed model is validated by empirical estimation. The proposal of such a model is the main contribution of this study.

1.2 Study Objective

This study has two main objectives. The first objective is to provide the structural model for consumers' information search with learning behaviors, which comprise the pre-purchase stage in the consumers' decision-making process. Since it would be desirable for the structural model to be based on the assumption of the utility-maximizing behavior of consumers and to be expressed in terms of the parameters of the utility function or its related constraints (Aguirregabiria & Mira, 2010; Erdem & Keane, 1996), the model proposed in this study should be consistent with utility maximization. To achieve this, the premise of this study is that consumers are uncertain about the utility obtainable from a product and search for product information in order to learn the true utility.

This uncertainty may be at a product attribute level or related to consumers' valuation

itself. Since some attributes are imperfectly observed, consumers are uncertain about their product valuations due to the existence of imperfect information (Erdem & Keane, 1996; Zhao et al., 2013). Hence, this study assumes that consumers are uncertain about which product matches their preferences due to attributes that are imperfectly observable by consumers. The part of utility affected by these attributes is defined as the *match value* in this study. In other words, the match value refers to the product's intrinsic value assumed by consumers based on how much the product matches up to their expectations or preferences. Since consumers are not able to know their exact match values prior to purchase, they search for online reviews to reduce the uncertainty associated with their match values.

Consumers' decision-making process raises the issue of how much information is needed to make a purchase decision. The economic perspective of consumers' information search, pioneered by Stigler (1961), assumes that consumers search for information when the marginal benefit of search exceeds its marginal cost. Literatures on information economics assume two approaches to consumer search behavior: simultaneous and sequential search (De los Santos et al., 2012; Kim, Albuquerque, & Bronnenberg, 2010). The simultaneous search strategy, also called fixed sample search strategy, assumes that consumers search for a fixed number of information of products at the same time, following Stigler's original suggestion. On the other hand, sequential search strategy assumes that consumers search for information of products one by one. Many researchers have supported the sequential search strategy as a better description of actual consumer search behavior (McCall, 1970; Nelson, 1970). In addition, recent

studies attempting to model consumer search behavior argue that consumers search for online information in a sequential manner (Häubl, Dellaert, & Donkers, 2010; Kim et al., 2010).

Accordingly, this study assumes that consumers search information of products sequentially. This assumption is adequate for the aim of this study, which is to provide a structural description of consumers' gradual learning from the online reviews they retrieve. The decision to stop searching is made in accordance with the economic perspective. In other words, consumers search for product information if the incremental utility, after updating their belief on match value from the additional information obtained, is expected to exceed the cost of searching for the information. This updating process is assumed to be conducted in a Bayesian manner. Since consumers' belief on a product's match value at the current time point depends on the past information processed, the proposed model is specified as a dynamic model.

The second objective of this study is to provide an estimation method for the proposed model and conduct an empirical study to validate the proposed model. The empirical investigation of dynamic discrete choice structural models has suffered from computational obstacles of estimation given that the formulation of the likelihood function or generalized method of moments (GMM) objective function requires dynamic programming solutions at every possible value of the associate variable. This essentially requires higher order integrals, and the complicated structure of the likelihood function makes it difficult to search for a global maximum (Aguirregabiria & Mira, 2010; Ching, Imai, Ishihara, & Jain, 2012; Erdem & Keane, 1996). In this respect, recent progress in

estimating dynamic discrete choice models based on the Bayesian Markov Chain Monte Carlo (MCMC) method allows us to reduce this computational burden by proposing methods to estimate parameters and solve the dynamic programming simultaneously (Imai, Jain, & Ching, 2009; Norets, 2009). Therefore, this study adopts the algorithm suggested by Imai et al. (2009) (IJC), one of the modified Metropolis-Hastings (MH) algorithms for MCMC sampling, to estimate the proposed model empirically.

1.3 Outline of the Study

This study consists of seven chapters. Chapter 2 discusses existing literature on consumers' behavior to search and process product information. Firstly, empirical studies analyzing factors affecting consumers' information search behavior are discussed. Then, structural modeling approaches to explain the consumer process of searching for information and reducing uncertainty based on the obtained information are reviewed. Lastly, limitations of previous studies are discussed, and the motivation and direction of this study is presented.

Chapter 3 proposes the dynamic discrete choice model for consumers' information search behavior from a theoretical perspective. The details of the model are described, including the specification of the utility function, the rule for deciding whether or not to search for additional product information, the manner of updating utility via a Bayesian learning mechanism, and the likelihood construction. Thereafter, important features of the proposed model are summarized with a discussion of the theoretical differences of the proposed model versus existing literature.

Chapter 4 documents the implementation of simulation studies to confirm the dynamics of consumer behavior in the decision making process based on the proposed model. Three Monte Carlo (MC) experiments are conducted for this purpose. The first MC experiment examines the change in consumers' information search behavior depending on the change in parameter values. To achieve this, a simulation is conducted for the scenario that a single consumer makes a decision for a single product. The second MC experiment examines the dynamics of searching between multiple products. For this purpose, this study simulates the decision of a single consumer across two products. The last MC experiment simulates the decision of multiple consumers with heterogeneous tastes in the match value. Decision-making processes of some representative consumers are tracked in this experiment.

Chapter 5 presents the method to estimate the proposed model, which is the Bayesian estimation method. Firstly, the model specification for structural estimation is introduced. Then, this chapter reviews Bayesian MCMC methods and their applications to dynamic discrete choice models, including the IJC algorithm that is adopted by this study. After presenting the estimation procedures, results of several Monte Carlo studies are presented, which are aimed at examining whether the proposed estimation method is able to recover true parameter values.

Chapter 6 shares the results of empirical studies that apply the proposed model. For these empirical studies, this study collects data on consumers' stated preferences by conducting choice experiments for several products. First, the design of the choice experiment is explained. Then, the data and empirical model are discussed. Lastly,

estimation results of consumers' behavior to search for information on the target product are presented with some discussion.

Finally, Chapter 7 summarizes the implications and contributions of this study. In addition, the limitations of the study are discussed along with possible topics for future research.

Chapter 2. Literature Review

The analysis of consumer search behavior has a long tradition in the research fields of Economics and Marketing. Pioneered by Stigler (1961), the research stream of consumer information search has developed remarkably from a theoretical perspective. However, few empirical studies have been conducted in this area due to model complexity and the lack of observable data (Ghose, Ipeirotis, & Li, 2012). The recent pervasiveness of the Internet allows consumers to access extensive information at little cost and provides researchers with a great opportunity to trace consumers' search behavior easily (Chorus & Timmermans, 2008). Therefore, the empirical investigation of consumers' information search behavior is of growing importance for both practitioners and researchers.

From this perspective, this chapter reviews the existing literature that attempts to verify consumers' search behavior empirically. First, studies that derive empirical findings without using structural models for consumers' behavior are reviewed. Then, two notable structural modeling approaches to consumers' information acquisition behavior, the consumer search framework and the consumer learning framework, are introduced in the second and third sub-sections. Lastly, the motivation of this study is outlined with a discussion on the limitations of existing research on this topic.

2.1 Empirical Study of Consumers' Information Search Behavior

For decades, most empirical researchers interested in consumers' decision-making

process have concentrated on deducing consumers' information search behavior from the data observed through surveys. According to Grant, Clarke, and Kyriazis (2007), existing literature may be classified into three groups. The first group focuses on analyzing the characteristics of information sources that affect consumers' search behavior (e.g., Bickart & Schindler, 2001; Daurer, Molitor, Spann, & Manchanda, 2015; Goh, Chu, & Wu, 2015; Klein & Ford, 2003). The second group of researchers analyze the consumer characteristics that lead consumers to make different choices and exhibit different information search behavior (e.g., Awasthy, Banerjee, & Banerjee, 2012; Levin, Huneke, & Jasper, 2000; Richard, Chebat, Yang, & Putrevu, 2010). The last group of researchers attempts to explain the difference in consumers' information search behavior in terms of product characteristics (e.g., Bei et al., 2004; Bhatnagar & Ghose, 2004; Detlor et al., 2003; Huang et al., 2009; Su, 2008; Wan, Nakayama, & Sutcliffe, 2012).

More specifically, researchers belonging to the first group are interested in examining the difference in consumers' information search behavior across specific information sources. For example, Klein and Ford (2003) investigated the difference in consumers' search behavior between online and offline information sources. Using the survey data from automobile shoppers and consumers, they confirmed that the earlier findings of information economics regarding the factors affecting consumers' total search time and total number of searched information sources, such as income, Internet use, age, and objective expertise, still hold when consumers search for product information via the Internet. Furthermore, based on evidence from the survey data, they argued that the traditional search behavior is substituted by Internet-based search, which is generally the

case nowadays. Bickart and Schindler (2001) compared the advertising effects of consumer-oriented online discussions including product reviews and marketer-generated web content, in terms of consumer perceived credibility and relevance of information based on survey data. They argued that consumer-generated information has a more positive effect on consumers' interest in products than marketer-generated information. On the other hand, some recent studies empirically analyzed consumers' search behavior on mobile Internet. Several studies suggested that consumers' dependency on their location is the most distinctive factor related to mobile information search (Daurer, Molter, Span, 2012; Daurer et al., 2015; Ghose, Goldfarb, & Han, 2013; Goh et al., 2015; Liu, Rau, & Gao, 2010). Most of these studies used location-based data to validate the difference in consumers' information search behavior with respect to their location.

The second group of researchers focused on analyzing the effects of consumers' personal attributes on their information search behavior, such as cognitive factors, personal skill factors, demographics, personality, and personal involvement. For example, Awasthy et al. (2012) examined the effect of consumers' prior product knowledge on their decision-making process using consumers' actual purchase data and surveyed prior knowledge data. Levin et al. (2000) described how consumers' need for cognition has an influence on their information search behavior by conducting choice experiments. Richard et al. (2010) examined the relationship between consumers' online search behavior and the Internet experience, website atmospherics, website involvement, and gender. Consumers' actual search data was collected from the website and data related to consumers' personal factors was gathered from a survey. From the analysis of

simultaneous equation models, they confirmed the differences in the web information search behavior between genders.

The last group of existing studies explains that consumer behavior to search for product information depends greatly on the characteristics of the product itself. Based on the framework suggested by Nelson (1970), products may be grouped into three categories based on the level of difficulty for consumers to evaluate their product quality: search goods, experience goods, and credence goods. Search goods are defined as products that may be evaluated easily and confidently in terms of quality by consumers prior to purchase. On the other hand, experience goods require the product to be purchased and experienced in order to evaluate product quality while credence goods are those whose quality cannot be evaluated precisely even a long time after their purchase. With the development of the Internet, the main interest of researchers in this field has shifted towards the analysis of the effect of the Internet on consumers' information search behavior by each product category. For instance, since the Internet enables consumers to observe other consumers' evaluation of product quality by obtaining product review information, products that were previously classified as experience or credence goods in the context of searching from traditional information sources may not be similarly categorized in the context of searching from the Internet. Bei et al. (2004) examined the effect of the product type, such as search goods or experience goods, on consumers' behavior to search. Based on the results of a web-based survey, they found that consumers search for more online information about experience goods than about search goods. Bhatnagar and Ghose (2004) suggested that consumers' pattern of searching for

information would differ by the type of products. They argued that when consumers search information about search goods, they reduce their search efforts in terms of time spent or frequency, as they get accustomed to more efficient search techniques over time. However, this tendency was not observed in the case of searching for information on experience goods.

As mentioned above, existing literature on empirical studies of consumer's information search behavior provides a glimpse of how and why consumers' information search behavior may differ. However, to explain consumers' behavior without constructing a structural model limits our ability to understand consumers' behavior deeply and predict it accurately. In addition, since information search is one of the important steps in consumers' decision-making process, the structural modeling approach should be adopted to consider the impact of searching for information on consumers' purchase decisions in all respects.

2.2 Structural Model for Consumer Search Behavior

Studies adopting the structural modeling approach to explain consumers' information search behavior are based on the theoretical background of information economics, developed from the original proposal by Stigler (1961) (e.g., De los Santos et al., 2012; Häubl et al., 2010; Kim et al., 2010; Koulayev, 2013). The premises of this approach may be summarized into two points: one is that information search comes at a cost, and the other is that consumers' search for information is based on their rationalization of costs and benefits. In other word, consumers search for information when the benefit of such

search exceeds the search cost.

As mentioned in Section 1.2, researchers in this field have assumed two different consumer search strategies: simultaneous and sequential search (Cheng, 2013; De los Santos et al., 2012; Kim et al., 2010). The simultaneous search, originally suggested by Stigler (1961), assumes that consumers search for a fixed number of information of products at the same time. Consumers' benefit from searching for fixed *n* number of information is the expected maximum gain of the sampling. Considering that a consumer searches for the best alternative in terms of utility, the benefit of the consumer from sampling alternatives in subset S ($j \in S$) is defined as follows:

$$B_i(S) = E\left[\max_{j \in S}\left\{U_{ij}\right\}\right] \cdots\cdots \text{Eq. (1)}$$

Specifying the number of elements in the subset *S* and the cost of search as n_s and *c* respectively, the net expected gain from searching alternatives in subset *S* at the same time can be described as follows:

$$R_i(S) = B_i(S) - n_s c \cdots\cdots \text{Eq. (2)}$$

Therefore, the optimal search strategy in simultaneous search is to find the subset that provides the greatest net expected gain to the consumer.

On the other hand, the sequential search strategy assumes that consumers search for product information sequentially. In this case, consumers keep searching for additional

information when the expected marginal benefit from search information is greater than the marginal search cost. Consider the basic case whereby consumers search for the product with the greatest utility. Let the current best alternative be product *k* and its utility be specified as U_k^*. Assuming that the distribution of utility of a non-searched product *j* is $f(U_j)$, the expected marginal benefit from searching for information on the product *j* is described as follows:

$$B_j\left(U_k^*\right) = \int_{U_k^*}^{\infty}\left(U_j - U_k^*\right) f\left(U_j\right) dU_j \quad \cdots\cdots \text{Eq. (3)}$$

Rational consumers include the products that are expected to give them positive, net expected gains in their consideration set of products to be searched. The consumer's optimal search strategy is to stop searching if the consideration set is empty, or else to continue searching for the product that has the maximum net expected gain among the products in the consideration set (Kim et al., 2010; Weitzman, 1979). The reservation utility allows researchers to describe this strategy very simply. By definition, reservation utility refers to the value of utility that equates the expected marginal benefit with the marginal cost, as follows:

$$z_j = B_j^{-1}(c) \quad \cdots\cdots \text{Eq. (4)}$$

where $c = B_j\left(z_j\right) = \int_{z_j}^{\infty}\left(u_j - z_j\right) f\left(u_j\right) du_j \quad \cdots\cdots \text{Eq. (5)}$

Therefore, the optimal search strategy in sequential search is whether to stop search if

$\max_{j \in S}\{z_j\} > U_k^*$ (where S is the set of non-searched products) or to find the product $l = \arg\max_{j \in S}\{z_j\}$.

Although there have been arguments for which search strategy is more appropriate to describe consumers' actual search behavior, the sequential search strategy is claimed as a superior strategy by many studies (Cheng, 2013; McCall, 1970; Nelson, 1970; Wiegmann, Seubert, & Wade, 2010). Recent empirical studies attempt to analyze consumers' information search process based on the sequential search strategy. Several notable studies are selected for review in this section. Firstly, Kim et al. (2010) analyzed consumers' online search and purchase behavior of durable goods by using aggregate level, view rank data of products collected from Amazon.com. They assumed that a consumer searches for the best product sequentially and a product once searched by the consumer is never re-searched. Although this assumption makes it impossible to explain consumers' information search behavior at the micro level, i.e., at the individual level or in terms of a single search behavior, their research contributes by predicting consumer preferences based on an observation of their search behavior by integrating the information search framework and a random utility choice framework.

On the other hand, De los Santos et al. (2012) utilized disaggregate level of data to examine consumers' search strategy in the context of searching prices of books sold by online bookstores. After estimating two structural models based on simultaneous and sequential search strategies respectively, the authors concluded that the simultaneous search strategy is more appropriate to explain consumers' behavior to search for pricing information.

While De los Santos et al. (2012) and Kim et al. (2010) specified the structural model for consumers' information search and purchase decisions by using the same utility function for all alternatives based on the random utility framework, Häubl et al. (2010) defined the utility for sequential information search and purchase decision differently. They assumed that consumers make purchase decisions based on the behavioral utility, using the difference of value centered on a reference point suggested by behavioral economics, and decisions to search based on the reservation utility. In addition, the authors used the data collected from choice experiments for empirical study.

Koulayev (2013) conducts the most notable research in this area. The author considered that consumers have no idea how the price of the product proposed by the seller is distributed and so they sequentially search for the seller who gives the best price for the product. In addition, consumers are assumed to learn the distribution of prices by searching for information in this framework. The author examined the validity of the model using aggregate level data. Although the author combined the consumers' information search and learning in a single research framework, this study has a similar limitation to the research of Kim et al. (2010) that once searched for by consumers the information of a seller is never re-searched. Furthermore, consumer learning is only related to the distribution of the unknown value, and not the unknown value itself.

2.3 Structural Model for Consumer Learning Behavior

The consumer learning models, pioneered by Erdem and Keane (1996), attempt to explain the effect of consumers' learning on their utility (e.g., Erdem & Keane, 1996;

Narayanan & Manchanda, 2009; Shin, Misra, & Horsky, 2012; Zhao et al., 2013). The premises of this approach are summarized as follows: consumers are uncertain about product attributes, including product quality, even after purchase, so they make decisions based on the perceived value of these attributes; consumers update their belief on these uncertain attributes steadily based on their experience through repeat purchases and advertising information over time (Ching, Erdem, & Keane, 2013). Therefore, the subject matter of consumer learning models has been generally restricted to non-durable goods that consumers buy repeatedly.

It is noteworthy to review the general framework proposed by Erdem and Keane (1996). They assumed that consumers are uncertain about the quality of product *j*, represented by Q_j. This assumption is the most distinctive feature of consumer learning models based on the conventional discrete choice models (Ching et al., 2013). In the conventional framework of discrete choice models, consumers have perfect information, implying that they know the exact value of utility without any uncertainty. However, in the actual choice situation, consumers have to make decisions with imperfect information. This uncertainty may be reduced by learning unobserved attributes from the informational signals. The authors assumed that informational signals received from both the purchasing experience and advertising are noisy measurements of product quality. Given this assumption, the informational signals received from purchasing the product *j* at period *t*, Q_{jt}^{E}, and those received from advertisements about product *j* at period *t*, A_{jt}, are defined as follows, respectively:

Purchasing experience: $Q_{jt}^{E} \sim N\left(Q_j, \sigma_E^2\right)$ for $t = 1, \cdots, T$ ····· Eq. (6)

Advertising: $A_{jt} \sim N\left(Q_j, \sigma_A^2\right)$ for $t = 1, \cdots, T$ ····· Eq. (7)

Suppose that the initial prior belief of unobserved quality is defined as $Q_j \mid \mathbf{I_0} \sim N\left(Q_{j0}, \sigma_{j0}^2\right)$, where $\mathbf{I_0}$ is the information set that contains Q_{j0} and σ_{j0}^2 for all products $j = 1, \cdots, J$. Hence, after obtaining the informational signals during *t* periods, the updated belief on the quality, $Q_j \mid \mathbf{I_t} \sim N\left(Q_{jt}, \sigma_{jt}^2\right)$, is specified as follows based on the Bayesian updating formula:

$$Q_{jt} = \frac{\left(1/\sigma_E^2\right)}{S_j(t)} \sum_{s=1}^{t} Q_{js}^{E} d_{js}^{E} + \frac{\left(1/\sigma_A^2\right)}{S_j(t)} \sum_{s=1}^{t} A_{js} d_{js}^{A} + \frac{\left(1/\sigma_{j0}^2\right)}{S_j(t)} Q_{j0}, \quad \text{Eq. (8)}$$

$$\frac{1}{S_j(t)} = \sigma_{jt}^2 = \frac{1}{\left(1/\sigma_{j0}^2\right) + N_j^E(t)\left(1/\sigma_E^2\right) + N_j^A(t)\left(1/\sigma_A^2\right)} \quad \text{Eq. (9)}$$

where d_{js}^{E} and d_{js}^{A} are indicators that consumers obtain the informational signal about product *j* at period *t* from the product experience and advertising, respectively. In addition, $N_j^E(t)$ and $N_j^A(t)$ are the number of informational signals about product *j* obtained by the consumer during *t* periods from the product experience and advertising, respectively. By definition, $N_j^E(t) = \sum_{s=1}^{t} d_{js}^{E}$ and $N_j^A(t) = \sum_{s=1}^{t} d_{js}^{A}$.

Based on this framework, consumers' belief on product quality depends on the past path of purchasing and advertisement exposure. This path dependency of the learning

model requires solving dynamic programming. Assuming that consumers are forward looking, the total value for consumers choosing the product *j*, $V_j\left(t \mid \mathbf{I_{t\text{-}1}}\right)$ is the sum of the current expected utility of the product, $E\left[U_{jt} \mid \mathbf{I_{t\text{-}1}}\right]$, and the expected present value of future payoffs, $\beta EV_j\left(\mathbf{I_t} \mid \mathbf{I_{t\text{-}1}}\right)$. Therefore, at period *t*, consumers' would decide to purchase the product that has the maximum total value, i.e., $\max_j V_j\left(t \mid \mathbf{I_{t\text{-}1}}\right)$.

Existing researches based on consumer learning framework utilized scanner panel data, which contains observation for purchasing products by consumers collected by retailers, for empirical study (Erdem & Keane, 1996; Shin et al., 2012). On the other hand, Narayanan and Manchanda (2009) analyzed the learning of physicians about new drugs using prescription observations and detailing calls of a panel of physicians. Furthermore, Zhao et al. (2013) examined the learning from online reviews of books, prior to purchase, using aggregate data. Although the work by Zhao et al. (2013) proposes the research framework to analyze the effect of online reviews on consumers' decision-making process including product learning, their work does not explain when consumers would stop searching for information.

2.4 Limitations of Previous Research and Research Motivation of the Study

As discussed in previous sections, while a structural model is required to understand the complex behavior of consumers in the pre-purchase stage and to predict their decision

accurately, most existing empirical literature does not adopt this approach due to the complexity of the model and lack of observed data. Specifically, although the basic models of consumer behavior describe the decision making process by consumers as the composition of the five stages (Figure 1) according to Mohr et al. (2010), previous empirical studies have focused on the information search stage by deducing factors affecting consumers' search behavior such as consumer characteristics, information sources, or products. Moreover, until recently, the use of the structural modeling approach to predict consumer behavior based on utility maximization has mainly been considered from a theoretical perspective.

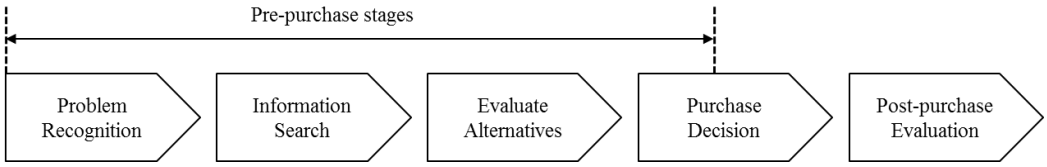

Figure 1. Conceptual model of the decision making process by consumers

Although the development of the Internet provides researchers with easy access to sufficient data, two main research streams of structural modeling approach for consumers' information search behavior have been developed, considering consumers' decision-making process in a bounded context. Literatures on structural modeling for consumer search assume perfect information about the acquiring information. In other words, assuming that consumers search for information about a product attribute, the data acquired by consumers contains the entire set of information about that attribute. Hence, the uncertainty regarding that attribute is removed. Based on this assumption, studies on

consumer search have not considered the case whereby consumers search for information on the product or product attribute repeatedly. In addition, assuming that the uncertainty is removed after searching for relevant information, consumers are assumed to know the exact, true value of the attribute from the acquired information. Hence, the remaining area of interest for researchers is whether consumers would stop or continue searching for information. In summary, existing studies on the structural modeling approach for consumer search have attempted to explain consumers' decision-making process from information search to purchase decision by considering that the evaluation of alternatives is concurrent with the information search, as depicted in Figure 2.

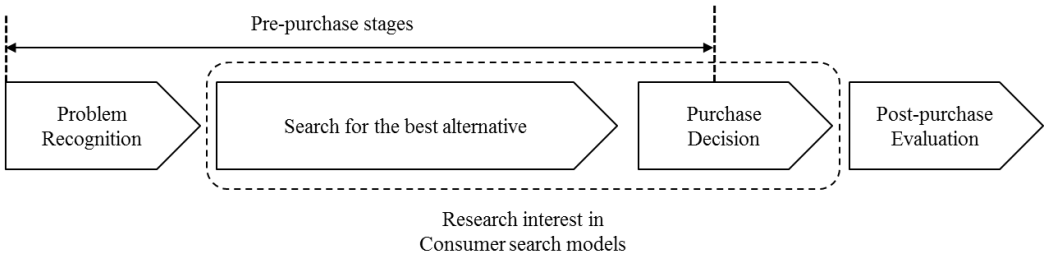

Figure 2. Assumed decision making process in consumer search models

On the other hand, consumer learning literature assumes that some uncertainty remains even after purchasing the product due to the existence of imperfectly observed attributes. Moreover, since the information obtained from commercial messages does not contain perfect information about the attribute, consumers are assumed to be uncertain about the product even after exposure to substantial advertising. As per this assumption, consumers keep modifying their utility after purchasing the product based on their usage

experience and commercial exposure for an infinite period, as described in Figure 3. Since the consumer learning model ignores the consumers' behavior to search for information actively, the stage of information search is usually not the focus of this modeling approach.

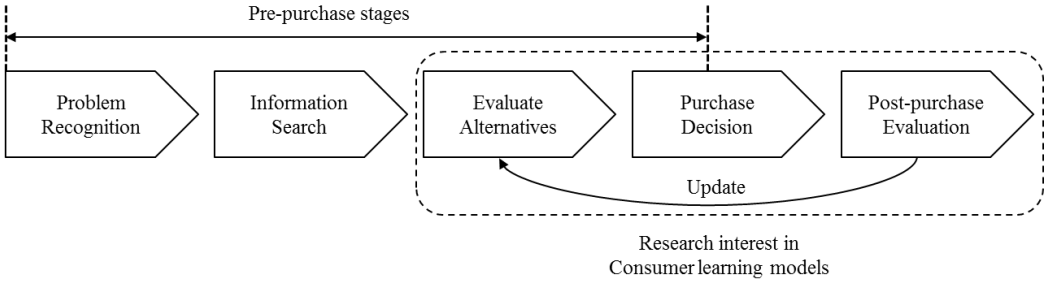

Figure 3. Assumed decision making process in consumer learning models

However, in reality, consumers make decisions despite the existence of uncertainty if the uncertainty is of a negligible value. Hence, consumers' information search behavior and updating of utility from acquired information occurs successively in the pre-purchase stage in consumers' actual decision-making process. Only a few, recent theoretical studies consider this point, although it has not been investigated empirically (Branco et al., 2012; Lelis & Howes, 2008). This study's objective is to describe consumers' decision-making process in accordance with their actual behavior as presented in Figure 4. Therefore, this study proposes a model which may be applied to empirical research by combining the consumer search and consumer learning frameworks.

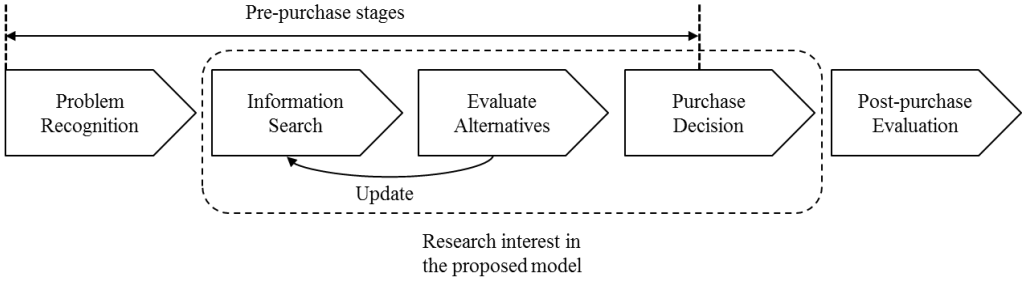

Figure 4. Assumed decision making process in the proposed model

Chapter 3. Model

In this chapter, the structural model for consumers' information search with learning behaviors is proposed from a theoretical perspective. The proposed model is derived from the framework of random utility theory and the assumption of the consumer's utility maximization. The following sections provide the detailed specifications of the proposed model. The model specifications from an econometric perspective are presented in Chapter 5.

This study considers a multinomial choice situation whereby consumers make a purchase decision among J products. As partially discussed in Section 1.2, the basic premises of this study are as follows:

(i) consumers are uncertain about the utility of products due to the existence of attributes that are imperfectly observed by both consumers and researchers, defined as the match value;

(ii) consumers search for product information and update their belief on match values sequentially;

(iii) at each time period, consumers decide whether to search for additional product information or to buy the most preferred product without any further search; and

(iv) consumer decisions to search for information are based on trade-offs between the costs and benefits of searching for additional information.

Accordingly, the conceptual framework of consumers' decision-making process described in this study is illustrated in Figure 5.

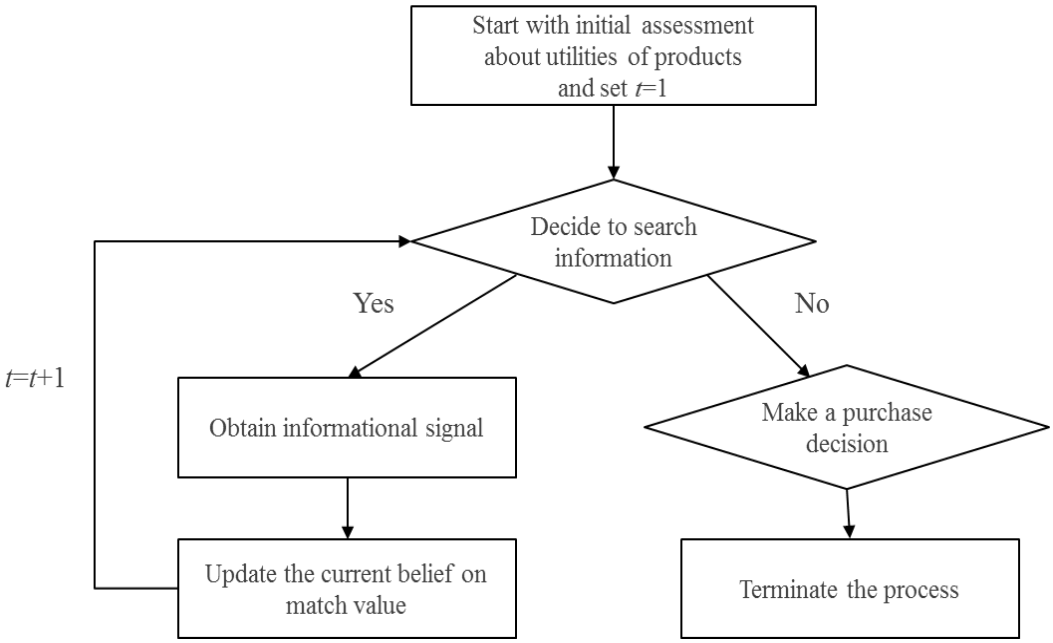

Figure 5. Illustration of consumers' decision-making process in the proposed model

3.1 Utility

Discrete choice models have been the normative framework to describe consumer purchase behavior. In discrete choice models, the utility that consumer *i* obtains from product *j* is defined as U_{ij}, for $j = 1, \cdots, J$. Based on the framework of random utility models, although consumers know their exact value of utility, researchers may observe consumers' utility maximizing behavior as well as some attributes of products (x_{ij}) and consumers (s_i), and presume the utility from these observations (Train, 2009). Based on

this concept, utility comprises two parts: one is the part that researchers can observe, denoted by V_{ij}, and the other is specified as a random variable ε_{ij} that represents the observation error.

$$U_{ij} = V_{ij}\left(x_{ij}, s_i\right) + \varepsilon_{ij} \quad \cdots\cdots \quad \text{Eq. (10)}$$

Based on the utility function described in Equation 10, the consumer purchases the product that gives him/her the greatest utility. Hence, the premise of the basic discrete choice model is that uncertainty exists only from the perspective of researchers, i.e., consumers are certain of their product valuations.

However, the uncertainty in the utility is caused not only by the regression error, but also by the bounded rationality of consumers and the incompleteness of information (Signorino, 2003). Bounded rationality indicates that consumers may misperceive their utilities or implement their action incorrectly due to either the limited information available or the cognitive limitations of the human brain (Simon, 1955). The uncertainty associated with incomplete information results from the fact that consumers are unable to precisely observe the payoff from the actions of other consumers, while the uncertainty captured by the regression error results from the fact that researchers are unable to precisely observe the payoff from the actions of all consumers.

The standard discrete choice models assume full rationality by consumers and complete information. Hence, they consider the regression error to be the only source of uncertainty. Since consumers are assumed to know the exact utilities before making purchase decisions and researchers are unable to perfectly observe them under any

circumstances, the purchasing decision problem is static. Suppose that the observed part of utility is determined as a linear function of k observed attributes. Assuming that $\mathbf{X_{ij}}$ is a k dimensional vector whose elements are the observable attributes of product j and $\mathbf{b}$ is a vector of corresponding coefficients, the utility obtained by consumer i from purchasing product j is specified as follows:

$$U_{ij} = \mathbf{X}'_{ij}\mathbf{b} + \varepsilon_{ij} \quad \cdots\cdots \text{Eq. (11)}$$

In Equation 11, the uncertainty is represented by the random component ε_{ij}. Although consumers know the exact value of ε_{ij}, researchers are unable to observe it and only know its distribution. If researchers specify the distribution of ε_{ij} as the type I extreme value distribution or joint normal distribution, the choice probabilities may be derived from the logit model or probit model, respectively.

On the other hand, this study allows the utility to include the uncertainty from bounded rationality and incomplete information. Given the existence of some imperfectly observed attributes, consumers are uncertain about their utility of products before purchase. The true value of the utility component depending on imperfectly observed attributes is defined as the match value, Q_{ij}. Consumers cannot know the true value of Q_{ij} before purchasing the product. For this reason, consumers perceive the utility based on their belief of the match value. Searching for product information allows consumers to revise their belief of the match value. As the consumers' beliefs of the match value at

search occasions obviously differ from each other, the perceived utility of consumer i for product j prior to searching for information at period t is specified as follows:

$$U_{ijt} = \mathbf{X}'_{ij}\mathbf{b} + \tilde{Q}_{ij,t-1} + \varepsilon_{ijt} \quad \cdots\cdots \quad \text{Eq. (12)}$$

Since $\tilde{Q}_{ij,t-1}$ is stochastic from the consumer's perspective, the perceived utility is also stochastic. Therefore, consumers make a decision using their expected utility with respected to their belief on the match value at the current time point, based on the set of information acquired from previous time points. Specifying consumer i 's set of information after he/she acquires additional information at time t as $\mathbf{I_{it}}$, the elements of $\mathbf{I_{it}}$ contain the information acquired from the initial period (that is, $\tau = 1$) to the current period ($\tau = t$) as well as consumers' perceived match values after acquiring information at period t. In the proposed model, consumers decide whether or not to search for additional information using the expected utility based on the set of information processed in the past, $\mathbf{I_{i,t-1}}$, and update their information set to $\mathbf{I_{it}}$ by acquiring the information only if they decide to delay the purchase decision and search for additional information. Hence, the expected utility prior to information search at time t is specified as follows:

$$U^E_{ijt} = E_{\tilde{Q}_{ijt}}\left[U_{ijt} \mid \mathbf{I_{i,t-1}}\right] = \mathbf{X}'_{ij}\mathbf{b} + E\left[\tilde{Q}_{ij,t-1} \mid \mathbf{I_{i,t-1}}\right] + \varepsilon_{ijt} \quad \cdots\cdots \quad \text{Eq. (13)}$$

Suppose that the consumer has an initial belief of the match value of the product

before acquiring any information, which is defined as $\tilde{Q}_{ij0}$. As specified in Equation 13, consumer i's expected utility of product j at the beginning of the decision making process may be described as follows:

$$U_{ij0}^{E} = \mathbf{X}_{ij}'\mathbf{b} + E\left[\tilde{Q}_{ij0} \mid \mathbf{I_{i0}}\right] + \varepsilon_{ij0} \quad \cdots\cdots \text{Eq. (14)}$$

At the beginning ($t = 1$), consumers know the distribution of their initial belief of the match values of products. Hence, the initial information set contains this information.

Before we discuss the process of updating the match values by consumers, one could consider the situation whereby consumers decide not to purchase any of the products even after searching for all product information. To capture this case, this study assumes the expected utility associated with the "no purchase" choice to be an error component normalized to zero, as shown in Equation 15.

$$U_{iot}^{E} = \varepsilon_{iot} \quad \cdots\cdots \text{Eq. (15)}$$

In this study, all regression error components of utility are assumed to be distributed independently and identically across alternatives and periods and based on the type I extreme value distribution.

3.2 Process of Updating Match Value Beliefs

This study assumes that consumers update their match value beliefs in a Bayesian

manner. This assumption is consistent with a recent stream of studies in the field of marketing and economics related to consumer learning models (e.g., Ackerberg, 2003; Erdem & Keane, 1996; Narayanan & Manchanda, 2009; Shin et al., 2012).

Consider that consumer i searches for information on product j at period t. The informational signal about product j's match value that consumer i acquires after searching for information on j is defined as S_{ijt}. Although consumers may not observe the true match values for product j, they may form an impression through the searched information. Accordingly, the informational signal may be interpreted as a noisy measurement of the match value, Q_{ij}. Hence, the informational signal about product j obtained by consumer i at period t may be specified as follows:

$$S_{ijt} \sim N\left(Q_{ij}, \sigma^2_{S_{ij}}\right) \quad \text{Eq. (16)}$$

According to Equation 16, the informational signal may be treated as a normal random variable having mean value of Q_{ij} and variance of $\sigma^2_{S_{ij}}$. A point to note is that while consumers know the exact value of the informational signal after acquiring it, researchers are unable to observe this.

Consider $t = 1$. Then, the informational signal about j, $S_{ij,t=1}$ is generated from the normal distribution, $N\left(Q_{ij}, \sigma^2_{S_{ij}}\right)$. Given that $\sigma^2_{S_{ij}}$ has a fixed value for every i and j, assuming that the prior belief on the match value, denoted as $\tilde{Q}_{ij0}$, is a normal random

variable whose mean and variance are μ_{ij0} and σ_{ij0}^2 respectively, the posterior belief on the match value $\tilde{Q}_{ij1}$ is specified as follows (DeGroot, 2004):

$$\tilde{Q}_{ij1} \sim N\left(\mu_{ij1}, \sigma_{ij1}^2\right) \quad \text{Eq. (17)}$$

where $\mu_{ij1} = \frac{\sigma_{ij1}^2}{\sigma_{ij0}^2}\mu_{ij0} + \frac{\sigma_{ij1}^2}{\sigma_{S_{ij}}^2} S_{ij,t=1}$ Eq.(18)

and $\frac{1}{\sigma_{ij1}^2} = \frac{1}{\sigma_{ij0}^2} + \frac{1}{\sigma_{S_{ij}}^2}$. Eq. (19)

It is noteworthy that the posterior belief about the match value of products other than product j is identical to their prior belief as no informational signal about these products has been obtained. In other words, $\tilde{Q}_{ik0} = \tilde{Q}_{ik1}$ for $\forall k \neq j$.

Now, consider the general case where $t \geq 1$. If consumer i searches for information of product j at period t, the informational signal about product j is denoted as S_{ijt} which is generated from the normal distribution $N\left(Q_{ij}, \sigma_{S_{ij}}^2\right)$ for a given $\sigma_{S_{ij}}^2$. At the current time point t, the prior belief on the match value of product j is $\tilde{Q}_{ij,t-1}$, which follows a normal distribution with mean $\mu_{ij,t-1}$ and variance $\sigma_{ij,t-1}^2$. Therefore, the posterior belief about the match value, $\tilde{Q}_{ijt}$, is specified as follows:

$$\tilde{Q}_{ijt} \sim N\left(\mu_{ijt}, \sigma_{ijt}^2\right) \quad \text{Eq. (20)}$$

where $\mu_{ijt} = \frac{\sigma_{ijt}^2}{\sigma_{ij,t-1}^2}\mu_{ij,t-1} + \frac{\sigma_{ijt}^2}{\sigma_{S_{ij}}^2} S_{ijt}$ Eq. (21)

and $\frac{1}{\sigma_{ijt}^2} = \frac{1}{\sigma_{ij,t-1}^2} + \frac{1}{\sigma_{S_{ij}}^2}$. ·····Eq. (22)

Equations 20 to 22 may be specified in the manner of successive substitutions. Denoting d_{it} as the variable that indicates that a product is selected by consumer i for information search at period t, then $d_{it} = j$ if the consumer i searches for information on product j at period t. Let $I(\cdot)$ be the indicator function whose value is 1 if the condition described in parenthesis is true or 0 otherwise. Then, $\sum_{\tau=1}^{t} I(d_{i\tau} = j)$ is identical to the number of informational signals acquired about product j up to the current period t. In addition, $\sum_{k=1}^{J}\sum_{\tau=1}^{t} I(d_{i\tau} = k) = t$ as this study assumes that consumers search for product information sequentially until they decide not to search for any further information. From this perspective, consumer i's posterior belief of the match value of product j at time t is explained as a combination of the prior belief $\tilde{Q}_{ij0} \sim N(\mu_{ij0}, \sigma_{ij0}^2)$ as well as $\sum_{\tau=1}^{t} I(d_{i\tau} = j)$ random samples from the normal distribution, $N(Q_{ij}, \sigma_{S_{ij}}^2)$. Therefore, the posterior belief of the match value $\tilde{Q}_{ijt}$ is derived as follows (DeGroot, 2004):

$\tilde{Q}_{ijt} \sim N(\mu_{ijt}, \sigma_{ijt}^2)$ ·····Eq. (23)

where $\mu_{ijt} = \frac{\sigma_{ijt}^2}{\sigma_{ij0}^2}\mu_{ij0} + \frac{\sigma_{ijt}^2}{\sigma_{S_{ij}}^2}\sum_{\tau=1}^{t}\left(I(d_{i\tau} = j)\cdot S_{ij\tau}\right)$ ·····Eq. (24)

and $$\frac{1}{\sigma_{ijt}^2} = \frac{1}{\sigma_{ij0}^2} + \frac{\sum_{\tau=1}^{t} I\left(d_{i\tau} = j\right)}{\sigma_{S_{ij}}^2} \quad \cdots\cdots \text{Eq. (25)}$$

Combining the above equations with the expected utility function described in Equation 13, the expected utility may be expressed as follows:

$$\begin{aligned} U_{ijt}^E &= \mathbf{X}_{ij}'\mathbf{b} + \mu_{ij,t-1} + \varepsilon_{ijt} \\ &= \mathbf{X}_{ij}'\mathbf{b} + \frac{\sigma_{S_{ij}}^2}{\sigma_{S_{ij}}^2 + \sigma_{ij0}^2 \sum_{\tau=1}^{t-1} I\left(d_{i\tau} = j\right)} \mu_{ij0} \\ &\quad + \frac{\sigma_{ij0}^2}{\sigma_{S_{ij}}^2 + \sigma_{ij0}^2 \sum_{\tau=1}^{t-1} I\left(d_{i\tau} = j\right)} \sum_{\tau=1}^{t-1} \left(I\left(d_{i\tau} = j\right) \cdot S_{ij\tau} \right) + \varepsilon_{ijt} \end{aligned} \quad \cdots\cdots \text{Eq. (26)}$$

3.3 Decision to Search for Information

At each time point, consumers make a decision whether or not to search for information by comparing the cost of searching with the benefit from searching, which may be expressed in terms of consumers' expected utilities. Within the framework of this study, the benefit from searching for additional information is modeled as the difference between the consumer's expected reward from making a purchase decision without further search in the current period and the anticipated reward in the next period after searching for additional information. The reward received by the consumer for making a purchase decision and terminating the information search process at current period t is

identical to the maximum expected utility obtained by the consumer under the assumption of utility maximization. This specification corresponds to the approach of previous studies (Ackerberg, 2003; Chorus & Timmermans, 2008).

Given the specification of the expected utility in Section 3.1, the reward received by consumer i at period t is described as follows:

$$R_{it} = \max\left[U_{iot}^{E}, U_{i1t}^{E}, \cdots, U_{iJt}^{E}\right] \cdots\cdots \text{Eq. (27)}$$

where $U_{ijt}^{E} = \mathbf{X}_{ij}'\mathbf{b} + \mu_{ij,t-1} + \varepsilon_{ijt}$, for $j = 1, \cdots, J$, $\cdots\cdots$ Eq. (28)

and $U_{iot}^{E} = \varepsilon_{iot}$. $\cdots\cdots$ Eq. (29)

Note that U_{ijt}^{E} is specified by Equation 26.

On the other hand, the expected reward obtained by consumer *i* when he/she searches for additional information at period *t* and makes a purchase decision at the next period is the maximum anticipated value of the expected utility of period *t+1*. Specifically, the expected utility of product j ($j = 1, \cdots, J$) obtained by a consumer i is specified as $U_{ij,t+1}^{E}$. As mentioned in Section 3.2, the posterior beliefs of the match values of products are derived conditional on the consumers' decision to search for product information, which is denoted by the variable d_{it}. As defined in Section 3.2, $d_{it} = j$ ($j = 1, \cdots, J$) indicates the case where consumer i searches for information on product j at period t. A point to highlight is that $d_{it} = 0$ indicates the case whereby the consumer decides to terminate the information search process.

Now, consider that a consumer *i* decides to search for information about product *k* at

period t ($d_{it} = k$). Firstly, we consider the reward at period $t+1$. Equations 30 and 31, respectively specify the expected utilities of products other than k ($\forall j \neq k$) and the "no purchase" option at period $t+1$:

$$\begin{aligned} U_{ij,t+1}^{E} &= \mathbf{X}_{ij}'\mathbf{b} + \mu_{ij,t-1} + \varepsilon_{ij,t+1} \\ &= \mathbf{X}_{ij}'\mathbf{b} + \frac{\sigma_{S_{ij}}^2}{\sigma_{S_{ij}}^2 + \sigma_{ij0}^2 \sum_{\tau=1}^{t-1} I\left(d_{i\tau} = j\right)} \mu_{ij0} \\ &+ \frac{\sigma_{S_{ij}}^2}{\sigma_{S_{ij}}^2 + \sigma_{ij0}^2 \sum_{\tau=1}^{t-1} I\left(d_{i\tau} = j\right)} \sum_{\tau=1}^{t-1} \left(I\left(d_{i\tau} = j\right) \cdot S_{ij\tau} \right) + \varepsilon_{ij,t+1} \end{aligned} \quad \cdots\cdots \text{Eq. (30)}$$

and $U_{io,t+1}^{E} = \varepsilon_{io,t+1}$ ····· Eq. (31)

From the consumers' perspective, all variables including $I\left(d_{i\tau} = j\right) \cdot S_{ij\tau}$, $\tau = 1, \cdots, t-1$, are deterministic at the time period $t+1$.

In addition, the expected utility of product k obtained by consumer i at period $t+1$ is described in Equation 32, as follows:

$$U_{ik,t+1}^{E} = \mathbf{X}_{ik}'\mathbf{b} + \mu_{ikt} + \varepsilon_{ik,t+1} \quad \cdots\cdots \text{Eq. (32)}$$

Substituting Equations 24 and 25, Equation 32 may be rewritten as follows:

$$
\begin{aligned}
U_{ik,t+1}^{E} &= \mathbf{X}_{ik}'\mathbf{b} + \frac{\sigma_{S_{ik}}^{2}}{\sigma_{S_{ik}}^{2} + \sigma_{ik0}^{2}\sum_{\tau=1}^{t} I\left(d_{i\tau}=k\right)}\mu_{ik0} \\
&\quad + \frac{\sigma_{ik0}^{2}}{\sigma_{S_{ik}}^{2} + \sigma_{ik0}^{2}\sum_{\tau=1}^{t} I\left(d_{i\tau}=k\right)}\sum_{\tau=1}^{t}\left(I\left(d_{i\tau}=k\right)\cdot S_{ik\tau}\right) + \varepsilon_{ik,t+1} \\
&= \mathbf{X}_{ik}'\mathbf{b} + \frac{\sigma_{S_{ik}}^{2}}{\sigma_{S_{ik}}^{2} + \sigma_{ik0}^{2}\cdot\left(\sum_{\tau=1}^{t-1} I\left(d_{i\tau}=k\right)+1\right)}\mu_{ik0} \\
&\quad + \frac{\sigma_{ik0}^{2}}{\sigma_{S_{ik}}^{2} + \sigma_{ik0}^{2}\cdot\left(\sum_{\tau=1}^{t-1} I\left(d_{i\tau}=k\right)+1\right)}\sum_{\tau=1}^{t-1}\left(I\left(d_{i\tau}=k\right)\cdot S_{ik\tau}\right) \\
&\quad + \frac{\sigma_{ik0}^{2}}{\sigma_{S_{ik}}^{2} + \sigma_{ik0}^{2}\cdot\left(\sum_{\tau=1}^{t-1} I\left(d_{i\tau}=k\right)+1\right)}S_{ikt} + \varepsilon_{ik,t+1}
\end{aligned}
\qquad \cdots\cdots\cdots\cdots\cdots\cdots \text{Eq. (33)}
$$

At period $t+1$, all variables in Equation 33 are deterministic. Therefore, the reward obtained by the consumer from terminating the information search and making a purchase decision is the maximized expected utility at period $t+1$.

However, at the current time point t, consumers are unable to evaluate the exact mean value of the posterior belief on the match value (μ_{ikt}) as they are yet to acquire any informational signals at this point. Hence, S_{ikt} is stochastic at period t having the information set $\mathbf{I_{i,t-1}}$. In addition, the regression errors of utility at period $t+1$ ($\mathbf{e_{t+1}} = \left\{\varepsilon_{i1,t+1},\cdots,\varepsilon_{iJ,t+1},\varepsilon_{io,t+1}\right\}$) are not realized values at point *t*. From these perspectives, the expected reward obtained by a consumer i at period $t+1$ conditional on $d_{it}=k$ is described as the expectation taken with respect to $\mathbf{e_{t+1}}$ and S_{ikt} based on the information set $\mathbf{I_{i,t-1}}$, as specified in Equation 34:

$$
\begin{aligned}
ER_{i,t+1|d_{it}=k} &= E_{S_{ikt},\mathbf{e_{t+1}}}\left[R_{i,t+1} \mid \mathbf{I_{i,t-1}}; d_{it}=k\right] \\
&= E_{S_{ikt},\mathbf{e_{t+1}}}\left[\max\left\{U^{E}_{io,t+1}, U^{E}_{i1,t+1}, \cdots, U^{E}_{iJ,t+1}\right\} \mid \mathbf{I_{i,t-1}}; d_{it}=k\right]
\end{aligned}
\quad \cdots\cdots\cdots\cdots\cdots \text{Eq. (34)}
$$

Since $\mathbf{e_{t+1}}$ and S_{ikt} are independent, Equation 34 may be rearranged as follows:

$$
\begin{aligned}
ER_{i,t+1|d_{it}=k} &= E_{S_{ikt}}\left[E_{\mathbf{e_{t+1}}}\left[\max\left\{U^{E}_{io,t+1}, U^{E}_{i1,t+1}, \cdots, U^{E}_{iJ,t+1}\right\}\right] \mid \mathbf{I_{i,t-1}}; d_{it}=k\right] \\
&= E_{S_{ikt}}\left[\ln\left(1+\sum_{j\neq k} e^{\mathbf{X}'_{ij}\mathbf{b}+\mu_{ij,t-1}} + e^{\mathbf{X}'_{ik}\mathbf{b}+\mu_{ikt}}\right) \mid \mathbf{I_{i,t-1}}; d_{it}=k\right] \\
&= E_{S_{ikt}}\left[\ln\left(\begin{array}{l}
1+\sum_{j\neq k}\exp\left(\begin{array}{l}
\mathbf{X}'_{ij}\mathbf{b}+\dfrac{\sigma^2_{S_{ij}}\cdot\mu_{ij0}}{\sigma^2_{S_{ij}}+\sigma^2_{ij0}\sum_{\tau=1}^{t-1} I\left(d_{i\tau}=j\right)} \\
+\dfrac{\sigma^2_{ij0}\cdot\sum_{\tau=1}^{t-1}\left(I\left(d_{i\tau}=j\right)\cdot S_{ij\tau}\right)}{\sigma^2_{S_{ij}}+\sigma^2_{ij0}\sum_{\tau=1}^{t-1} I\left(d_{i\tau}=j\right)}
\end{array}\right) \\
+\exp\left(\begin{array}{l}
\mathbf{X}'_{ik}\mathbf{b}+\dfrac{\sigma^2_{S_{ik}}\cdot\mu_{ik0}}{\sigma^2_{S_{ik}}+\sigma^2_{ik0}\left(\sum_{\tau=1}^{t-1} I\left(d_{i\tau}=k\right)+1\right)} \\
+\dfrac{\sigma^2_{ik0}\cdot\sum_{\tau=1}^{t-1}\left(I\left(d_{i\tau}=k\right)\cdot S_{ik\tau}\right)}{\sigma^2_{S_{ik}}+\sigma^2_{ik0}\left(\sum_{\tau=1}^{t-1} I\left(d_{i\tau}=k\right)+1\right)} \\
+\dfrac{\sigma^2_{ik0}\cdot S_{ikt}}{\sigma^2_{S_{ik}}+\sigma^2_{ik0}\left(\sum_{\tau=1}^{t-1} I\left(d_{i\tau}=k\right)+1\right)}
\end{array}\right)
\end{array}\right) \mid \mathbf{I_{i,t-1}}; d_{it}=k\right]
\end{aligned}
$$

$$\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots \text{Eq. (35)}$$

Since consumers behave in a manner that maximizes their utility, they choose to search for information on the product that allows them to obtain the maximum expected

reward after acquiring the information. Based on this insight, the benefit of information search is expressed as $\max_k ER_{i,t+1|d_{it}=k} - R_{it}$.

Finally, the marginal cost of search is defined as the consumer specific variable c_i in this study. The cost of search comprises either monetary or non-monetary value. Literature on the economics of search describes cost as the time spent on searching or the effort to ascertain and process information (Chorus & Timmermans, 2008; De los Santos et al., 2012; Kim et al., 2010; Ratchfold, 1982). Therefore, the cost specified in this study represents the value that consumers are perceived to pay for additional information search. Under the assumption of trade-offs between the costs and benefits of search, consumers' decision to search for information is described as follows:

(i) Decide to search for additional information at period t if and only if

$$\max_k ER_{i,t+1|d_{it}=k} - R_{it} \geq c_i \quad \cdots\cdots \text{Eq. (36)},$$

or

(ii) Decide to terminate the information search process and purchase the most preferred product under the following condition:

$$\max_k ER_{i,t+1|d_{it}=k} - R_{it} < c_i \quad \cdots\cdots \text{Eq. (37)}$$

3.4 Likelihood

In the framework of this study, researchers are able to observe consumers' choice to search for information, which is $\left\{\left\{d_{it}\right\}_{t=1}^{T_i}\right\}_{i=1}^{I}$, and their final purchase decision specified by the variable $\left\{Y_i\right\}_{i=1}^{I}$, where the index i denotes the identification number of the consumers and T_i indicates the number of periods spent by consumer i in searching for product information.

Suppose the probability that a consumer i chooses the product j for searching information at period t is defined as $\Pr\left(d_{it}=j\right)$. For $d_{it}=j$, at least two conditions should be satisfied. The first condition is that the expected reward from acquiring information about product j is greater than that expected from searching for information on other products. In other words, $j=\underset{k\in\{1,\cdots,J\}}{\arg\max}\, ER_{i,t+1|d_{it}=k}$. The probability satisfying this condition may be specified as per Equation 38:

$$\Pr\left(ER_{i,t+1|d_{it}=j} > ER_{i,t+1|d_{it}=k},\ \forall k \neq j\right) \cdots\cdots \text{Eq. (38)}$$

The second condition corresponds to Equation 37. Substituting the first condition (Equation 38) in Equation 37, the probability satisfying the second condition is described as follows:

$$
\begin{aligned}
&\Pr\left(\max_k ER_{i,t+1|d_{it}=k} - R_{it} \geq c_i\right) \\
&= \Pr\left(ER_{i,t+1|d_{it}=j} - \max_k U_{ikt}^E \geq c_i\right) \\
&= \Pr\left(ER_{i,t+1|d_{it}=j} - c_i \geq U_{ikt}^E,\ \text{for } \forall k \in \{o,1,\cdots,J\}\right) \qquad \cdots\cdots \text{Eq. (39)} \\
&= \Pr\left(ER_{i,t+1|d_{it}=j} - c_i \geq \varepsilon_{iot}\right) \times \prod_{k=1}^{J} \Pr\left(ER_{i,t+1|d_{it}=j} - c_i - \mathbf{X}_{ik}'\mathbf{b} - \mu_{ik,t-1} \geq \varepsilon_{ikt}\right) \\
&= \exp\left(-\exp\left(-ER_{i,t+1|d_{it}=j} + c_i\right) \times \left(1 + \sum_{k=1}^{J} \exp\left(\mathbf{X}_{ik}'\mathbf{b} + \mu_{ik,t-1}\right)\right)\right)
\end{aligned}
$$

Therefore, the probability of $d_{it} = j$ is calculated by multiplying probabilities specified in Equations 38 and 39.

At T_i, the information search termination period of consumer i, the variable indicating consumers' decision to search is observed to be zero. The probability of $d_{i,T_i} = 0$ is specified as $\Pr\left(\max_k ER_{i,t+1|d_{it}=k} - R_{it} < c_i\right)$. At this termination period, researchers may observe the final purchase decision by consumers. Hence, R_{it} becomes the expected utility of the product that the consumer chooses to purchase. The variable indicating consumers' final decision to purchase is defined as follows:

$$
\begin{cases}
Y_i = j, \ \text{if } U_{ij,T_i}^E \geq U_{ik,T_i}^E \text{ and } U_{ij,T_i}^E > U_{io,T_i}^E \ \text{ for } \forall \mathrm{k} \neq \mathrm{j} \in \{1,\cdots,J\} \\
Y_i = 0, \ \text{if } U_{io,T_i}^E > U_{ik,T_i}^E \ \text{ for } \forall k \in \{1,\cdots,J\}
\end{cases}
\qquad \cdots\cdots\cdots\cdots\cdots\cdots \text{Eq. (40)}
$$

Hence, the probabilities of $Y_i = j$ for $j = 1,\cdots,J$, and $Y_i = 0$ are specified as logit probabilities as follow:

$$\Pr\left[Y_i = j\right] = \frac{\exp\left(\mathbf{X}'_{ij}\mathbf{b} + \mu_{ij,T_i-1}\right)}{1 + \sum_{k=1}^{J} \exp\left(\mathbf{X}'_{ik}\mathbf{b} + \mu_{ik,T_i-1}\right)}, \quad \text{Eq. (41)}$$

and $$\Pr\left[Y_i = 0\right] = \frac{1}{1 + \sum_{k=1}^{J} \exp\left(\mathbf{X}'_{ik}\mathbf{b} + \mu_{ik,T_i-1}\right)} \quad \text{Eq. (42)}$$

Suppose consumer i decides to purchase product j at the termination period. Then, $R_{i,T_i} = U^E_{ij,T_i} = \mathbf{X}'_{ij}\mathbf{b} + \mu_{ij,T_i-1} + \varepsilon_{ijT_i}$. Based on this assumption, the conditional probability of $d_{i,T_i} = 0$ may be specified as follows:

$$\begin{aligned}
&\Pr\left(\max_k ER_{i,T_i+1|d_{iT_i}=k} - R_{iT_i} < c_i \mid Y_i = j\right) \\
&= \Pr\left(\max_k ER_{i,T_i+1|d_{iT_i}=k} - U^E_{ij,T_i} < c_i \mid Y_i = j\right) \\
&= \Pr\left(ER_{i,T_i+1|d_{iT_i}=k} < \mathbf{X}'_{ij}\mathbf{b} + \mu_{ij,T_i-1} + c_i + \varepsilon_{ij,T_i},\ \text{for } \forall k \in \{1,\cdots,J\} \mid Y_i = j\right) \\
&= \prod_{k=1}^{J} \Pr\left(ER_{i,T_i+1|d_{iT_i}=k} - \mathbf{X}'_{ij}\mathbf{b} - \mu_{ij,T_i-1} - c_i < \varepsilon_{ij,T_i} \mid Y_i = j\right)
\end{aligned} \quad \cdots\cdots \text{Eq. (43)}$$

The unconditional probability of $d_{i,T_i} = 0$ is specified as follows:

$$\Pr\left(d_{iT_i}=0\right)=\sum_{j=0}^{J}\left(\begin{array}{l}\Pr\left(\max_k ER_{i,T_i+1|d_{iT_i}=k}-R_{iT_i}<c_i\,|\,Y_i=j\right)\\ \times\Pr\left(Y_i=j\right)\cdot I\left(Y_i=j\right)\end{array}\right)$$

$$=\sum_{j=1}^{J}\left(\begin{array}{l}\prod_{k=1}^{J}\left(1-\exp\left(-\exp\left(-ER_{i,T_i+1|d_{iT_i}=k}+\mathbf{X}_{ij}'\mathbf{b}+\mu_{ij,T_i-1}+c_i\right)\right)\right)\\ \times\dfrac{\exp\left(\mathbf{X}_{ij}'\mathbf{b}+\mu_{ij,T_i-1}\right)}{1+\sum_{q=1}^{J}\exp\left(\mathbf{X}_{iq}'\mathbf{b}+\mu_{iq,T_i-1}\right)}\cdot I\left(Y_i=j\right)\end{array}\right)$$

$$+\prod_{k=1}^{J}\left(1-\exp\left(-\exp\left(-ER_{i,T_i+1|d_{iT_i}=k}+c_i\right)\right)\right)\cdot\frac{I\left(Y_i=0\right)}{1+\sum_{q=1}^{J}\exp\left(\mathbf{X}_{iq}'\mathbf{b}+\mu_{iq,T_i-1}\right)}$$

…… Eq. (44)

In Equations 39 and 44, $\mu_{ij,t-1}$ for $t=2,\cdots,T_i$ and $j=1,\cdots,J$ may be substituted by Equations 24 and 25. Then, the likelihood increment of consumer i is:

$$L_i=\prod_{t=1}^{T_i-1}\left[\sum_{j=1}^{J}\Pr\left(d_{it}=j\right)\cdot I\left(d_{it}=j\right)\right]\times\Pr\left(d_{iT_i}=0\right)$$ …… Eq. (45)

Therefore, the likelihood of observed data is:

$$L=\prod_{i=1}^{I}L_i$$ ……Eq. (46)

Although this study assumes that researchers are able to observe both consumers' choice to search for information and their final purchase decisions, it would be quite possible that researchers may only be able to observe one kind of data set in reality. Given that the case whereby researchers can observe only consumers' final purchase

decisions would be exactly the same for the decision problem using the standard discrete choice model, this study considers the case whereby researchers may only access the observations related to the consumers' information search behavior. As concluded later in this study, the proposed model could be adapted for this case without any theoretical obstacles. Since consumers' decisions to search for information and to purchase the product are both expressed in terms of the same utility function in the proposed model, researchers may analyze consumer preference structures by using only the observed data for information search with a minimal change in the likelihood function.

In the proposed model, the observed data for the final purchase decision only affects the probability to terminate information search. In this case, the unconditional probability of $d_{iT_i} = 0$ may be amended as follows:

$$\begin{aligned}
\Pr\left(d_{iT_i} = 0\right) &= \sum_{j=0}^{J}\left(\Pr\left(\max_k ER_{i,T_i+1|d_{iT_i}=k} - R_{iT_i} < c_i \mid R_{iT_i} = U_{ijT_i}^{E}\right)\cdot \Pr\left(R_{iT_i} = U_{ijT_i}^{E}\right)\right) \\
&= \sum_{j=1}^{J}\left(\begin{array}{l}\prod_{k=1}^{J}\left(1-\exp\left(-\exp\left(-ER_{i,T_i+1|d_{iT_i}=k} + \mathbf{X}_{ij}'\mathbf{b} + \mu_{ij,T_i-1} + c_i\right)\right)\right) \\ \times \dfrac{\exp\left(\mathbf{X}_{ij}'\mathbf{b} + \mu_{ij,T_i-1}\right)}{1+\sum_{q=1}^{J}\exp\left(\mathbf{X}_{iq}'\mathbf{b} + \mu_{iq,T_i-1}\right)}\end{array}\right) \\
&\quad + \prod_{k=1}^{J}\left(1-\exp\left(-\exp\left(-ER_{i,T_i+1|d_{iT_i}=k} + c_i\right)\right)\right)\cdot \frac{1}{1+\sum_{q=1}^{J}\exp\left(\mathbf{X}_{iq}'\mathbf{b} + \mu_{iq,T_i-1}\right)}
\end{aligned} \quad \cdots$$

…… Eq. (47)

With this probability, the likelihood function may be calculated based on Equations 45 and 46, in the same manner as the original suggestion.

3.5 Summary and Discussion

This chapter proposes the structural model for consumers' decision-making process, specifically information search with learning behavior at the pre-purchase stage. The proposed model considers consumers' choices under conditions of uncertainty due to the part of utility that is imperfectly observed by consumers. Hence, the proposed model contains the uncertainty of utility arising from consumers' bounded rationality while the standard discrete choice model only explains the uncertainty on the part of researchers by assuming that consumers are fully rational.

This difference in the premise of the proposed model poses some advantages over the standard discrete choice model when analyzing consumer preference structures. The first advantage is that the proposed model allows researchers to avoid biased and spurious results. In statistical analysis, uncertainty must be specified as a random variable. According to Signorino (2003), since different types of uncertainty generate different specifications of the choice probability, constructing a model with incorrectly specified uncertainty may lead to incorrect inferences resulting from biased and inconsistent estimates. Hence, the proposed model could increase the chance of analyzing consumer preference structures more accurately.

Regarding the uncertainty from the consumer's perspective, the proposed model assumes that the match value captures the entire utility that consumers are unable to observe directly. Since consumers cannot know the true match value before purchasing a product, they perceive the utility based on their beliefs of the match value that vary

depending on idiosyncratic informational signals obtained by individual consumers. In addition, the match value beliefs are defined by a time-varying variable in the framework of the proposed model, whereby consumers update their current beliefs by combining their prior beliefs with the acquired information. In fact, existing literature on consumer learning frameworks have introduced a utility component that contains the uncertainty from the consumer's perspective that is similar to the concept of match value. Although the name varies by researchers, e.g., true mean quality assessment (Shin et al., 2012) or experience utility (Ackerberg, 2003), the uncertain property of the utility component is generally caused by the existence of attributes that are imperfectly observed by consumers.

Consider that uncertainty may also arise from the consumer's limited ability to assess the value of the observed attributes due to limitations of memory or cognitive ability. Assuming that consumers are uncertain about the value of the observed attribute l, the perceived utility obtained by consumer i for product j may be described as follows:

$$U_{ijt} = \sum_{k \neq l} \beta_k x_{ijk} + \tilde{\beta}_{il,t-1} x_{ijl} + \tilde{Q}_{ij,t-1} + \varepsilon_{ijt} \quad \cdots\cdots \text{Eq. (48)}$$

In Equation 48, $\tilde{\beta}_{il,t-1}$ is consumer i's belief on the coefficient of the observed attribute l prior to information search at period t, comprising the true mean coefficient β_l and random error $\zeta_{il,t-1}$; that is, $\tilde{\beta}_{il,t-1} x_{ijl} = \beta_l x_{ijl} + \zeta_{il,t-1} x_{ijl}$. In this problem setting, consumers update their beliefs, on the coefficient of attribute l or on the match values, based on the acquired information. Hence, the informational signal may be (i) only

related to the coefficient of l, (ii) only related to the match value, or (iii) related to both values. However, it is difficult to distinguish between these three cases empirically given the observed data set containing consumer decisions related to the product, including information search and purchase decisions. In addition, to include the uncertainty in the coefficient of a certain attribute in the structural model may be too strict assumption that could possibly lead to the misspecification of the uncertainty. Therefore, it is recommended to adopt a more flexible framework that aggregates the entire utility containing the uncertainty by consumers into the match value belief by setting $\zeta_{il,t-1}$ as zero in the proposed model.

On the other hand, Equation 48 may suggest that allowing heterogeneous consumer preferences in the standard discrete choice model could capture all the effects of uncertainty as explained by the proposed model. However, the standard discrete choice model with heterogeneous consumer preferences can only capture the mean consumer behavior, and not the dispersion generated from idiosyncratic information acquired by individual consumers (Ackerberg, 2003). To discuss this more specifically, the utility described in Equation 12 should be re-specified by defining some additional variables: (i) the deviation in the perception, ω_{ijt}, defined as $\omega_{ijt} = \tilde{Q}_{ijt} - \mu_{ijt}$ with probability density $\omega_{ijt} \sim N\left(0, \sigma_{ijt}^2\right)$ based on Equations 23 to 25, (ii) perception bias, ν_{ijt}, defined as $\nu_{ijt} = \mu_{ijt} - Q_{ij}$, and (iii) signal noise, η_{ijt}, defined as $\eta_{ijt} = S_{ijt} - Q_{ij}$ with probability density $\eta_{ijt} \sim N\left(0, \sigma_{S_{ij}}^2\right)$ based on Equation 16. One must carefully understand that $\tilde{Q}_{ijt}$ is not an unbiased measurement of the true match value Q_{ij} because consumers have

bounded rationality. Therefore, the per-period unbiased mean of the perception $\tilde{Q}_{ijt}$ is μ_{ijt}, and not Q_{ij}. Thus, the utility of consumer i for product j at time t may be specified as follows:

$$\begin{aligned} U_{ijt} &= \mathbf{X}_{ij}'\mathbf{b} + \tilde{Q}_{ij,t-1} + \varepsilon_{ijt} \\ &= \mathbf{X}_{ij}'\mathbf{b} + \mu_{ij,t-1} + \omega_{ij,t-1} + \varepsilon_{ijt} \\ &= \mathbf{X}_{ij}'\mathbf{b} + \frac{\sigma_{S_{ij}}^2 \mu_{ij0} + \sigma_{ij0}^2 \sum_{\tau=1}^{t-1} \left(I\left(d_{i\tau} = j\right) \cdot S_{ij\tau} \right)}{\sigma_{S_{ij}}^2 + \sigma_{ij0}^2 \sum_{\tau=1}^{t-1} I\left(d_{i\tau} = j\right)} + \omega_{ij,t-1} + \varepsilon_{ijt} \qquad \cdots\cdots\cdots\cdots\cdots \text{ Eq. (49)} \\ &= \mathbf{X}_{ij}'\mathbf{b} + Q_{ij} + \frac{\sigma_{S_{ij}}^2 \nu_{ij0} + \sigma_{ij0}^2 \sum_{\tau=1}^{t-1} \left(I\left(d_{i\tau} = j\right) \cdot \eta_{ij\tau} \right)}{\sigma_{S_{ij}}^2 + \sigma_{ij0}^2 \sum_{\tau=1}^{t-1} I\left(d_{i\tau} = j\right)} + \omega_{ij,t-1} + \varepsilon_{ijt} \end{aligned}$$

On the other hand, the utility under heterogeneous consumer preferences in the standard discrete choice model is defined as follows:

$$U_{ijt} = \alpha_{ij} + \mathbf{X}_{ij}'\mathbf{b}_i + \varepsilon_{ijt}, \text{ where } \alpha_{ij} \sim N\left(\alpha_j, \sigma_\alpha^2\right) \text{ and } \mathbf{b}_i \sim MVN\left(\mathbf{b}, \Sigma\right) \cdots\cdots \text{ Eq. (50)}$$

Comparing both equations, the consumers' mean utility function may be specified as follows:

$$\mathbf{X}_{ij}'\mathbf{b} + E\left[Q_{ij} + \frac{\sigma_{S_{ij}}^2 \nu_{ij0}}{\sigma_{S_{ij}}^2 + \sigma_{ij0}^2 \sum_{\tau=1}^{t-1} I\left(d_{i\tau} = j\right)} \right] = \mathbf{X}_{ij}'\mathbf{b} + \alpha_j \cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots \text{ Eq. (51)}$$

From Equation 51, the average effect of searching for information may be captured by the mean in the standard discrete choice model with heterogeneous consumer preferences. However, while the variances of utility in Equation 49 depend on the number of informational signals acquired or, effectively, the time spent in information search, the variance of utility in Equation 50 does not. Therefore, allowing heterogeneity in the standard discrete choice model cannot explain either the shift in variance by searching for information, or the dispersion in consumer behavior depending on the difference in information possessed by different consumers, unlike the proposed model. This may lead to inefficient estimates and spurious results in the empirical study (Ackerberg, 2003).

The second advantage comes from the premise related to the information search behavior of consumers. This study assumes that consumers decide whether to search for a single, additional piece of information by comparing the marginal cost of search versus the additional reward from searching that is expected to be gained at the next period. This strategy of searching may be labeled as the sequential strategy under the one-step look-ahead rule (Bather, 2000). On the other hand, most studies on either consumer search models or consumer learning models assume full forward-looking consumer behavior (e.g., De los Santos et al., 2012; Erdem & Keane, 1996; Kim et al., 2010; Shin et al., 2009; Weitzman, 1979). In fact, the reservation utility, the optimal strategy in sequential search model that was introduced in Section 2.2, is the solution of the dynamic programming over all the alternatives available to consumers (Weitzman, 1979). In addition, the expected utility in the consumer learning model is the solution of the dynamic programming over the entire period (Erdem & Keane, 1996). However, this

assumption of full forward-looking behavior has been criticized for inadequately describing the actual behavior of human agents who have limited cognitive and processing abilities (Assunção & Meyer, 1993; Camerer, Ho, & Chong, 2004; Gabaix, Laibson, Moloche, & Wienberg, 2006; Hutchinson & Meyer, 1994; Yang, Toubia, & De Jong, 2015). Furthermore, Camerer et al. (2004) and Gabaix et al. (2006) find evidence that the model that assumes forward-looking behavior of only one or two steps shows a better fit with the actual data than the model that assumes full forward-looking behavior of consumers. In addition, Hauser (2011) observes that consumers only look one-step ahead from the information. These studies support the claim that the one-step look-ahead search strategy proposed in this study is more appropriate for explaining actual consumer search behavior.

Assuming consumer search strategy to be sequential may be similarly justified. If the assumption of simultaneous search strategy were more appropriate to explain consumer behavior, consumers would be able to observe sets of information about all the products and evaluate their utility at the same time. However, due to the limitations of the human brain, this may not be possible without the aid of information sources that organize and provide the search results, such as a table that gathers and structures the information of all available products (Kundisch, 2000). In addition, Kundisch (2000) claims that surfing through different web sources to evaluate products is akin to a sequential search while a search for the seller providing the best price based on a price database represents a simultaneous search. Since this study considers the aim of consumer search behavior to be to reduce the uncertainty in their utility by obtaining reviews provided by other consumers and given that most reviews contain information for a single product, the

sequential search strategy is adequate for the context of this study.

In summary, the first premise allows the proposed model to adapt the uncertainty in consumers' perspective that is ignored by the standard discrete choice model for making purchase decisions. Furthermore, the first and second premises allow the proposed model to explain consumer behavior to update their utility while they search for information based on the structural model for consumer search as explained in Section 3.2. Lastly, the third and fourth premises enable explaining consumer decisions to actively search for product information based on the tradeoff between the benefits and costs of search. The structural model for consumer learning is similar to a passive information search, focusing on the information that could be acquired prior to purchase through incidental exposure to commercial messages. Hence, Section 3.3 highlights the major differences of the proposed study from existing literature in the domain of consumer learning models.

Chapter 4. Simulation studies

In this chapter, the results of simulation studies are discussed to examine the consumer behavior dynamics based on the proposed model. Monte Carlo experiments are conducted to simulate consumer decisions, based on assumed parameter values or choice situations.

4.1 Design of Monte Carlo Experiments

This study conducts Monte Carlo experiments to simulate consumer behavior. To reveal the effect of the match value on the utility directly, the simulation studies are conducted based on the expected utility function substituting informational signals for the summation of the match value and signal noise in Equation 26. Hence, the expected utility function obtained by consumer i for product j at period t is specified as follows:

$$\begin{aligned} U_{ijt}^{E} &= \mathbf{X}_{ij}'\mathbf{b} + \mu_{ij,t-1} + \varepsilon_{ijt} \\ &= \mathbf{X}_{ij}'\mathbf{b} + \frac{\sigma_{S_{ij}}^2}{\sigma_{S_{ij}}^2 + \sigma_{ij0}^2 \sum_{\tau=1}^{t-1} I\left(d_{i\tau} = j\right)} \mu_{ij0} + \frac{\sigma_{ij0}^2 \sum_{\tau=1}^{t-1} I\left(d_{i\tau} = j\right)}{\sigma_{S_{ij}}^2 + \sigma_{ij0}^2 \sum_{\tau=1}^{t-1} I\left(d_{i\tau} = j\right)} Q_{ij} \quad \cdots\cdots \text{ Eq. (52)} \\ &+ \frac{\sigma_{ij0}^2}{\sigma_{S_{ij}}^2 + \sigma_{ij0}^2 \sum_{\tau=1}^{t-1} I\left(d_{i\tau} = j\right)} \sum_{\tau=1}^{t-1} \left(I\left(d_{i\tau} = j\right) \cdot \eta_{ij,\tau} \right) + \varepsilon_{ijt} \end{aligned}$$

At each period t in each Monte Carlo simulation, signal errors $\{\eta_{ij1},\cdots,\eta_{ij,t-1}\}$ are created as they are randomly generated from the normal distribution with zero mean and variance $\sigma^2_{S_{ij}}$ independently at previous time periods. The expected utility for the "no purchase" option is also assumed to be normalized to zero as per Equation 15.

As discussed in Section 3.3, consumer's decision to search for information based on the expected function in Equation 52 is simulated using the expected reward conditional on $d_{it}=k$. The expected reward obtained by a consumer i at period $t+1$ conditional on $d_{it}=k$ is specified in Equation 53. Similar to Equation 35, the expectation of the reward at the next period is taken with respect to $\mathbf{e_{t+1}}$ and η_{ikt} based on the information set $\mathbf{I_{i,t-1}}$ since η_{ikt} is stochastic for consumers in this case, as the informational signal S_{ikt} is substituted by the sum of the deterministic mean Q_{ik} and the stochastic error η_{ikt}.

$$
\begin{aligned}
ER_{i,t+1|d_{it}=k} &= E_{\eta_{ikt}}\left[E_{\mathbf{e_{t+1}}}\left[\max\left\{U^{E}_{io,t+1},U^{E}_{i1,t+1},\cdots,U^{E}_{iJ,t+1}\right\}\right]|\,\mathbf{I_{i,t-1}};d_{it}=k\right] \\
&= E_{\eta_{ikt}}\left[\ln\left(1+\exp\left(\begin{array}{l}
\mathbf{X}'_{ik}\mathbf{b}+\dfrac{\sigma^2_{S_{ik}}\mu_{ik,0}}{\sigma^2_{S_{ik}}+\sigma^2_{ik,0}\left(\sum_{\tau=1}^{t-1}I\left(d_{i\tau}=k\right)+1\right)} \\
+\dfrac{\sigma^2_{ik,0}\cdot Q_{ik}\cdot\left(\sum_{\tau=1}^{t-1}I\left(d_{i\tau}=k\right)+1\right)}{\sigma^2_{S_{ik}}+\sigma^2_{ik,0}\left(\sum_{\tau=1}^{t-1}I\left(d_{i\tau}=k\right)+1\right)} \\
+\dfrac{\sigma^2_{ik,0}\sum_{\tau=1}^{t-1}\left(I\left(d_{i\tau}=k\right)\cdot\eta_{ik\tau}\right)}{\sigma^2_{S_{ik}}+\sigma^2_{ik,0}\left(\sum_{\tau=1}^{t-1}I\left(d_{i\tau}=k\right)+1\right)} \\
+\dfrac{\sigma^2_{ik,0}\cdot\eta_{ikt}}{\sigma^2_{S_{ik}}+\sigma^2_{ik,0}\left(\sum_{\tau=1}^{t-1}I\left(d_{i\tau}=k\right)+1\right)}
\end{array}\right)\right. \\
&\qquad\left.+\sum_{j\neq k}\exp\left(\begin{array}{l}
\mathbf{X}'_{ij}\mathbf{b}+\dfrac{\sigma^2_{S_{ij}}\mu_{ij0}}{\sigma^2_{S_{ij}}+\sigma^2_{ij,0}\sum_{\tau=1}^{t-1}I\left(d_{i\tau}=j\right)} \\
+\dfrac{\sigma^2_{ij,0}\cdot Q_{ij}\cdot\sum_{\tau=1}^{t-1}I\left(d_{i\tau}=j\right)}{\sigma^2_{S_{ij}}+\sigma^2_{ij,0}\sum_{\tau=1}^{t-1}I\left(d_{i\tau}=j\right)} \\
+\dfrac{\sigma^2_{ij,0}\sum_{\tau=1}^{t-1}\left(I\left(d_{i\tau}=j\right)\cdot\eta_{ij\tau}\right)}{\sigma^2_{S_{ij}}+\sigma^2_{ij,0}\sum_{\tau=1}^{t-1}I\left(d_{i\tau}=j\right)}
\end{array}\right)\right)\Bigg]\,|\,\mathbf{I_{i,t-1}};d_{it}=k\Bigg]
\end{aligned}
$$

…… Eq. (53)

Equation 53 is calculated based on Monte Carlo integration, which is the method of calculating the integration of the Lebesgue integrable function $f(x)$ by taking the

expected value, or average, of $f\left(x_n\right)$ at random locations x_n, $n=1,\cdots,N$, sampled from $U\left[0,1\right]$. The number of random points for the Monte Carlo integration is 1,000,000.

Since consumers choose to search for information of the product that is expected to yield maximum expected gains, this study simulates consumers' search decisions by computing the expected returns in Equation 53 for all products and comparing their values. In other word, if $j=\underset{k\in\{1,\cdots,J\}}{\arg\max}\, ER_{i,t+1|d_{it}=k}$, set $d_{it}=j$ with probability as follows:

$$
\begin{aligned}
&\Pr\left(ER_{i,t+1|d_{it}=j}-\max\left\{U_{iot}^{E},U_{i1t}^{E},\cdots,U_{iJt}^{E}\right\}\geq c_i\right)\\
&=\Pr\left(ER_{i,t+1|d_{it}=j}-c_i\geq\varepsilon_{iot}\right)\prod_{k=1}^{J}\Pr\left(ER_{i,t+1|d_{it}=j}-c_i\geq U_{ikt}^{E}\right)\\
&=\exp\left(-\exp\left(-ER_{i,t+1|d_{it}=j}+c_i\right)\times\left(1+\sum_{k=1}^{J}\exp\left(\begin{array}{l}\mathbf{X}_{ik}'\mathbf{b}+\dfrac{\sigma_{S_{ik}}^{2}}{\sigma_{S_{ik}}^{2}+\sigma_{ik0}^{2}\sum_{\tau=1}^{t-1}I\left(d_{i\tau}=k\right)}\\ +\dfrac{\sigma_{ik0}^{2}Q_{ik}\sum_{\tau=1}^{t-1}I\left(d_{i\tau}=k\right)}{\sigma_{S_{ik}}^{2}+\sigma_{ik0}^{2}\sum_{\tau=1}^{t-1}I\left(d_{i\tau}=k\right)}\\ +\dfrac{\sigma_{ik0}^{2}\sum_{\tau=1}^{t-1}\left(I\left(d_{i\tau}=k\right)\cdot\eta_{ik\tau}\right)}{\sigma_{S_{ik}}^{2}+\sigma_{ik0}^{2}\sum_{\tau=1}^{t-1}I\left(d_{i\tau}=k\right)}\end{array}\right)\right)\right)
\end{aligned}
$$

……Eq. (54)

Hence, this study generates the random variable u from the uniform distribution $U\left[0,1\right]$ and compares it with the probability calculated from Equation 54. If the value of u is less than the value computed by Equation 54, the consumer's decision at the period

t is simulated as searching for additional information about product j and delaying the purchase decision to the next period. Thus, we proceed to the next period ($t = t+1$), generating η_{ikt} from the normal distribution with mean zero and variance $\sigma^2_{S_{ik}}$ and setting η_{ijt} to be zero for $j \neq k$.

On the other hand, the consumer decides to stop searching for information and makes a purchase decision when u is greater than the probability calculated from Equation 54. To simulate the consumer's final purchase decision, this study samples random variable z from a multinomial distribution whereby $z = j$ ($j = 1, \cdots, J$) with probability $p_j = \dfrac{e^{\mathbf{X}'_{ij}\mathbf{b} + \mu_{ij,T_i-1}}}{1 + \sum_{k=1}^{J} e^{\mathbf{X}'_{ik}\mathbf{b} + \mu_{ik,T_i-1}}}$, and $z = 0$ with the probability $1 - \sum_{j=1}^{J} p_j$. After this, we set $Y_{i,T_i} = z$.

Following the above procedure, this study simulates artificial paths of the consumer decision-making process. Technically, simulated paths in this study depend on the sequence of the random variable $\{\eta_{ij,t}\}_{i,j,t \geq 1}$ and probabilities to search for information and to purchase. In addition, these probabilities depend on the parameters introduced in the proposed model. Hence, the paths simulating consumer's decision-making behavior tend to differ if one of the parameters in the proposed model is changed.

This study implements three Monte Carlo experiments. The first experiment examines the direct effect of parameters related to a particular product on consumers' decision making for the product. In order to investigate this, the study assumes a choice situation of a single consumer for a single product. Given that consumers choose which product to

search for information on by comparing the expected reward after acquiring the information about the product, a change in the parameters related to the product may influence the consumers' behavior towards other products. Hence, the second experiment simulates the choice decision of a single consumer between two products in order to confirm the dynamics of information search among multiple products. Lastly, the third experiment simulates the decisions of multiple consumers with heterogeneous tastes in the match values to track the dispersion in the decision making process of consumers.

4.2 Monte Carlo Experiment 1: Changes in Parameters

Compared to conventional discrete choice models, the proposed model in this study contains parameters related to the consumers' information search and learning behavior such as Q_{ij}, μ_{ij0}, σ^2_{ij0}, $\sigma^2_{S_{ij}}$, and c_i. Since these parameters are newly introduced in this study, it is necessary to investigate them, including how a change in these parameters affects consumers' behavior to search for information. To simulate consumer decisions based on arbitrary values of parameters is helpful to understand the dynamics between these parameters.

This section considers behavioral changes of one specific consumer I in making purchase decisions for one specific product J, in order to examine the direct effect of parameters on consumer behavior. For simplicity, the observable part of utility is assumed to be normalized to zero. Following the procedure explained in Section 4.1, this study simulates 10,000 paths of the consumer's decision-making process. To examine how the

decision of a consumer changes depending on a change in parameters in the proposed model, this experiment compares the simulation results from different scenarios that assume a combination of various parameter values. Specifically, this experiment focuses on revealing the change in the volume of information searched by the consumer given different scenarios. In the baseline scenario, parameter values are defined as described in Table 1.

Table 1. Parameter values in the baseline scenario (Experiment 1)

Parameter	Definition	Value
Q_{IJ}	Match value obtained by consumer I from product J, which is unobserved by the consumer and researchers	2.0
μ_{IJ0}	Consumer I's initial expectation (mean) of the match value: Mean of consumer I's initial match value belief	1.0
σ^2_{IJ0}	Variance of consumer I's initial match value belief	1.0
$\sigma^2_{S_{IJ}}$	Variance of the informational signal around the match value	1.0
c_I	Search cost	0.5

In the baseline scenario, the average number of periods that the consumer searches for information over 10,000 paths is 0.8348, with a corresponding maximum and minimum of 9 and 0, respectively. The histogram showing the frequency of paths based on the

amount of information searched by the consumer is presented in Figure 6.

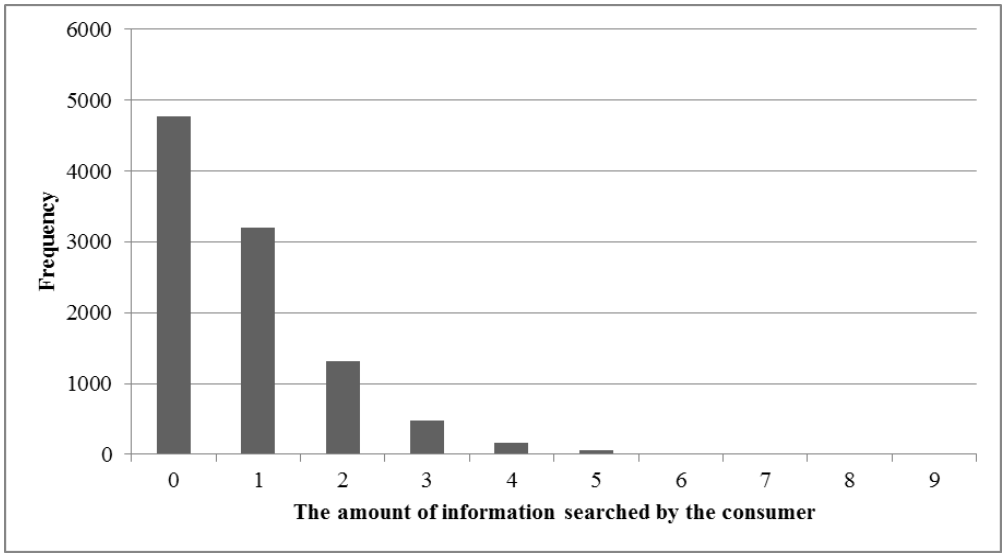

Figure 6. Histogram of the amount of information searched in the baseline scenario of Experiment 1

To confirm the influence of parameters on consumers' decisions, this study firstly considers the search cost c_I. It is well established that higher search costs reduce the consumer's motivation to search, thereby decreasing the amount of information searched by the consumer (e.g., Bei et al., 2004; Stigler, 1961; Su, 2008). To confirm that the proposed model is able to reflect this finding, this study compares the simulation results based on scenarios of different values of search costs, ceteris paribus. When the values of search cost are changed from 0 (no cost incurred) to 2 (four times the effort needed to search for information compared to the baseline scenario), the change in the average amount of information searched by the consumer is presented in Figure 7. As shown in

this figure, consumers tend to reduce their search for information as search cost increases.

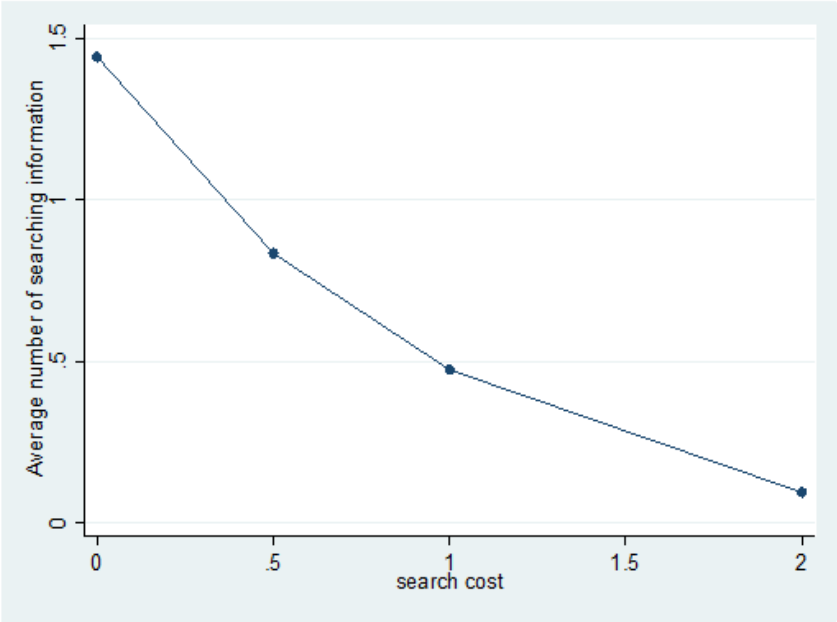

Figure 7. Changes in the average amount of information searched based on changes in search cost

Secondly, this study considers the variance of the consumer's initial match value belief, σ_{IJ0}^2. Technically, if the consumer has a large variance of their initial match value belief, it means that the consumer presumes the match value within broad intervals at the beginning of the decision making process. In other words, the consumer's initial perception of the uncertainty of the match value is greater when the consumer has a larger value of σ_{IJ0}^2. This explanation corresponds to the specification in existing studies on consumer learning (Erdem and Keane, 1996; Shin et al., 2012). Given that one of the consumer's objectives of searching for information is to reduce uncertainty by definition,

this study hypothesizes that consumers attempt to gather more information under conditions of higher uncertainty. Figure 8 shows the simulation results of the average amount of information searched by the consumer by changing the value of σ^2_{IJ0}. Based on Figure 8, consumers tend to search for more information as the uncertainty increases, supporting the hypothesis.

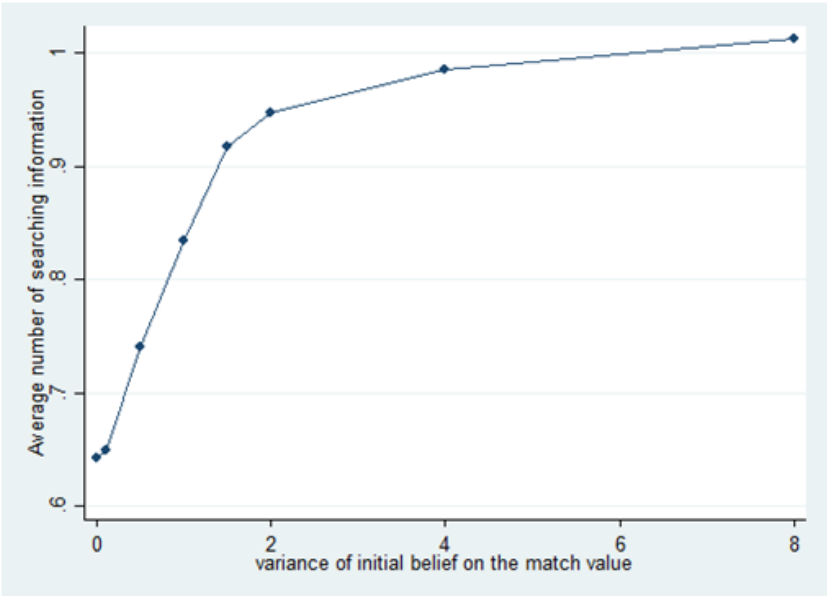

Figure 8. Changes in the average amount of information searched based on changes in the variance of initial match value belief

Thirdly, the variance of the informational signal of the match value, $\sigma^2_{S_{IJ}}$, is discussed. By definition, a greater value of $\sigma^2_{S_{IJ}}$ implies that the informational signal transfers a value that is further away from the true match value. Hence, this value is related to the accuracy of the information source (Erdem & Keane, 1996; Erdem, Keane,

Öncü, & Strebel, 2005). Figure 9 shows that consumers tend to search for less information as the variance of informational signal increases. This result could be explained intuitively; consumers are unlikely to search for information when they are doubtful about its accuracy. Mathematically, the weight of the informational signal in the updating process is reduced when the variance of the informational signal increases in Equation 52. Therefore, consumers' motivation to search for information is reduced as they place lower or no importance on the informational signal.

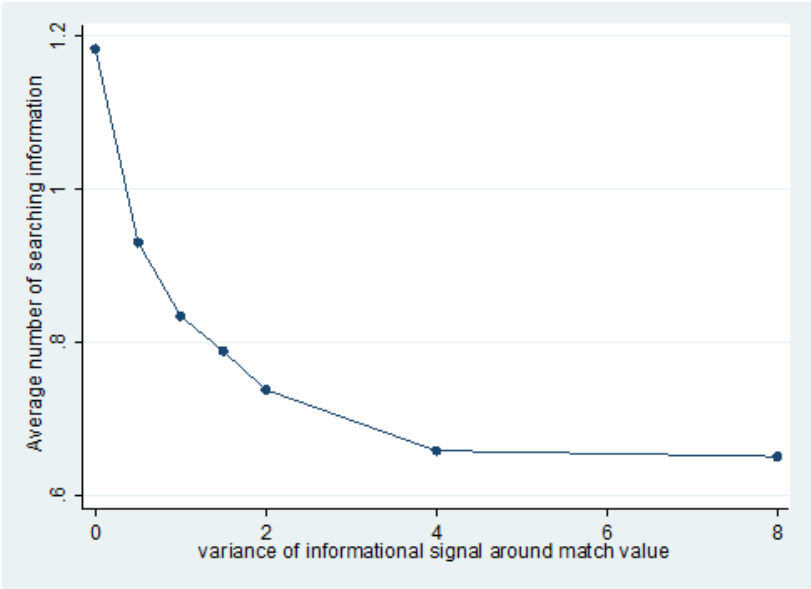

Figure 9. Changes in the average amount of information searched based on changes in the variance of the informational signal of the match value

Now, consider a change in the mean of the initial match value belief, μ_{IJ0}. Figure 10 shows the simulation results. Based on the results, how far μ_{IJ0} is from the match value

appears to be important in this case. It appears that consumers have little motivation to search for information if they overestimate the value obtained by purchasing the product compared to the actual match value. This result may be explained by the assumption of the utility maximizing behavior of rational consumers. Although searching for information and learning the match value decreases the uncertainty associated with the match value, the expected product utility obtained by the consumer decreases as well. Hence, the benefit from searching for additional information may not be sufficient to exceed the search cost. Consider the case whereby the consumer's initial expectation of the match value is positive but smaller than the true match value. In this case, if the difference between the consumer's initial expectation and the match value is reduced, the consumer's motivation to search for information is also reduced. The benefit obtained from the additional information search reduces as the consumer accurately guesses the match value at the initial period.

On the other hand, one notable finding may be deduced from simulation results on scenarios assuming negative values for μ_{IJ0}. The simulation results show that consumers tend to search for less information the more pessimistic they are about the product. This is the contrary proof for the intuitive hypothesis that consumers are more likely to search for information if the difference between the match value and the initial expectation is greater as this also increases the potential benefit from information search. This result seems to support the biased learning behavior by the consumer. According to Alloy and Tabachnik (1984), biased learning means that the learning process fails to converge to the true probability distribution. This biased processing of information may be caused by consumers' confirmatory behavior, i.e., the desire of consumers to confirm their thoughts

or decisions (Chylinski, Roberts, & Hardie, 2012; Hoeffler, Ariely, & West, 2006). Hence, it may be explained that consumers are likely to be less motivated to search for information the more negatively they assess the product even if the true match value is positive, as they tend to resist disconfirmation.

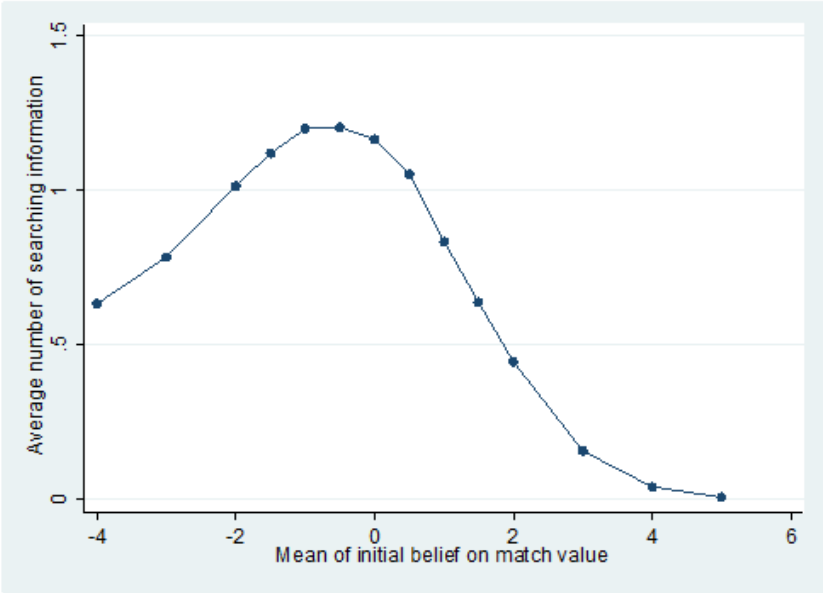

Figure 10. Changes in the average amount of information searched based on changes in the mean initial match value belief (positive)

This leaves us with the scenario of negative match values. Figure 11 shows the change in the average number of information searched by the consumer based on different mean values of initial belief given the match value of -2.0. The graph in Figure 11 appears to be a mirror image of the graph in Figure 10. Consumers are less likely to search for information when they have more optimistic expectations about the product as updating the match value may oppose or disconfirm their initial belief.

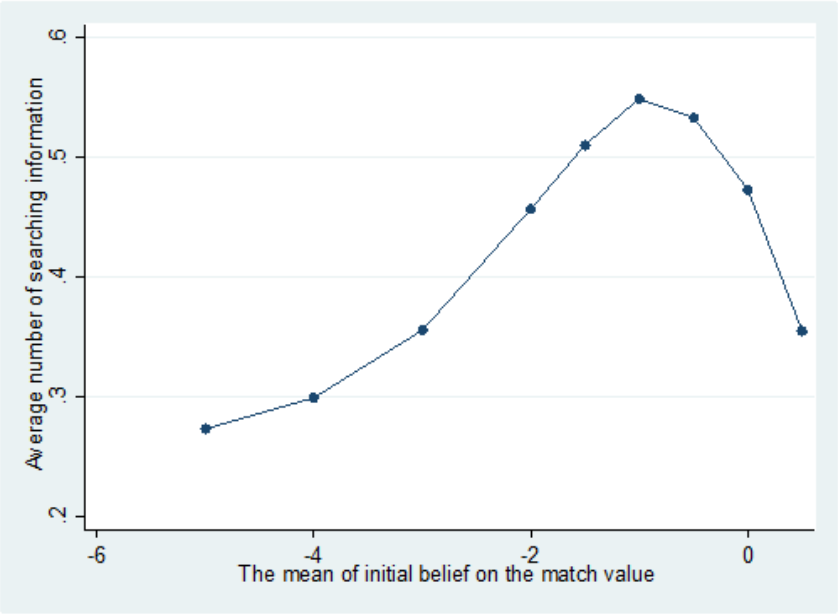

Figure 11. Changes in the average amount of information searched based on changes in the mean initial match value belief (negative)

4.3 Monte Carlo Experiment 2: Dynamics of Search

In the proposed model, the consumer's decision to select a product for information search at the current period depends on a comparison of the expected gains from searching for information on each product as per Equation 53. Based on this equation, the expected gains depend on parameters related to the information search and information acquired up to that period. Hence, consumer behavior varies with a change in the parameters of an individual product. To facilitate the understanding of consumers'

decisions among products, this section simulates consumer behavior assuming the scenario whereby a single consumer I makes purchase decisions between two products, $j=1$ and $j=2$.

Following the procedures explained in Section 4.1, this study simulates 1,000 paths of the consumer's decision-making process. Similar to the first experiment, this experiment examines how the consumer decision changes depending on changes in the parameters of each product by comparing the simulation results from various scenarios whereby the product parameters take on different combinations of values.

Specifically, this experiment examines the decision of the consumer in the case where two products have identical preference structures (i.e., identical preference scenario). In other words, all the parameters of expected utility, such as parameters in the observed part of utility, match value, the mean and variance of initial perception of match value, variance of informational signal, are identical for both products in this scenario as described in Table 2. For simplicity, the observed parts of utility for both products are assumed to be zero and the search cost c_I is assumed to be 0.5.

Table 2. Parameter values in the identical preference scenario

Parameter	Product 1	Product 2
Q_{Ij}	2.0	2.0
μ_{Ij0}	1.0	1.0
σ^2_{Ij0}	1.0	1.0
$\sigma^2_{S_{Ij}}$	1.0	1.0

From a practical perspective, the average tendency to search for information and to purchase is similar for both products even if individual paths show different results. For example, some paths search for more information on product 1 than product 2 and make a final purchase decision in favor of product 1 while other paths behave contrary to this. For 1,000 simulated paths, the average of the total number of periods that the consumer searches for information is 2.546. In addition, the average number of informational signals acquired about products 1 and 2 is similar at 1.265 and 1.281 respectively. The histogram showing the frequency of paths based on the number of information searched by the consumer for each product is presented in Figure 12. The median of the total number of information searched is two, with the corresponding number for each product at one each. This implies that 50% of the simulated paths search for product information no more than once. This result verifies that the average tendency of the consumer to search for information coincides with general opinion. Figure 13 shows the histogram of the final purchase decision in 1,000 simulated paths. Similar to the average tendency of the information search, the ratio of product purchase to the total number of paths is similar across both products.

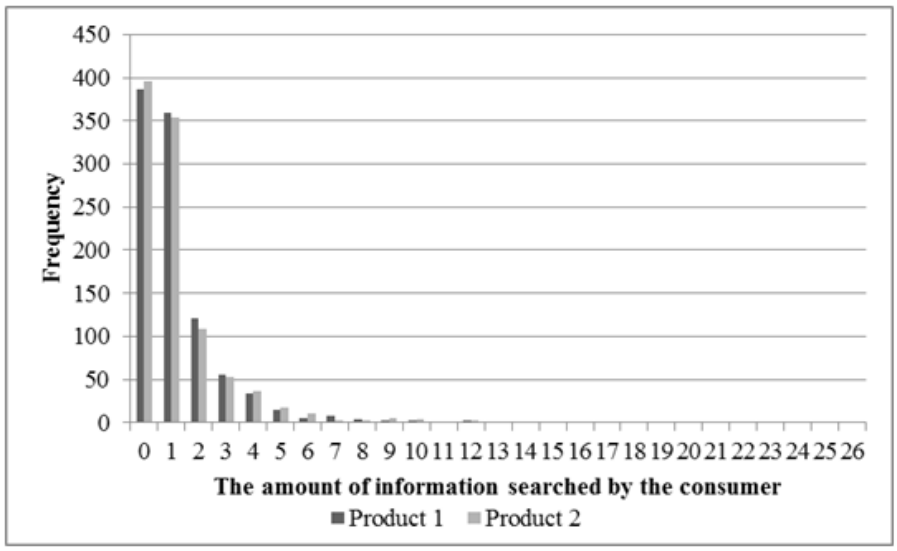

Figure 12. Histogram of the amount of information searched for each product in the identical preference scenario

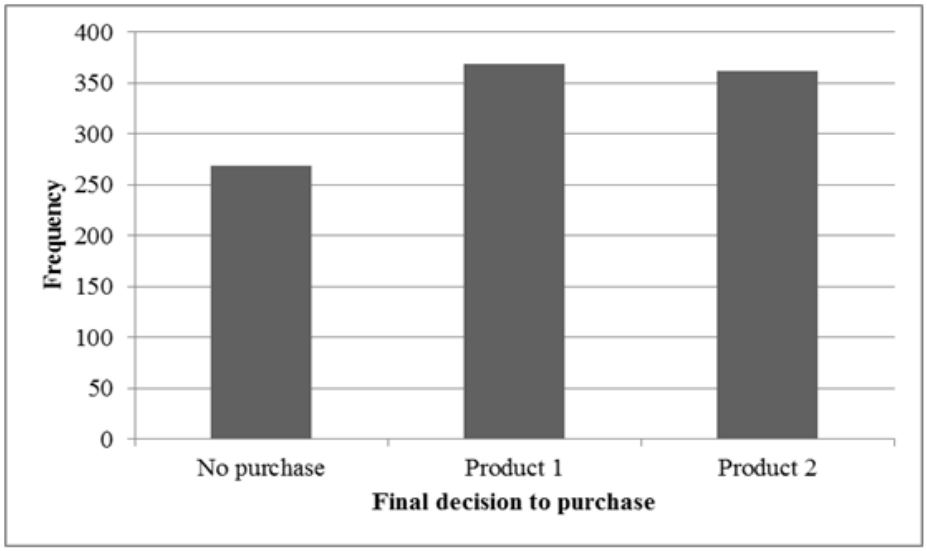

Figure 13. Histogram of the final purchase decision in the identical preference scenario

To confirm the dynamics of information search between two products, this study presents the simulation results of several scenarios, assuming differences in the preference structures between two products.

Firstly, this study examines scenarios whereby the consumer has a certain utility for one of the products while he/she is uncertain about the match value of the other product, ceteris paribus. Assume that the consumer knows the exact match value of product 1 even before acquiring any information (i.e., scenario of full certainty about product 1). As shown in Table 3, the parameter values are assumed to be the same as in the identical preference scenario except for the mean and variance of the initial match value belief of product 1. Since consumers know their exact preference for product 1 right from the beginning, the mean of the initial match value belief has the same value as the match value and the variance of the initial belief becomes zero.

Table 3. Parameter values in the scenario of full certainty about product 1

Parameter	Product 1	Product 2
Q_{Ij}	2.0	2.0
μ_{Ij0}	**2.0**	1.0
σ^2_{Ij0}	**0.0**	1.0
$\sigma^2_{S_{Ij}}$	1.0	1.0

For 1,000 simulated paths in the scenario of full certainty about product 1, the average of the total number of periods that the consumer searches for information is 2.381. Compared to the identical preference scenario, the total amount of information searched

by the consumer seems to be similar. However, the average number of acquired information about product 1 is 0.437 while that about product 2 is 1.944. Hence, the assumption of certainty about product 1 by the consumer changes not only the amount of information searched about product 1, but also of product 2. This may be interpreted as consumers being able to afford searching for more information about product 2 under the same constraints of searching costs as in the identical scenario, as they are less motivated to search for information about product 1.

Figure 14 shows the histogram of the frequency of paths based on the number of information searched about each product by the consumer. As shown in the figure, almost 90% of the simulated paths do not search for any information about product 1 while about 70% of the simulated paths search for information about product 2 at least once. Figure 15 presents the frequency of the final purchase decision in 1,000 simulated paths. Unlike the identical preference scenario, the proportion of purchase decisions for product 1 is significantly higher than product 2. This may be explained as the consumer is more likely to purchase product 1 given that the mean of the initial belief on the match value of product 1 is greater than that of product 2 and the consumer is certain of his/her belief about product 1.

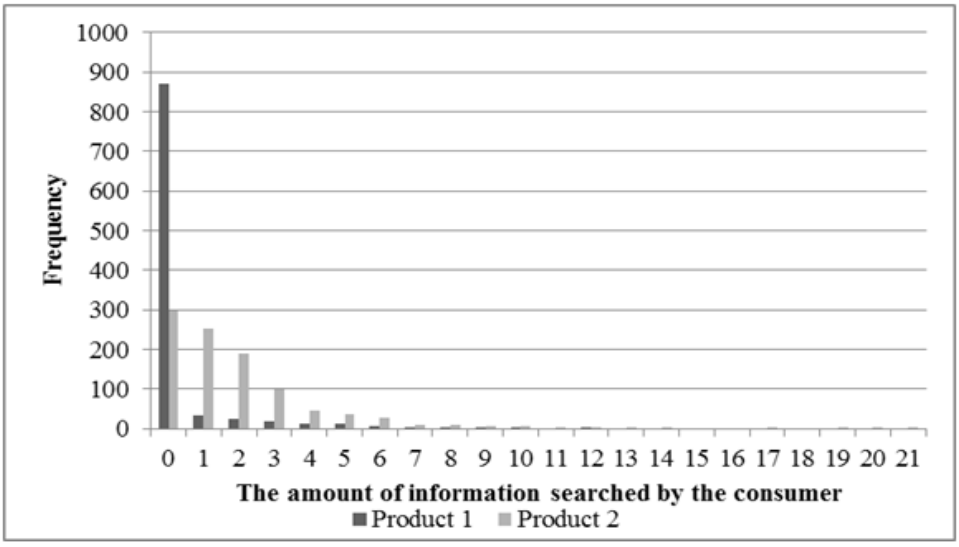

Figure 14. Histogram of the amount of information searched for each product in the scenario of full certainty about product 1

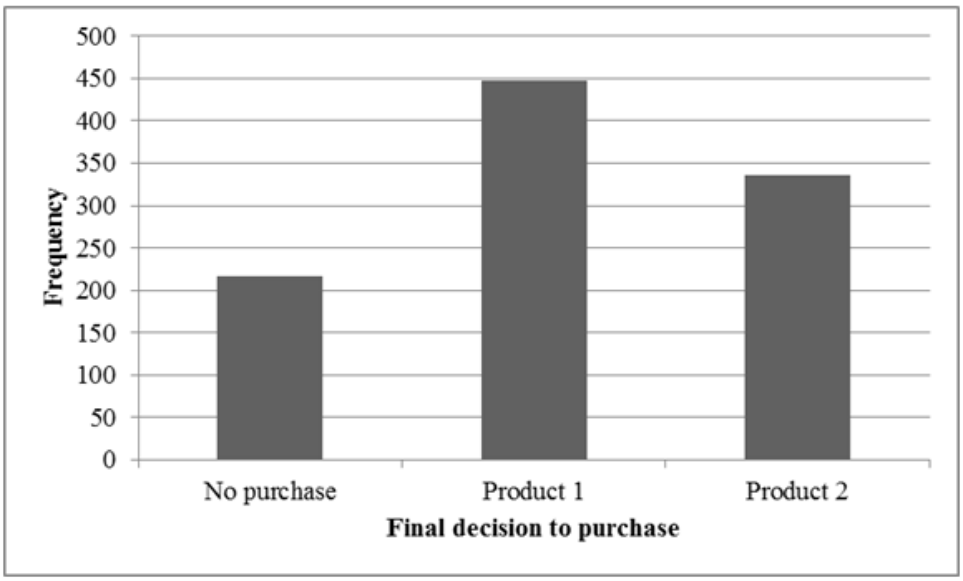

Figure 15. Histogram of the final purchase decision for each product in the scenario of full certainty about product 1

Before we proceed further, this study examines two additional scenarios for the conditions related to the uncertainty associated with product 1. The first one is that the consumer presumes the match value of product 1 accurately but uncertainly before acquiring any information. This scenario implies that the consumer has an initial perception of the uncertainty about product 1 without initial bias (i.e., scenario of unbiased but uncertain belief about product 1). On the other hand, the second scenario assumes that the consumer is sure that an incorrect initial match value is true (i.e., scenario for biased but certain belief about product 1). Table 4 documents the assumed parameter values for these two scenarios. Specifically, the second and third columns show the parameter values in the scenario for unbiased but uncertain belief about product 1 and the scenario of biased but certain belief about product 1, respectively.

Table 4. Parameter values in the scenarios related to uncertainty about product 1

Parameter	Product 1		Product 2
	Unbiased and uncertain	Biased and certain	
Q_{Ij}	2.0	2.0	2.0
μ_{Ij0}	2.0	1.0	1.0
σ^2_{Ij0}	1.0	0.0	1.0
$\sigma^2_{S_{Ij}}$	1.0	1.0	1.0

After simulating 1,000 paths for each scenario, the average of the total number of periods that the consumer searches for information is 2.495 and 2.678 in the two scenarios, respectively. These values are similar to the total amount of information

searched by the consumer in the identical preferences scenario and the scenario of full certainty about product 1. However, the average number of acquired information about product 1 for the scenario for unbiased but uncertain belief about product 1 is derived to be 0.594 while the corresponding figure for the scenario of biased but certain belief about product 1 is 0.829. This may be explained by the informational signal. As the informational signal provides a hint of the true match value of the product, consumers are likely to be affected by the difference between the informational signal they obtain and their match value belief. Hence, the consumer tends to search for more information about product 1 when he/she is certain about an incorrect initial belief about its match value compared to the case when the consumer is uncertain about an initial belief that is correct. On the other hand, the average number of acquired information about product 2 is similar in both scenarios, at 1.901 and 1.849 respectively. These values are slightly lower compared to those of the full certainty scenario because the motivation to search for information about product 1 slightly increases as the level of uncertainty about product 1 increases. Hence, consumer's affordability to search for information about product 2 reduces slightly given the constraints of searching costs.

Figures 16 and 17 present the histograms of the frequency of paths based on the number of information searched about each product by the consumer in each scenario. Over 70% and over 80% of the simulated paths do not search for any information on product 1 in the scenarios of unbiased but uncertain belief about product 1 and biased but certain belief about product 1, respectively. This result may seem at odds with the average number of information acquired about product 1, which is greater in the scenario of biased but certain belief than in the scenario of unbiased but uncertain belief. Again, this

may be explained in terms of the informational signal. In the scenario of unbiased but uncertain belief about product 1, consumers may be motivated to search for information about product 1 due to their initial uncertainty. However, after acquiring information, consumers confirm that their initial belief is close to the true match value by observing the informational signal and hence, are less motivated to continue searching for additional information on product 1. In contrast, in the other scenario, consumers may be less motivated to search for information at the beginning as they are certain about their incorrect belief on the match value. However, after acquiring information, their motivation to continue searching increases by observing the informational signal that their initial match value belief may be wrong. Hence, consumers tend to search for more information about product 1 once they have started to search for this information. In fact, the maximum number of searched information about product 1 is 35 in the scenario of biased but certain belief about product 1 while it is 14 in the scenario of unbiased but uncertain belief about product 1. In addition, as shown in Figures 16 and 17, more than 70% of the simulated paths search for information about product 2 at least once, similar to the scenario of full certainty about product 1.

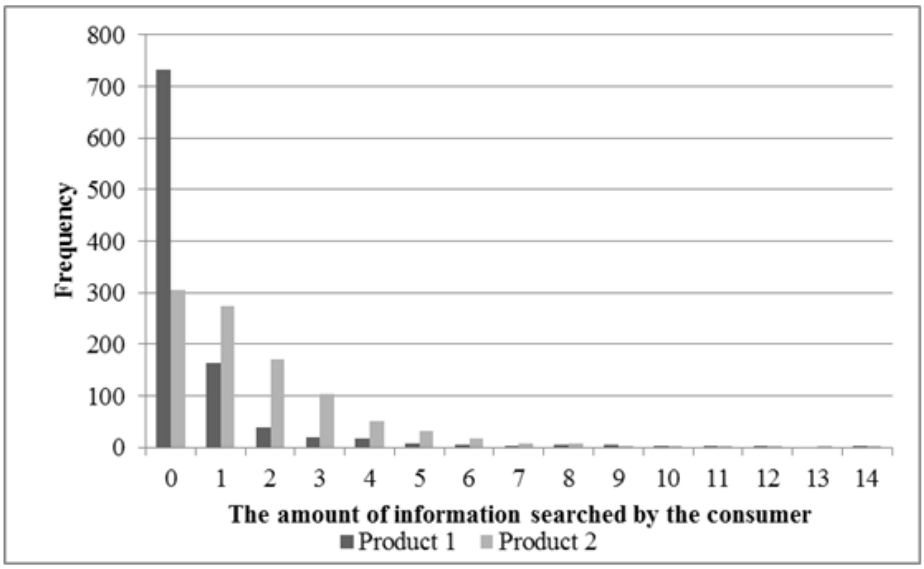

Figure 16. Histogram of the amount of information searched for each product in the scenario of unbiased but uncertain belief about product 1

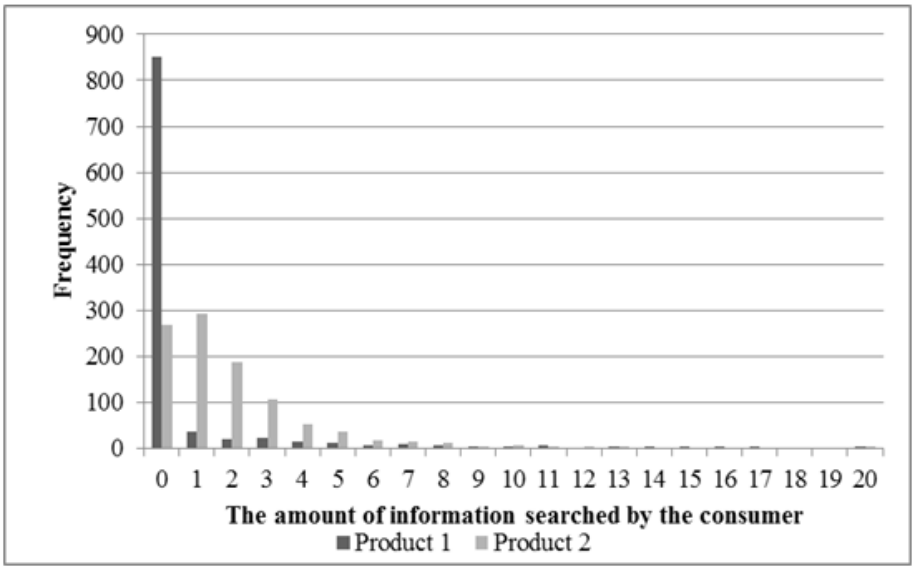

Figure 17. Histogram of the amount of information searched for each product in the scenario of biased but certain belief about product 1

Figures 18 and 19 present the histograms of the final purchase decisions in about 1,000 simulated paths for each scenario respectively. Similar to the scenario of full certainty about product 1, the proportion of purchase is significantly higher for product 1 than for product 2 in the scenario of unbiased but uncertain belief about product 1. Hence, although consumers are just as uncertain about their initial belief on the match value with regards to product 1 as they are with regards to product 2, they are more likely to purchase product 1 than product 2 because the initial expectation of product 1 is greater than that of product 2. On the other hand, Figure 19 shows that the proportion of purchase is significantly lower for product 1 than for product 2 in the scenario of biased but certain belief about product 1. In this scenario, consumers are certain of the match value of product 1 at the beginning. As this means that they are less motivated to search for information on product 1, they do not have sufficient opportunity to learn its true match value right up to the final decision period. For this reason, even though the match values of both products are the same, consumers tend to perceive the match value of product 2 to be greater than that of product 1. Hence, the probability to purchase product 2 is higher than the probability to purchase product 1.

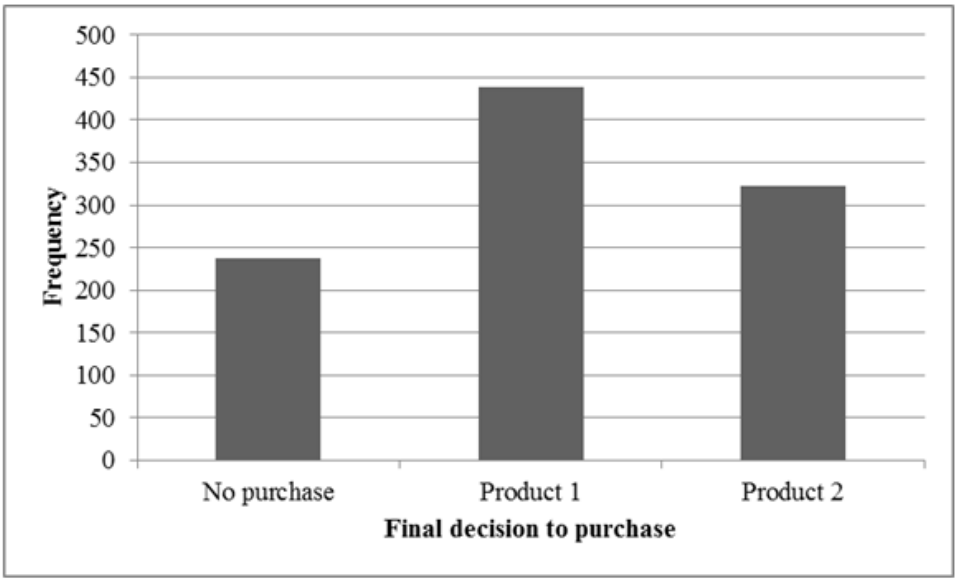

Figure 18. Histogram of the final purchase decision in the scenario of unbiased but uncertain belief about product 1

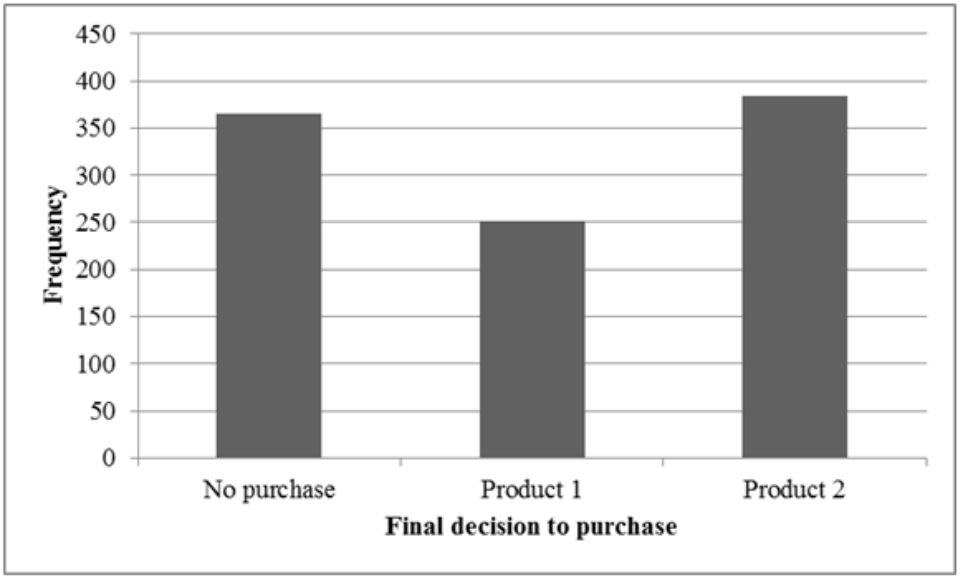

Figure 19. Histogram of the final purchase decision in the scenario of biased but certain belief about product 1

Finally, this study examines scenarios whereby the consumer has uncertain utilities for both products. From the discussion of the simulation results in earlier scenarios, consumer behavior varies greatly depending on the difference between the mean of the initial belief on the match value and the actual match value. Hence, this study focuses on describing the change in consumer behavior resulting from a change in the match value of one product. To capture the general case, this experiment assumes that products 1 and 2 have different match values (i.e., scenario of different match values with uncertainty about both products). Table 5 shows the assumed parameter values in the baseline scenario. According to Table 5, the parameter values of both products are identical except for the match value, which is greater for product 2 than for product 1. For simplicity, the observed parts of utility for both products are assumed to be zero and the search cost is set to be 0.5. In order to examine the dynamics of search based on the difference between initial expectation and true match value, this study compares the results of the baseline scenario with other scenarios that change the mean initial belief about product 1 similar to the first experiment in Section 4.2.

Table 5. Parameter values in the baseline scenario of different match values with uncertainty about both products

Parameter	Product 1	Product 2
Q_{Ij}	2.5	3.5
μ_{Ij0}	1.5	1.5
σ^2_{Ij0}	1.0	1.0
$\sigma^2_{S_{Ij}}$	1.0	1.0

In the baseline scenario, the average total number of periods that the consumer searches for information over 1,000 paths is 4.561. In addition, the averages of the number of acquired information about products 1 and 2 are 1.844 and 2.717, respectively. The histogram showing the frequency of paths based on the number of information searched by the consumer for each product is presented in Figure 20. As shown in this figure, more than 50% of the simulated paths do not search for any information about product 1 while more than 80% of the simulated paths search for information about product 2 at least once. This may be interpreted as consumers have lower motivation to search for information about product 1 than for product 2 as the difference between the initial expectation and the true match value is smaller in the case of product 1 compared to product 2. Hence, the benefit from searching for additional information about product 1 is smaller than the benefit from searching for information about product 2, as discussed in Section 4.2.

Figure 21 shows the histogram of the final purchase decision in 1,000 simulated paths. Consistent with general theory, the proportion of purchase is greater for product 2 than for product 1 because the true match value of product 2 is greater than that of product 1.

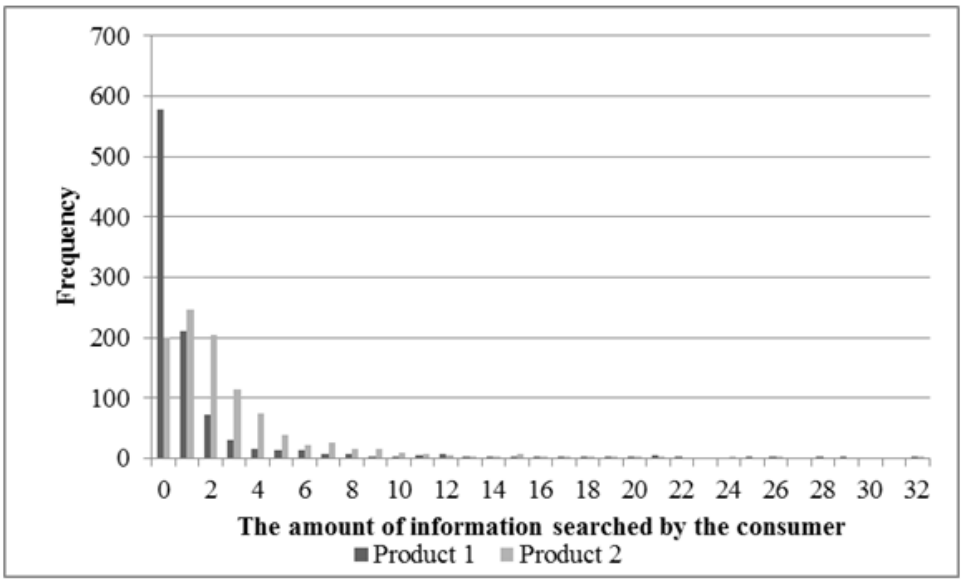

Figure 20. Histogram of the amount of information searched for each product in the baseline scenario of different match values with uncertainty for both products

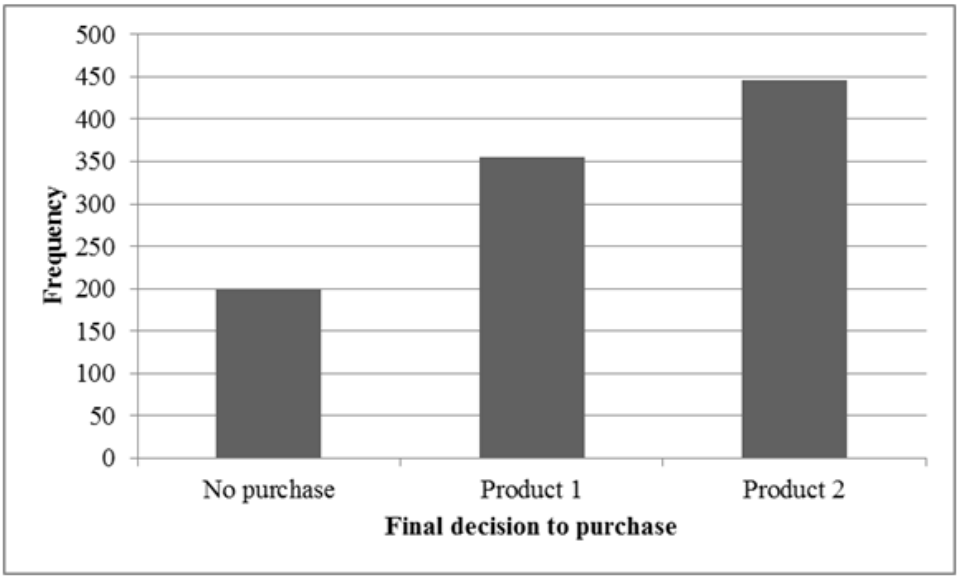

Figure 21. Histogram of the final purchase decision for each product in the baseline scenario of different match values with uncertainty for both products

Now, this study considers the change in the average number of information searched depending on a change in the mean of the initial match value belief of product 1. Figure 22 shows the simulation results.

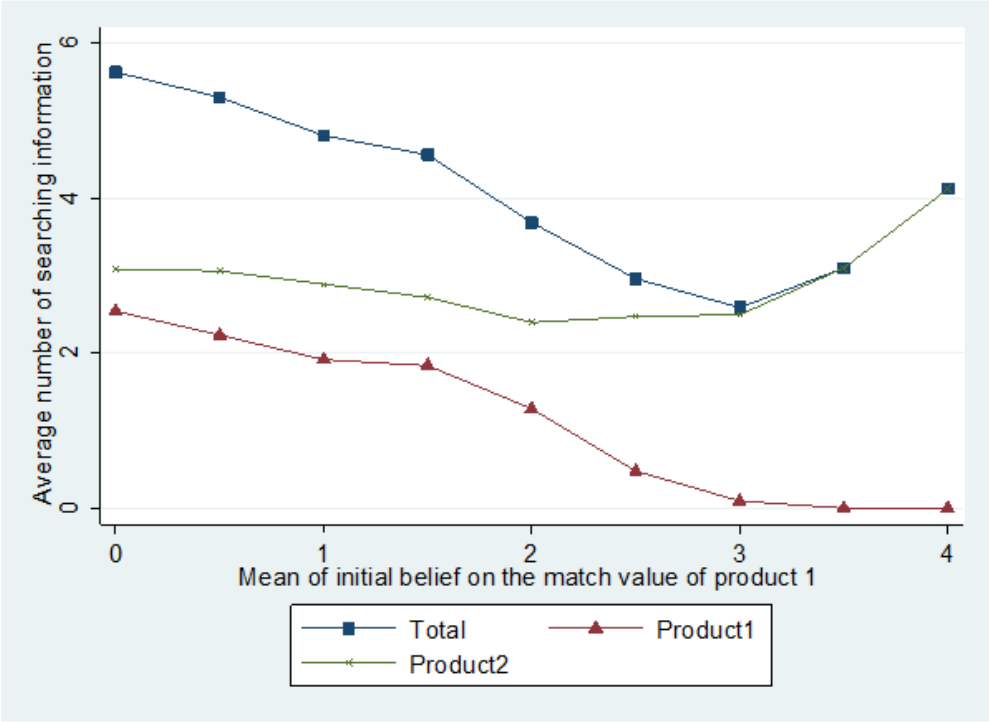

Figure 22. Changes in the average amount of information searched depending on the changes in the mean of the initial match value belief of product 1

Regarding product 1, consumers are likely to be less motivated to search for information about the product as the mean of the initial belief on the match value of product 1 increases from 0 to 2.5. This result may be explained by the benefit from searching for additional information, which is smaller when the mean of the initial match

value belief is larger as discussed in Section 4.2. In addition, when the initial expectation of product 1 exceeds its true match value (2.5), consumers' motivation to search for information is reduced dramatically as searching for information may decrease the expected utility of the product, which is contrary to utility maximization. Specifically, the average number of acquired information on product 1 is close to zero as the mean initial belief on the match value of product 1 exceeds 3.5, which is the true match value of product 2.

Regarding product 2, the average number of acquired information about this product seems relatively stable when the initial perception of the match value of product 1 changes from 0 to 2.5. However, consumers' motivation to search for information on product 2 tends to increase as that for product 1 decreases. Especially, consumers tend to search for more information about product 2 if the initial expectation about product 1 exceeds the true match value of product 2. As consumers are unlikely to search for information on product 1 in this case, they may focus on searching for information about product 2. In addition, consumers need to acquire more information about product 2 in order to evaluate products accurately as the information signal for product 2 hints that the actual match value of product 2 may be similar to consumers' perception about the match value of product 1.

Figure 23 shows the change in the frequency of final decision among 1,000 paths depending on the change in initial expectation of product 1. As the initial expectation of the match value of product 1 increases, the share of purchases of product 1 increases and decreases for "no purchase" and purchasing product 2. Specifically, the ratio of purchasing product 1 increases dramatically as the initial expectation about product 1

exceeds the mean of the initial match value belief of product 2. Especially in the case where the initial expectation about product 1 exceeds the actual match value of product 2, almost 50% of the paths decide to purchase product 1. However, although the proportion of purchase for product 2 is slightly lower as the initial expectation about the product 1 increases, it remains over 40% of the total paths, which cannot be explained by the standard logit probability. Therefore, Figure 23 proves that consumers could make the right purchase decision by searching for information even if their initial perceptions are severely incorrect.

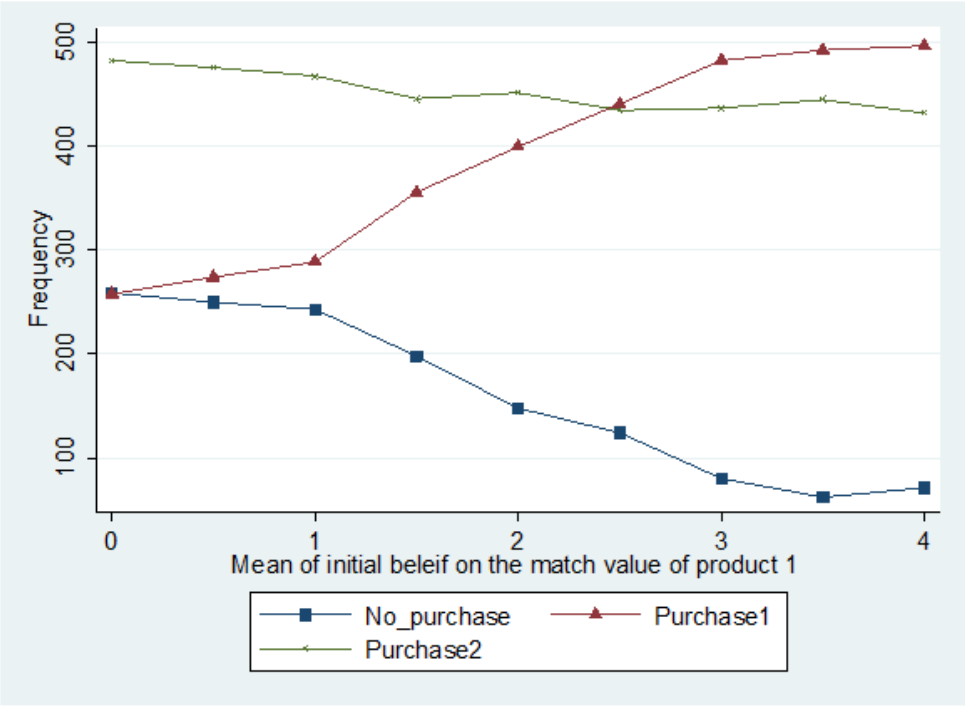

Figure 23. Changes in the frequency of final purchase decisions depending on the changes in the mean of the initial match value belief of product 1

4.4 Monte Carlo Experiment 3: Heterogeneity in Match Values

It is more realistic that consumers have heterogeneous rather than homogeneous match values. As discussed in the earlier section, consumers' behavior are severely dispersed in this case compared to the assumption of homogeneity in match values as the differences between the initial expectation and the actual match values vary by consumer. To examine the dispersion of consumer behavior resulting from heterogeneous match values, this study compares the case where consumers have homogeneous match values of two products (i.e., scenario of homogeneity) to the case where consumers have heterogeneous match values for one product and a homogeneous match value for the other product (i.e., scenario of heterogeneity). Table 6 shows the assumed parameter values for the two scenarios. The second and third columns show the parameter values for the scenarios of homogeneity and heterogeneity, respectively.

Table 6. Parameter values for scenarios of homogeneity and heterogeneity of product 1

Parameter	Product 1		Product 2
	Homogeneity	Heterogeneity	
$\mathbf{X}_{ij}'\mathbf{b}$	1.0	1.0	2.0
Q_{ij}	2.0	$N\left(2.0, 0.5^2\right)$	2.0
μ_{ij0}	1.0	1.0	1.0
σ_{ij0}^2	1.0	1.0	1.0
$\sigma_{S_{ij}}^2$	1.0	1.0	1.0

As shown in Table 6, the match value for each consumer is generated from a normal distribution in the scenario of heterogeneity. After simulating the decisions 10 paths per consumer for 1,000 consumers for each scenario, the averages of the total number of periods that consumers search for information are 5.61 and 6.03 in the two scenarios, respectively. In addition, while the average number of acquired information about product 2 is of similar value in both scenarios (2.38 for homogeneity scenario and 2.43 for heterogeneity scenario), the average number of acquired information about product 1 shows a relatively greater difference between the two cases (3.23 for homogeneity scenario and 3.60 for heterogeneity scenario). These results are predictable because the assumption of heterogeneity in match values may allow the difference between the initial expectation and the match value to be in the region that provides greater motivation to consumers to search for information.

Figures 24 and 25 present the histograms of the number of information searched on each product by the consumers for both scenarios. Regarding product 1, over 50% of the consumers do not search for any information about the product in the scenario of homogeneity while over 85% of the consumers search for information about the product at least once in the scenario of heterogeneity. In addition, approximately 40% of all consumers search for information on product 1 once or twice in the scenario of heterogeneity. On the other hand, over 70% of the consumers search for information about product 2 at least once in the scenario of homogeneity while almost all consumers do so in the scenario of heterogeneity.

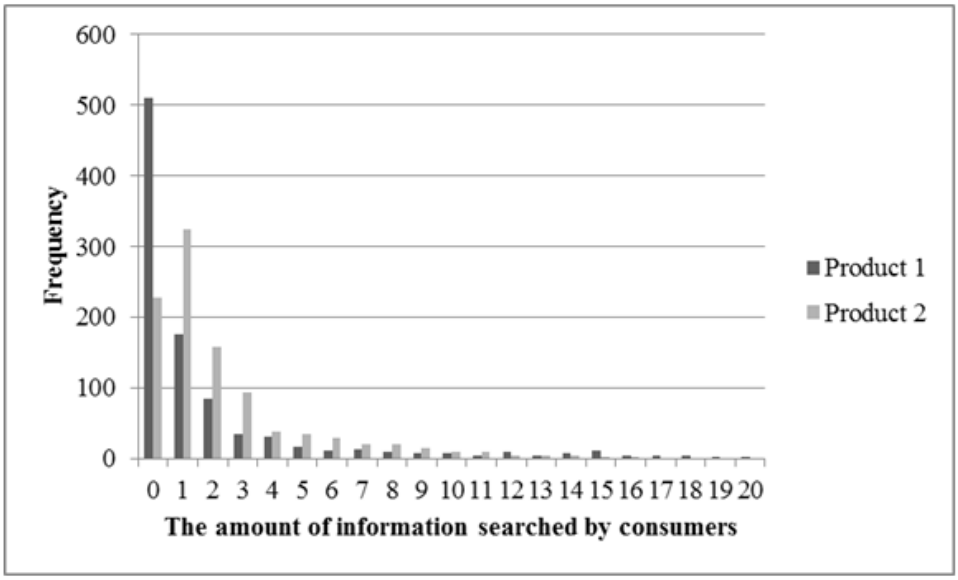

Figure 24. Histogram of the amount of information searched for each product in the scenario of homogeneity

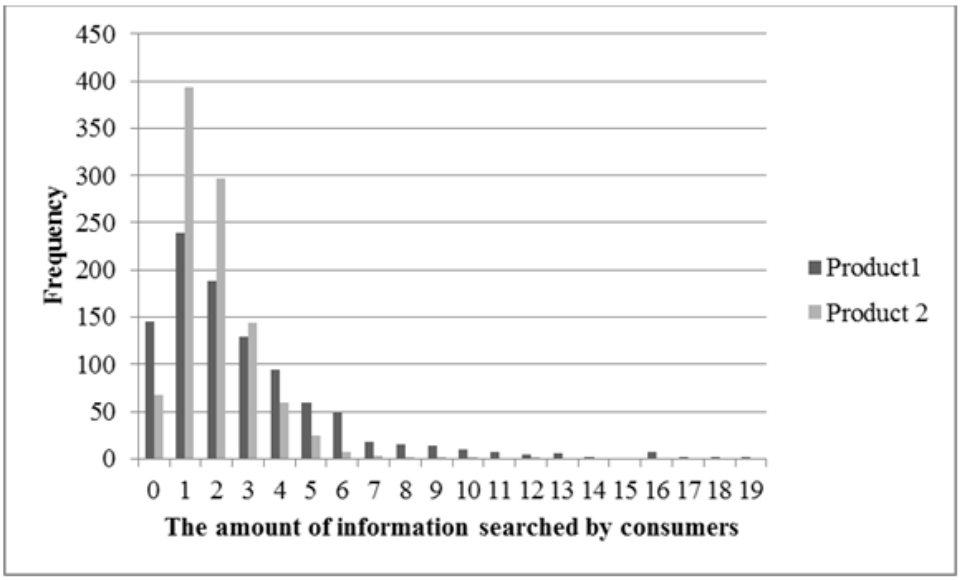

Figure 25. Histogram of the amount of information searched for each product in the scenario of heterogeneity

As discussed earlier, the change in the parameters of individual products may have an effect on consumers' decisions regarding other products. Figures 24 and 25 prove that the heterogeneity assumption of product 1 induces changes in consumers' behavior regarding not only product 1, but also product 2. Especially, consumers' behavior to search for information tends to be more concentrated on product 2 while product 1 shows more dispersed behavior by consumers in the scenario of heterogeneity. On the other hand, the final purchase decisions tend to be similar for both scenarios as shown in Figure 26. This implies that the assumption of heterogeneity has a greater effect of dispersion on the information search behavior than on the final purchase decision.

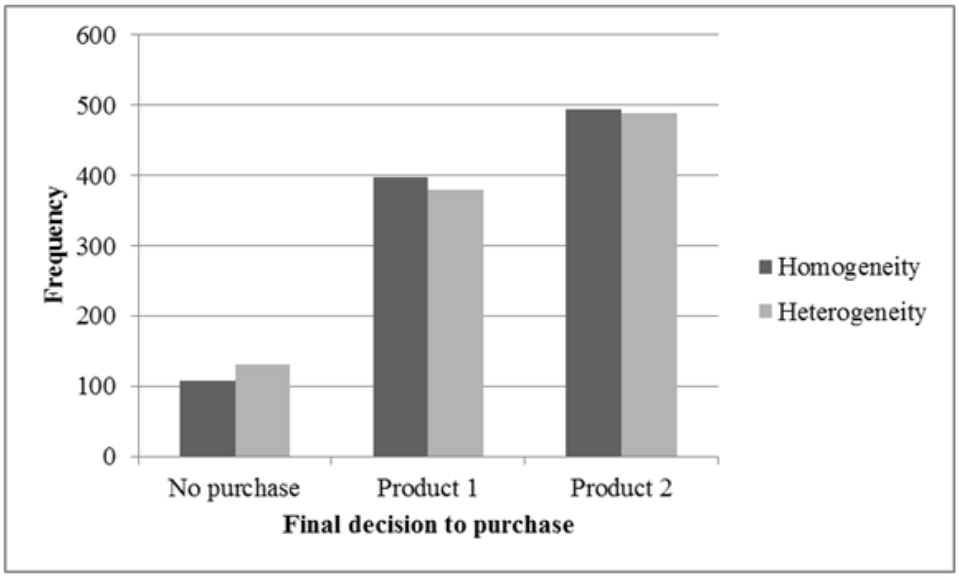

Figure 26. Histogram of the final purchase decisions in the homogeneity and heterogeneity scenarios

To illustrate the dispersion of consumers' information search behavior more

specifically, this study selects two consumers who make the same final purchase decision but have different match values of product 1. Consumer #15 and consumer #800 choose to purchase product 2 at the final period of information search. However, the match value assessment about product 1 by consumer #15 is 2.627 while that by consumer #800 is 1.709. In addition, as shown in Figures 27 and 28, their paths of searching for information are different: consumer #15 searches for information about product 1 five times and about product 2 just once while consumer #800 searches for information about products 1 and 2 only once and twice, respectively. Understandably, at the final time period, the means of the belief on match values for both products vary by consumers despite them making the same purchase decision. Hence, ignoring the heterogeneity in match values may not be able to explain the dispersion of consumers' information search behavior correctly.

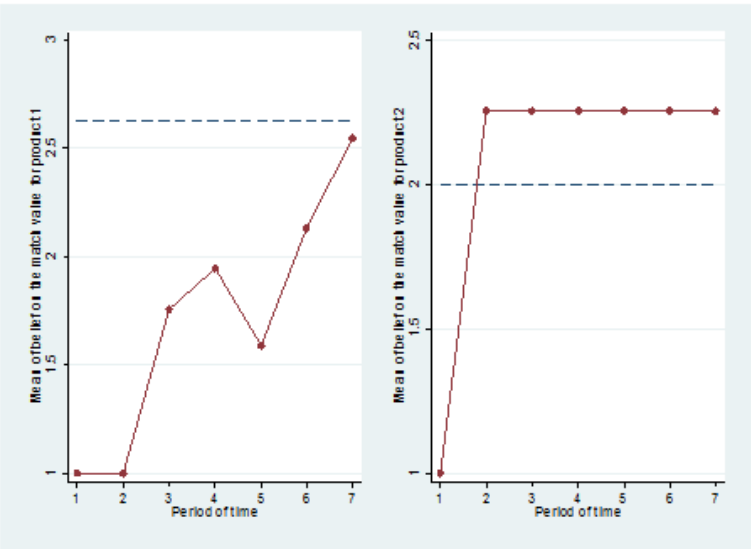

Figure 27. Updating process of the mean perception of match value for consumer #15

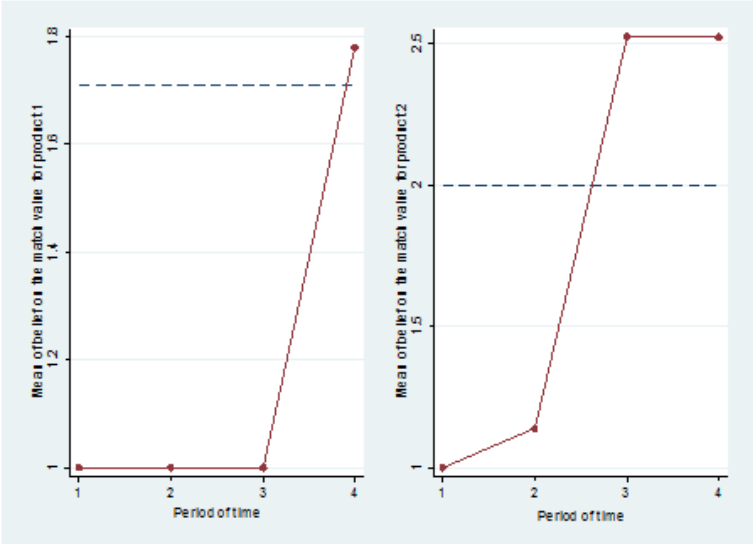

Figure 28. Updating process of the mean perception of match value for consumer #800

4.5 Summary and Discussion

This chapter examines the dynamics of consumer behavior explained by the proposed model by conducting three Monte Carlo experiments. The first experiment confirms the direct effects of the parameters suggested in the proposed model on consumers' decision-making process. Towards this objective, this study simulates the decision making process of a single consumer about a single product and compares the results across different combinations of parameter values. As the search cost increases, consumers are likely to be less motivated to search for information about the product as searching for information requires increased efforts. On the other hand, as the variance of the initial match value belief increases, the consumer's motivation to search for product information increases as the initial scale of uncertainty perceived by the consumer increases. Conversely,

consumers are less likely to search for information as the variance of the informational signal increases. This may be explained as the increased variance of the informational signal based on the match value implies a less accurate information source. Hence, if the informational signal provides incorrect information about the match value, consumers tend to refrain from searching for information from that source of information. Lastly, the mean of initial match value belief has an effect on consumers' behavior by substituting the difference from the true match value. For example, if consumers have similar mean perceptions of the true match value, they are not likely to search for information due to the small benefit from searching. In addition, if consumers overestimate the match value initially, they tend to not search for information as updating the match value by acquiring information may reduce the expected utility of the product. Finally, if consumers have opposing opinions of the match value belief compared to its true value, they are not likely to search for information due to their resistance to the disconfirmation of their initial perception.

The second experiment examines the decision-making behavior of a single consumer across two products. Since the uncertainty perceived by consumers is a major motivation for information search, this study attempts to confirm various scenarios related to the uncertainty. The parameters related to this uncertainty include the variance of the initial match value belief, mean of the initial match value belief, and the true match value. Consumers are unlikely to search for information on a product if they are certain of their initial match value belief of the product (i.e., zero variance of the initial match value belief). However, if they doubt their initial mean perception, they are likely to search for more information after acquiring information once. In addition, this study examines the

effect of the difference between the mean perception of match value of a specific product and its true match value on the consumer's decisions, specifically, regarding other products. From the results, if consumers have mean perceptions that are similar to the true match value of the product at the initial period, they are more motivated to search for information on other products. In addition, if consumers overestimate the match value of a certain product, they may purchase that product even if there are other superior products available. However, consumers still have the opportunity to make the right decision by learning from acquired information.

Lastly, the third experiment examines the dispersion of consumers' behavior caused by heterogeneous match values. From the results, the assumption of heterogeneity in match values has a greater effect on consumers' decision to search for information than to purchase the product. This suggests that ignoring this heterogeneity may result in an incorrect explanation of the dispersion in consumer behavior thereby causing inefficient model estimates.

Chapter 5. Estimation

This chapter proposes the model specification for structural estimation and explains the Bayesian inference method applied to estimate the proposed dynamic discrete choice model. Specifically, this study adopts the modified Metropolis-Hastings algorithm to avoid the computational burden of solving dynamic programming problems that are required to construct the likelihood function.

5.1 Model Specification for Structural Estimation

For the structural estimation, this study considers the simplest case whereby consumers make a purchase decision about a newly introduced product *A*. In this problem context, consumers are assumed to search for information only on product *A*. Then the utility perceived by a consumer *i* from purchasing the product *A* may be defined as follows:

$$U_{iAt} = \mathbf{X}'_{iA}\mathbf{b} + \tilde{Q}_{iA,t-1} + \varepsilon_{iAt} \quad \cdots\cdots \text{Eq. (55)}$$

where $\mathbf{X}_{iA}$ is a k dimensional vector containing attributes that are observable by researchers and $\mathbf{b}$ is the corresponding coefficient vector. ε_{iAt} is the random component of utility with zero mean, which remains unobserved by researchers, and $\tilde{Q}_{iA,t-1}$ is consumer i's match value belief of product A, whereby consumers are

uncertain of the match value due to imperfectly observed product attributes prior to product purchase. As discussed in Chapter 3, consumers' match value beliefs vary according to the time spent acquiring product information. Therefore, the match value belief is a time-varying variable. In addition, since the belief at time t prior to information search $\tilde{Q}_{iA,t-1}$, is stochastic for consumers, consumer decisions are made utilizing the expected value of the utility at current time point t based on the set of information acquired up to the previous period $\mathbf{I_{i,t-1}}$, similar to Equation 13. In addition, the expected utility associated with the "no purchase" choice is specified in the same manner as Equation 15.

Similar to Equation 16, S_{iAt}, the informational signal about product A received at period t by consumer i, is specified as the noisy measurement of the true match value and is normally distributed with mean Q_{iA} and variance $\sigma^2_{S_{iA}}$. Assuming that the initial prior match value belief $\tilde{Q}_{iA0}$ follows a normal distribution with mean μ_{iA0} and variance σ^2_{iA0}, the posterior match value belief at period t, $\tilde{Q}_{iAt}$, is derived in a similar manner to Equations 23 to 25, as follows:

$$\tilde{Q}_{iAt} \sim N\left(\mu_{iAt}, \sigma^2_{iAt}\right) \quad \text{Eq. (56)}$$

where $$\mu_{iAt} = \frac{\sigma^2_{iAt}}{\sigma^2_{iA0}}\mu_{iA0} + \frac{\sigma^2_{iAt}}{\sigma^2_{S_{iA}}}\sum_{\tau=1}^{t} S_{iA\tau} \quad \text{Eq.(57)}$$

and $$\frac{1}{\sigma^2_{iAt}} = \frac{1}{\sigma^2_{iA0}} + \frac{t}{\sigma^2_{S_{iA}}} \quad \text{Eq. (58)}$$

To facilitate the identification of parameters and to provide a more comprehensive

interpretation, the Bayesian learning process could be reparameterized by defining two new variables: perception bias and signal noise (Erdem & Keane, 1996; Shin et al., 2012), as introduced in Section 3.5. Consistent with Section 3.5, perception bias is defined as the difference between the mean of the match value belief by the consumer and the true match value, i.e., $\nu_{iAt} = \mu_{iAt} - Q_{iA}$. On the other hand, signal noise is defined as the difference between the informational signal and the true match value, i.e., $\eta_{iAt} = S_{iAt} - Q_{iA}$. Substituting both these variables and Equation 58 into Equation 57, the mean perception of the match value belief after receiving product information at period t may be derived as follows:

$$\mu_{iAt} = Q_{iA} + \frac{\sigma^2_{S_{iA}}}{\sigma^2_{S_{iA}} + t\sigma^2_{iA0}} \nu_{iA0} + \frac{\sigma^2_{iA0}}{\sigma^2_{S_{iA}} + t\sigma^2_{iA0}} \sum_{\tau=1}^{t} \eta_{iA\tau} \quad \cdots\cdots\text{Eq. (59)}$$

Then, the expected utility obtained by consumer i from product A at period t is specified as per Equation 60:

$$U^E_{iAt} = \mathbf{X}'_{iA}\mathbf{b} + Q_{iA} + \frac{\sigma^2_{S_{iA}}}{\sigma^2_{S_{iA}} + (t-1)\sigma^2_{iA0}} \nu_{iA0} + \frac{\sigma^2_{iA0}}{\sigma^2_{S_{iA}} + (t-1)\sigma^2_{iA0}} \sum_{\tau=1}^{t-1} \eta_{iA\tau} + \varepsilon_{iAt} \quad \cdots\cdots\text{Eq. (60)}$$

At each time point, consumers decide between searching for information on product A or making a purchase decision without further information search, based on the trade-offs between the cost of search and the benefit from searching. In the binary choice situation, the reward received by consumer i for making a purchase decision and terminating

information search at period t is denoted as $R_{it} = \max\left[U_{iot}^{E}, U_{iAt}^{E}\right]$, and the expected reward of the next period is specified as $ER_{i,t+1} = E_{\eta_{iAt},\mathbf{e_{t+1}}}\left[\max\left\{U_{io,t+1}^{E}, U_{iA,t+1}^{E}\right\} \mid \mathbf{I_{i,t-1}}\right]$. This is similar to the case considered in Section 4.1. Thus, the expected reward does not depend on the consumer's choice of the product selected for information search, as the information search is conducted only for product *A*.

Similar to Equation 53, the expected reward may be specified as follows:

$$
\begin{aligned}
ER_{i,t+1} &= E_{\eta_{iAt}}\left[E_{\mathbf{e_{t+1}}}\left[\max\left\{U_{iA,t+1}^{E}, U_{io,t+1}^{E}\right\}\right] \mid \mathbf{I_{i,t-1}}\right] \\
&= E_{\eta_{iAt}}\left[\ln\left(1+\exp\left(\begin{array}{l}\dfrac{\sigma_{S_{iA}}^{2}}{\sigma_{S_{iA}}^{2}+t\sigma_{iA0}^{2}}\mu_{iA0}+\dfrac{t\sigma_{iA0}^{2}}{\sigma_{S_{iA}}^{2}+t\sigma_{iA0}^{2}}Q_{iA} \\ +\dfrac{\sigma_{iA0}^{2}}{\sigma_{S_{iA}}^{2}+t\sigma_{iA0}^{2}}\sum_{\tau=1}^{t-1}\eta_{iA\tau}+\dfrac{\sigma_{iA0}^{2}}{\sigma_{S_{iA}}^{2}+t\sigma_{iA0}^{2}}\eta_{iAt}\end{array}\right)\right) \mid \mathbf{I_{i,t-1}}\right] \quad \cdots\text{Eq. (61)}
\end{aligned}
$$

As shown in Equation 61, η_{iAt}, the signal noise at time t, is the random variable prior to information search. Based on the above specification, consumer i decides to search for additional information if $ER_{i,t+1} - R_t \geq c_i$, or to terminate the information search process and make a purchase decision if $ER_{i,t+1} - R_t < c_i$.

In the binary choice situation, researchers may observe consumers' decisions on whether to search for information, which is specified as $\left\{\left\{d_{it}\right\}_{t=1}^{T_i}\right\}_{i=1}^{I}$, and their final purchase decision, specified by the variable $\left\{Y_i\right\}_{i=1}^{I}$.

At period *t*, the probability that a consumer *i* searches for additional information is

specified as follows:

$$
\begin{aligned}
&\Pr\left(d_{it}=1\right)\\
&\Pr\left(ER_{i,t+1}-R_{it}\geq c_i\right)\\
&=\Pr\left(ER_{i,t+1}-c_i\geq U_{iAt}^{E}\mid R_{it}=U_{iAt}^{E}\right)\Pr\left(R_{it}=U_{iAt}^{E}\right) \cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots \text{Eq. (62)}\\
&+\Pr\left(ER_{i,t+1}-c_i\geq \varepsilon_{iot}^{E}\mid R_{it}=U_{iot}^{E}\right)\Pr\left(R_{it}=U_{iot}^{E}\right)\\
&=e^{-e^{-ER_{i,t+1}+c_i+\mathbf{X}_{iA}'\mathbf{b}+\mu_{iA,t-1}}}\frac{\mathrm{e}^{\mathbf{X}_{iA}'\mathbf{b}+\mu_{iA,t-1}}}{1+\mathrm{e}^{\mathbf{X}_{iA}'\mathbf{b}+\mu_{iA,t-1}}}+\frac{e^{e^{-e^{-ER_{i,t+1}+c_i}}}}{1+\mathrm{e}^{\mathbf{X}_{iA}'\mathbf{b}+\mu_{iA,t-1}}}
\end{aligned}
$$

At T_i consumer i terminates the information search. Since researchers can observe the final purchase decision by consumers, the probability of $d_{iT_i}=0$ depends on Y_i. The probability that a consumer i decides to purchase product *A* is specified by the logit probability, $\Pr\left(Y_i=1\right)=\dfrac{\exp\left(\mathbf{X}_{iA}'\mathbf{b}+\mu_{iA,T_i-1}\right)}{1+\exp\left(\mathbf{X}_{iA}'\mathbf{b}+\mu_{iA,T_i-1}\right)}$, and the probability that the consumer decides not to buy the product is described as $\Pr\left(Y_i=0\right)=\dfrac{1}{1+\exp\left(\mathbf{X}_{iA}'\mathbf{b}+\mu_{iA,T_i-1}\right)}$.

Therefore, the probability to terminate the information search process may be defined as follows:

$$
\begin{aligned}
\Pr\left(d_{iT_i}=0\right) &= \Pr\left(ER_{i,T_i+1}-R_{iT_i}<c_i\right) \\
&= \Pr\left(ER_{i,T_i+1}-c_i<U^{E}_{iAT_i}\mid Y_i=1\right)I\left(Y_i=1\right) \\
&+\Pr\left(ER_{i,T_i+1}-c_i<\varepsilon_{ioT_i}\mid Y_i=0\right)I\left(Y_i=0\right) \qquad \cdots\cdots \text{ Eq. (63)} \\
&= \left(1-\exp\left(-\exp\left(-ER_{i,T_i+1}+c_i+\mathbf{X}'_{iA}\mathbf{b}+\mu_{iA,T_i-1}\right)\right)\right)\cdot I\left(Y_i=1\right) \\
&+\left(1-\exp\left(-\exp\left(-ER_{i,T_i+1}+c_i\right)\right)\right)\cdot I\left(Y_i=0\right)
\end{aligned}
$$

Based on the set of observed data $\left\{\left\{d_{it}\right\}_{t=1}^{T_i},Y_i\right\}_{i=1}^{I}$, the likelihood increment of a consumer i is specified as follows:

$$
\begin{aligned}
&L_i\left(\left\{d_{it}\right\}_{t=1}^{T_i},Y_i\mid\mathbf{X}_{iA};\mathbf{b},Q_{iA},\sigma^2_{iA0},\sigma^2_{S_{iA}},\nu_{iA0},c_i,\left\{\eta_{iAt}\right\}_{t=1}^{T_i-1}\right) \\
&=\prod_{t=1}^{T_i}\begin{pmatrix}\Pr\left(d_{it}=1\right)\cdot I\left(d_{it}=1\right) \\ +\Pr\left(d_{it}=0\right)\cdot\Pr\left(Y_i=1\right)\cdot I\left(d_{it}=0\right)\cdot I\left(Y_i=1\right) \\ +\Pr\left(d_{it}=0\right)\cdot\Pr\left(Y_i=0\right)\cdot I\left(d_{it}=0\right)\cdot I\left(Y_i=0\right)\end{pmatrix}\times\prod_{t=1}^{T_i-1}f\left(\eta_{iAt}\right) \qquad \cdots\cdots\cdots\text{Eq. (64)}
\end{aligned}
$$

Then, the likelihood across all consumers is computed based on Equation 46.

5.2 Overview of Bayesian MCMC Method

In the Bayesian approach, the parameters to be estimated are drawn from the posterior distribution of parameters, $\pi(\theta)=p(\theta\mid x)$, which is proportional to the multiplication of the prior distribution of parameters, $p(\theta)$, and the likelihood function derived from the observed data x, $l(\theta)=f(x\mid\theta)$, based on Bayes' theorem (Gamerman and Lopes, 2006). The Markov Chain Monte Carlo (MCMC) method is generally used in the

Bayesian approach, to draw from the posterior distribution.

The Gibbs sampling algorithm is the most prominent MCMC method in Bayesian literature, which allows researchers to draw a parameter successively from the full conditional distribution by fixing the values of other parameters.

However, if the full conditional distributions have a complex and awkward form making it difficult to draw parameters directly from the distribution, the implementation of the Metropolis-Hastings (MH) algorithm is more efficient in terms of computational time and much easier than Gibbs sampling (Gamerman and Lopes, 2006; Train, 2009).

The basic process of the MH algorithm is as follows:

Step 1) Start the iteration ($m = 1$) with an arbitrary initial value $\theta^{(0)}$.

Step 2) Move the chain with a new parameter value ϕ, which is generated from the proposal density $q\left(\theta^{(m-1)}, \phi\right)$.

Step 3) Calculate the acceptance probability of ϕ moved from $\theta^{(m-1)}$, denoted as $\alpha\left(\theta^{(m-1)}, \phi\right)$, based on the following equation:

$$\alpha\left(\theta^{(m-1)}, \phi\right) = \min\left\{1, \frac{\pi(\phi) q\left(\phi, \theta^{(m-1)}\right)}{\pi\left(\theta^{(m-1)}\right) q\left(\theta^{(m-1)}, \phi\right)}\right\} \quad \text{Eq. (65)}$$

where the posterior distribution $\pi(\theta) = l(\theta) p(\theta)$ is computed in the Bayesian scheme. If the proposal density is symmetric, i.e.,

$q\left(\theta^{(m-1)},\phi\right)=q\left(\phi,\theta^{(m-1)}\right)$, then the acceptance probability is reduced to $\min\left\{1,\pi(\phi)\big/\pi\left(\theta^{(m-1)}\right)\right\}$.

Step 4) Accept the move with a probability of $\alpha\left(\theta^{(m-1)},\phi\right)$ and set $\theta^{(m)}=\phi$, or reject the move with the probability of $1-\alpha\left(\theta^{(m-1)},\phi\right)$ and set $\theta^{(m)}=\theta^{(m-1)}$.

Step 5) Set $m=m+1$. Go to Step 2 and repeat until convergence is reached.

For the estimation of dynamic discrete choice models, researchers generally compute the expected value of maximum utility, $E_{\varepsilon',s'}\left[\max U(\varepsilon',s')\,|\,\varepsilon,s\right]$, where the set of random variables (ε,s) is with reference to the current period and (ε',s') is with reference to the next period, to calculate the likelihood function, which is required to derive the acceptance probability in the MH algorithm. In this study, the corresponding term is $ER_{i,t+1}=E_{\eta_{iAt},\mathbf{e_{t+1}}}\left[\max\left\{U^{E}_{io,t+1},U^{E}_{iA,t+1}\right\}|\,\mathbf{I_{i,t-1}}\right]$. The conventional way to obtain this is to solve the dynamic programming for all possible states and parameter values. In addition, the expectation taken with respect to the random variable is approximated by repeatedly taking the average over simulated random variables until convergence is reached. Therefore, two loops are required for the conventional Bayesian MCMC algorithm, one for the MH algorithm and the other for the dynamic programming solution (Imai et al., 2009).

As solving the dynamic programming by conducting these loops imposes substantial

computational burden, this study uses an approximating method for expected maximum utility, which allows the estimation of parameters without solving the dynamic programming. Two recent studies suggest the Bayesian estimation method for dynamic discrete choice models using the approximation of the expected maximum utility (Imai et al., 2009; Norets, 2009). Both studies calculate the expected value based on the kernel smoothing method, which is one of the methods to approximate the value in a non-parametric setting. Imai et al. (2009) suggest a modified MH algorithm applicable to general dynamic discrete choice models using the approximation method based on the Gaussian kernel with values from past iterations. On the other hand, Norets (2009) suggests the Gibbs sampling method focused on dynamic discrete choice models with a serially correlated unobserved state variable, using the approximation method based on the nearest neighbor kernel among the past iteration values.

5.3 Estimation Procedure

In this study, the random grid method (Rust, 1987) is applied to the Gaussian kernel to approximate the expected reward. Signal noise is an unobserved value, which should be imputed in the estimation algorithm. Hence, this study computes the average value of the expected reward with candidate parameters in past iterations and the signal noise at random grid point, then computes the weighted average of these values based on the Gaussian kernel function in order to increase the accuracy per iteration and to reduce the computational burden for reaching convergence.

At each iteration k in the MH algorithm, this study generates L random grid

points for signal noise: $\eta_{iAt}^{k,l}$, $l=1,\cdots,L$. The number of past iterations used for calculating the expected reward is assumed to be N. Let the candidate parameter at iteration m be specified as $\tilde{\theta}^m$. Then, the expected reward is calculated based on Equation 66, as follows:

$$E_{it}^m = \frac{\sum_{k=m-N}^{m-1}\left(\frac{1}{L}\sum_{l=1}^{L}\ln\left(1+\exp\left(U_{iA,t+1}^E \mid \eta_{iAt}^{k,l},\tilde{\theta}^k,\left\{\eta_{iA\tau}^k\right\}_{\tau=1}^{t-1}\right)\right)\right)\times K_h\left(\tilde{\theta}^m-\tilde{\theta}^k\right)\prod_{\tau=1}^{t-1}K_h\left(\eta_{iA\tau}^m-\eta_{iA\tau}^k\right)}{\sum_{k=m-N}^{m-1}K_h\left(\tilde{\theta}^m-\tilde{\theta}^k\right)\prod_{\tau=1}^{t-1}K_h\left(\eta_{iA\tau}^m-\eta_{iA\tau}^k\right)} \quad \text{Eq. (66)}$$

For simplicity, this study assumes that consumers have homogeneous preferences. Then, the following parameters of the proposed model would be identical for all consumers:

$$Q_{iA} = Q_A \quad \text{Eq. (67)}$$
$$\nu_{iA0} = \nu_{A0} \quad \text{Eq. (68)}$$
$$\sigma_{S_{iA}}^2 = \sigma_S^2 \quad \text{Eq. (69)}$$
$$\sigma_{iA0}^2 = \sigma_{A0}^2 \quad \text{Eq. (70)}$$
$$c_i = c \quad \text{Eq. (71)}$$

In this assumption, σ_S^2 and σ_{A0}^2 may be unidentifiable separately as they are revealed as per the ratio in Equation 59. Moreover, ν_{A0}, which is related to σ_{A0}^2, is a parameter to be estimated and η_{iAt}, which is related to σ_S^2, is an unobserved variable to be imputed in the algorithm. Given this, Erdem and Keane (1996) fixed one of the

variances as one, in order to ensure individual identifiability. Hence, this study fixes the value of σ_S^2 as one and imputes the signal noises from a standard normal distribution.

The prior distributions of parameters are assumed to be diffused except for σ_{A0}^2. To ensure that $\sigma_{A0}^2 > 0$, this study transforms it as a precision ($\tau_{A0} = \frac{1}{\sigma_{A0}^2}$) and estimates τ_{A0}, instead, assuming a gamma prior distribution and a gamma proposal density function. Proposal densities for other parameters are assumed to follow a normal random walk.

Based on the expected reward as per Equation 66, the parameters are estimated by the MH algorithm. The estimation algorithm is developed in C99 and compiled by Intel C compiler.

5.4 Simulation Study

A Monte Carlo experiment is conducted to confirm the performance of the estimation algorithm in terms of its ability to recover true parameter values. To do this, artificial datasets are simulated for the homogeneous model.

Assuming consumers have homogeneous preferences, the true values of parameters are set as follow: $v_{A0} = -1.0$, $Q_A = 2.0$, $\sigma_{A0}^2 = 1.0$, and $c = 0.5$. Assuming that there are two observed attributes, x_{i1} and x_{i2}, and both are randomly generated from the uniform distribution $U[-1,1]$, the true values of corresponding parameters are set as follows: $\beta_1 = 1.0$ and $\beta_2 = -0.5$. As discussed in Section 5.3, σ_S^2 is fixed as 1.

Based on this approach, this study simulated the observed data for a sample size of 1,000 consumers. The procedure to simulate artificial data is similar to the procedure described in Section 4.1 Then, this study estimates the parameter values from the simulated artificial observed dataset. Set the number of the past iterations used for calculating the expected reward as 1,000 and the number of random grids as 100. After iterating 20,000 periods, this study reports the results of parameter estimation based on the 10,000th to 20,000th iterations.

Table 7 summarizes the estimation results and Figure 29 presents the MCMC plots of parameters from the 10,000th to 20,000th iterations. All parameters are significant at the 1% significance level.

Table 7. Estimation results of the proposed model

Parameter	True	Estimated value	
		mean	SD
v_{A0}	-1.0	-0.967	0.2349
Q_A	2.0	2.0227	0.2339
σ_{A0}^2	1.0	1.0939	0.3485
β_1	1.0	1.1533	0.0667
β_2	-0.5	-0.4770	0.0494
c	0.5	0.4975	0.0545

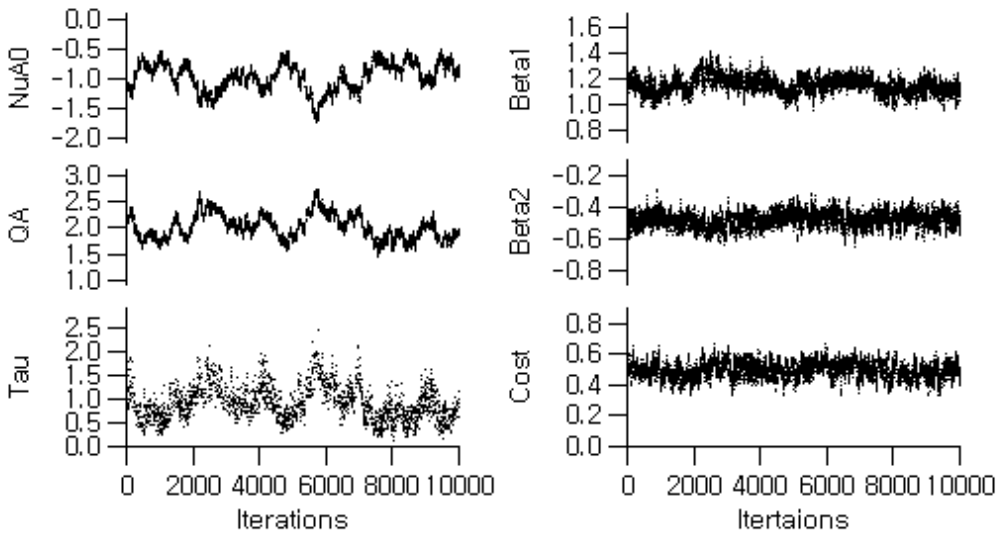

Figure 29. MCMC plots

Section 3.5 argues that the estimation results from the standard discrete choice model may lead to incorrect inferences due to misspecification of the uncertainty in the utility. To confirm this argument empirically, this study estimates the binary logit model based on the simulated data of consumers' final purchasing decision. Under the framework of the standard logit model, the utility obtained by consumer i from purchasing product A is specified as follows:

$$U_{iA} = \alpha_A + \beta_1 x_{i1} + \beta_2 x_{i2} + \varepsilon_{iA} \quad \cdots\cdots \text{Eq. (72)}$$

In Equation 72, the alternative specific constant (ASC), α_A, is included to capture the average effect of excluding factors from the observable attributes. Technically, ASC represents the mean of the unobserved part of utility from researchers' perspective. Hence,

the regression error term in Equation 72, ε_{iA}, should be defined as a random variable with mean zero for the identification. Table 8 documents the estimation results of the binary logit model.

Table 8. Estimation results of the binary logit model

Variable	Coefficient	Std.	t-value	p-value
α_A	2.1366***	0.1333	16.034	0.000
β_1	1.7144***	0.1525	11.242	0.000
β_2	-0.81145***	0.1229	-6.604	0.000

Note: *, **, and *** implies p < 0.1, p < 0.05, and p < 0.01 respectively.

As discussed in Section 3.5, the estimation results in Table 8 are biased against the true value of parameters. For deeper investigation, this study compares the root mean square errors (RMSE) of estimates to draw a direct comparison between the proposed model and the binary logit model, relative to the true parameter values. The results are presented in Table 9.

Table 9. Comparison of root mean square error of estimates

		β_1	β_2
True value		1.0	-0.5
The proposed model	Estimated value	1.1533	-0.477
	RMSE	0.1672	0.0545
Binary logit model	Estimated value	1.7144	-0.8115
	RMSE	0.7305	0.3349

Table 9 shows that the coefficient RMSEs of the binary logit model are more than four times greater than the corresponding RMSEs of the proposed model. Hence, the proposed model is proven more powerful in predicting consumer preferences accurately.

Under the assumption of homogenous preferences, the idea that the match value belief captures the average effect of the unobserved utility by the consumer may delude researchers into thinking that the match value in the proposed model corresponds to the ASC in the standard discrete choice model. As mentioned before, the ASC for the product captures the average effect of the part of utility unobserved by researchers in the standard discrete choice model. Since the regression error ε_{iA} is the only unobserved part of utility in the standard logit model, the ASC is interpreted as the mean of the regression error given that no other constant term is included in the utility function. On the other hand, the part of utility unobserved by researchers in the proposed model is the sum of the match value belief $\tilde{Q}_{iA,t-1}$ and the regression error. As this study normalizes the mean of the regression errors to zero for identification purposes and the match value belief is the error-compounded perception of the true match value Q_A, one may misunderstand that the mean of the unobserved part of utility in the proposed model is identical to the match value, i.e., $Q_A = ASC$ in the proposed model.

However, the unbiased mean of $\tilde{Q}_{iA,t-1}$ is μ_{iAt} and not Q_A, as discussed in Section 3.5. Hence, the mean of the unobserved part of utility may be specified as $Q_A + \nu_{A0}\sigma_{S_A}^2 \cdot E\left[\frac{1}{\sigma_{S_A}^2 + (t-1)\cdot\sigma_{A0}^2}\right]$, which is the component corresponding to the ASC. Notably, consumers have heterogeneous match value beliefs after acquiring information,

even though their initial match value belief and the true match value are assumed to be homogeneous. Therefore, $E\left[\frac{1}{\sigma_{S_A}^2+(t-1)\sigma_{A0}^2}\right]$ should be computed by taking the average over time and across consumers.

Lastly, one may opine that an additional constant term should be included in order to distinguish the average effect of unobserved utility from the researchers and consumers perspectives. However, while consumers could distinguish between them since they can observe the additional constant, researchers are unable to do so. Therefore, the additional constant and the mean perception of the match value cannot be identified separately from the researchers' perspective. Due to this reason, the proposed model does not contain an additional ACS.

Chapter 6. Empirical Study

This chapter validates the proposed model by applying it to several empirical studies. The information search and purchasing data used in this chapter is obtained by web-based choice experiments. The following sections present the design of the choice experiments and discuss the estimation results.

6.1 Design of Choice Experiment

To analyze consumers' information search behavior, this study designs a choice experiment that tracks consumers' decision-making process. The choice experiment comprises several stages, which are summarized as follows:

Stage1: Start the experiment by sharing basic information of the target product with respondents.

Stage2: Respondents make a choice decision between searching for additional information after waiting for 5 seconds and terminating the information search immediately. Go to Stage 3 if the decision is to search for additional information, otherwise go to Stage 4.

Stage3: Read additional reviews, which are obtained randomly, and go to the previous stage.

Stage4: Make a purchase decision.

For the choice experiment, this study chose two target products, Samsung's new smartwatch, "Gear 2," and Hyundai's new hybrid car, "Sonata Hybrid." Target products were chosen based on two criteria. The first criterion is that the product is introduced very recently at the point of the experiment, such that consumers have product uncertainty. This condition is necessary to ensure that respondents search for additional information prior to purchase. The second criterion is that a sufficient number of reviews of the target product must be available on the Internet. For the experiment to resemble a real search situation, it is critical that consumers decide whether or not to search for information based on the trade-offs between the costs and benefits of search. Hence, the experiment should provide as much information to the respondents as they desire. The second condition enables access to sufficient number of information to be shared with respondents in Stage 3.

Within the framework of this study, consumers are certain of the observed attributes of the product but are uncertain about the corresponding match value. To fulfill this assumption in choice experiments, the basic information of the target product is shared with respondents at the first stage. Figure 30 shows an example of the information shared at Stage 1 of the choice experiment.

다음은 현대 쏘나타 하이브리드 자동차 제품에 대한 질문입니다. 아래 설명문과 제품 기본정보를 보신 후 응답해 주십시오.

다음은 신제품에 대한 구매 여부를 결정하기 위해 관련 정보를 찾는 과정에 대한 질문입니다.

※ 지금부터 귀하는 기본적으로 제시되는 제품 정보(가격, 기능 등) 이외에 **추가적인 정보(다른 소비자의 구매 후기)를 원하는 만큼 보실 수 있습니다.** 단, 추가 정보를 한 번 얻기 위해서는 **약 10초 정도의 시간을 기다리셔야 합니다.**

① 귀하는 우선 제시된 제품의 **기본정보를 확인하신 후, 제품 구매의향에 대해 응답**해 주십시오.

② 제품 구매여부에 대한 최종 결정을 내리기 전에 추가정보를 보시려면 '추가정보 확인'을 선택해 추가정보를 확인합니다. **추가정보를 확인하신 후, 제품 구매의향에 대해 다시 응답**해 주십시오.

③ 정보탐색을 그만두고 현재의 의향대로 구매를 확정하기를 원하실 경우에는 '정보 확인 중단 및 구매확정'을 선택해 주십시오.

▣ 현대 쏘나타 하이브리드 자동차 제품 기본정보

구분	제품 기본정보
1. 이미지	
2. 차종	중형차
3. 연료	하이브리드 (가솔린+전기 모터)
4. 연비	17.7 ~ 18.2 km/l
5. 이산화탄소 배출량	91 ~ 94g/km
6. 제로백 (시속 0-100km 가속시간)	9.3초
7. 가격	2,829 ~ 3,139만원

Figure 30. Example of the basic information shared at Stage 1 of the choice experiment

At Stage 2, respondents are required to answer the following question: "If you need additional information before making a final purchase decision, you may confirm this after waiting for about 5 seconds. Do you need additional information"? The waiting time for acquiring additional information is intended to realize consumers' search cost. The 5 seconds wait is decided based on a rule of thumb. An example of a product review given to respondents at Stage 3 is shown in Figure 31. In preparation for sharing with respondents at Stage 3, approximately 20 online reviews are collected from online news,

forums, or blogs. The reviews are arranged in a consistent format, revising review lengths to be similar to each other.

추가정보 제목	**리얼후기] 삼성 '기어S2', 생활 밀접형 "혁신"..."과연 그럴까?"** 디자인 '세련', 편리성 '충분'.."근데 꼭 필요해?"
추가정보 내용	**최근 삼성전자에서 야심 차게 출시한 '기어S2'의 인기가 하늘 높은 줄 모르고 치솟고 있다.** **이에 연일 미디어와 블로그 등 각종 사이트에는 "혁신을 혁신한 시계", "애플워치를 능가하는 기어S2", "S2 실생활 밀접, 이제 스마트워치 시대" 등 칭찬 일색의 각양각색 글들이 하루에도 수십 개씩 쏟아지고 있다. 아울러 '기어S2'와 관련한 새로운 소식이 전해질 때마다 실시간 검색순위에 오르는 등 그야말로 사회 전반에 걸쳐 엄청난 센세이션을 불러일으키고 있다.** **"기어S2가 정말 혁신을 혁신한 생활 밀접형 제품일까?" 기자가 직접 '기어S2'를 사용해봤다.** 디자인 일반 시계와 '유사', 편리성 '충분', "근데 꼭 필요해?" 삼성 '기어S2'는 지난 2일부터 판매에 들어간 삼성의 차세대 스마트워치로 동그란 원형 디자인으로 출시돼 일반 시계와 비교해도 어색하지 않을 정도로 세련된 디자인을 자랑한다. 이 때문에 애플의 '애플워치' 사각디자인보다 한층 진보한 스마트워치라는 평가를 받음은 물론, 실생활에 편리한 '티머니', '캐시비', 운동량 관리 앱인 'S 헬스', ATM 출금이 가능한 '우리은행'등 다양한 앱이 지원된다는 장점이 있다. 기자가 직접 사용해보니 이번 삼성 '기어S2'는 소문대로 실제 시계와 비교해도 어색하지 않을 만큼 세련된 디자인을 가지고 있었다. 조금 더 세밀히 묘사하자면 전자시계로 유명한 '지샥'의 기본 모델과 매우 흡사해 일반시계라고 봐도 무방할 정도로 깔끔한 외관을 지니고 있다. 무게 역시 상당히 가벼워 일반 전자시계와 무게가 비슷하다. 오히려 남성의 메탈 시계와 비교하면 더 가벼워 여성들이 착용하기에도 별 무리가 가지 않을 것으로 예상된다. 여기에 스마트폰 앱 중 가장 많이 사용되는 '카카오톡', '라인' 등도 지원해 업무상 스마트폰을 볼 수 없을 경우 '기어S2'는 충실한 알리미 역할을 해낼 것으로 전망된다. 또한, 사진 역시 별다른 조작을 하지 않아도 쉽게 볼 수 있어 가족들의 사진을 항상 볼 수 있다. 'S 헬스' 역시 만보기 기능 등을 지원해 하루 동안 걷는 양, 움직인 거리 등을 자동으로 측정하는 등 다양한 편리성을 제공한다. 하지만 이러한 다양한 장점들에도 불구하고 '기어S2'가 실생활에 반드시 필요한 제품이라고 보기에는 다소 무리가 있다. 우선 메신저 기능을 지원하기는 하지만 '기어S2'를 통해 의사 전달이 100% 불가능하고 시계를 볼 수 있는 상황에서 스마트폰을 보지 못한다는 건 억측에 가깝기 때문이다. 만일 사무실의 경우라 PC용 메신저를 사용하는 게 오히려 편하다. 사진의 경우 역시 시계라는 작은 디스플레이 안에서 구현해야 하기 때문에 장시간 보기에는 불편한 감이 있다. 한편, '기어S2'를 총괄적으로 사용해본 결과 실생활에 꼭 필요하다고는 느끼지 못했지만, 생활을 좀더 편리하게 해줄 수 있다는 의견에는 공감했다. 특히, 배터리 시간이 최대 3일에서 최소 2일 이상 지속되는 점은 상당히 긍정적인 요인이다. 가격 역시 '기어 S2' 33만3300원, '기어 S2 클래식' 37만4000원 수준으로 웬만한 고가의 시계보다 훨씬 저렴한 가격에 구매가 가능하기 때문에 한 번쯤 스마트워치를 사용해보고 싶은 유저라면 삼성의 '기어S2'를 구매하는 것도 좋은 방법이 될 듯하다. 다만, 실생활에서 '꼭' 필요한 기기를 원하는 유저라면 아직까지 스마트워치 구매는 권유하고 싶지 않다.
추가정보 사진	
추가정보 출처	브레이크뉴스 기사 (2015.10.13)

Figure 31. Example of product review information shared at Stage 3 of the choice experiment

At Stage 4, respondents make a final decision whether or not to purchase the target product. After completing the experiment, respondents are required to provide information on their demographics or other characteristics related to the target product.

6.2 Data Description

To obtain data, this study conducted web-based choice experiments with 798 adults aged between 20 and 59 years, in the Republic of Korea in November 2015. Table 10 shows the key characteristics of the respondents.

Table 10. Respondent Profile

	Value	Number of respondents (%)
Gender	Male	353 (44.2%)
	Female	445 (55.8%)
Ages (years)	20-29	267 (33.5%)
	30-39	288 (36.1%)
	40-49	156 (19.5%)
	50-59	87 (10.9%)
Average income per month (million KRW)	Under 1	25 (3.1%)
	1-2	61 (7.6%)
	2-3	135 (16.9%)
	3-4	141 (17.7%)

	4-5	146 (18.3%)
	5-6	106 (13.3%)
	6-7	61 (7.6%)
	7-8	51 (6.4%)
	8-9	27 (3.4%)
	9-10	14 (1.8%)
	Over 10	31 (3.9%)
OS of possessed smartphone	Android	650 (81.5%)
	iOS	138 (17.3%)
	Windows mobile	6 (0.8%)
	Blackberry	2 (0.3%)
	Bada OS	1 (0.1%)
	Tigen	1 (0.1%)
Possession for car	Yes	627 (78.6%)
	No	171 (21.4%)

Excluding missing data, the number of observations used in the empirical study of the smartwatch and the hybrid car are 787 and 789, respectively. The average amount of information searched by respondents is 1.7103 for the smartwatch and 1.7199 for the hybrid car.

6.3 Results and Discussion

For the hybrid car, the expected utility of respondent *i* at period *t* is specified as

follows:

$$U_{i,Hyb,t}^{E} = \beta_{car} X_{i,car} + \beta_{pi-ratio} X_{i,pi-ratio} + Q_{Hyb} + \frac{1}{1+(t-1)\sigma_{Hyb,0}^{2}} \nu_{Hyb,0} + \frac{\sigma_{Hyb,0}^{2}}{1+(t-1)\sigma_{Hyb,0}^{2}} \sum_{\tau=1}^{t-1} \eta_{i,Hyb,\tau} + \varepsilon_{i,Hyb,t} \quad \cdots\cdots\cdots\cdots\text{Eq. (73)}$$

In Equation 73, $X_{i,car}$ represents whether the respondent *i* possesses a car or not and $X_{i,pi-ratio}$ represents the ratio of the price of the hybrid car over average monthly income, i.e., price/income.

Table 11 shows the estimation results for the information search on the hybrid car.

Table 11. Estimation results for the hybrid car

Variable	Estimated result
$\nu_{Hyb,0}$	-0.2658**
Q_{Hyb}	-0.20354
$\sigma_{Hyb,0}^{2}$	1.7492**
β_{car}	0.2291*
β_{pi_ratio}	-0.0112**
c	0.02235***

Note: *, **, and *** implies p < 0.1, p < 0.05, and p < 0.01 respectively.

Based on the estimation results, all parameter estimates except the match value are significant at 10% significance level. Parameters related to the observed part of utility may be interpreted in the same manner as the classical discrete choice model, e.g., consumers who possess a car are more likely to make a purchase decision for the new hybrid car than consumers who do not own a car. In addition, the coefficient of price-income ratio is significantly negative. This means that if the price of the hybrid car increases for a given value of average income, then the product utility of the consumer decreases. Conversely, given the same price of the hybrid car for all consumers, the consumers with large average monthly income are more likely to purchase the car.

Now, consider the parameters suggested by the proposed model, $\nu_{Hyb,0}$, Q_{Hyb}, $\sigma^2_{Hyb,0}$, and c. In this result, the exact value of the consumers' average match value, Q_{Hyb}, and consumers' initial expectation about the match value, $\mu_{Hyb,0}$, both are undeterminable as the estimate of the true match value, Q_{Hyb}, is insignificant. However, the estimate of $\nu_{Hyb,0}$ may provide useful information about their initial gap, as it represents the difference between the true match value and the initial expectation, by definition. In this result, the initial perception bias has a negative value. This implies that consumers have lower expectations of the hybrid car than the actual value that it provides to consumers. Hence, consumers who do not search for additional information or search only a little are likely to have a lower preference for the hybrid car in comparison to consumers who search extensively. This initial gap may affect the initial sales of the product.

The parameter $\sigma^2_{Hyb,0}$ is interpreted relative to σ^2_S given that we have fixed $\sigma^2_S = 1$

in this study for identification purposes. Hence, the estimated value $\hat{\sigma}^2_{Hyb,0} = 1.7492$ implies that $\sigma^2_{Hyb,0}$ is 1.7492 times σ^2_S, which implies that the informational signal tends to be more concentrated around the true match value than the initial perception.

For the target product of smartwatch, the expected utility of respondent *i* at period *t* is specified as follows:

$$U^E_{i,Smart,t} = \beta_{Android} X_{i,Android} + \beta_{pi-ratio} X_{i,pi-ratio} + Q_{Smart} + \frac{1}{1+(t-1)\sigma^2_{Smart,0}} \nu_{Smart,0} + \frac{\sigma^2_{Smart,0}}{1+(t-1)\sigma^2_{Smart,0}} \sum_{\tau=1}^{t-1} \eta_{i,Smart,\tau} + \varepsilon_{i,Smart,t} \quad \text{Eq. (74)}$$

In Equation 74, $X_{i,Android}$ represents whether or not respondent *i* possesses a smart phone based on the Android OS and $X_{i,pi-ratio}$ represents the ratio of the price of the smartwatch against the average monthly income, i.e., price/income. The estimation results of Equation 74 are presented in Table 12.

Table 12. Estimation results for the smartwatch

Variable	Estimated result
$\nu_{Smart,0}$	-0.1895**
Q_{Smart}	-0.1569
$\sigma^2_{Smart,0}$	1.5576***
$\beta_{Android}$	-0.0063
$\beta_{pi-ratio}$	-0.2158
c	0.0044***

Note: *, **, and *** implies p < 0.1, p < 0.05, and p < 0.01 respectively.

Based on the results, the parameters related to the observed utility are insignificant at all significance levels considered. On the other hand, except for Q_{smart}, all parameters related to the information search process are significant at 5% significance level or better, similar to the case of the hybrid car. For both products, the match value estimates are shown to be insignificant. This may be interpreted as the inadequacy of the assumption of homogeneity in match values to describe actual consumer preferences. In fact, from a common sense perspective, it is more realistic to assume that each consumer feels differently about his/her product match values. This necessitates further research to adapt the assumption of heterogeneity in the match value.

Some practical implications may be suggested based on the estimation results of the proposed model. First, for both the hybrid car and the smartwatch, consumers tend to underestimate the product before they search for product information. Hence, firms'

marketing strategies should focus on encouraging consumers to discover the true value of the product. Such marketing strategies could include conducting a campaign to increase the probability of consumer exposure to the product or to provide consumers with a motive for searching for product information. Secondly, it may require more effort on the part of consumers to search for information on hybrid cars as the estimate of search cost for the hybrid car is greater than that for the smartwatch. As discussed in Section 2.1, consumers' behavior to search for information depends on product characteristics, especially on the difficulty for consumers to evaluate product quality. Generally, cars are categorized as experience goods (Huang et al., 2009; Nelson, 1970; Wan et al., 2012) while electric devices such as personal computers (PCs) or cell phones are classified as search goods (Bei et al., 2004; Wan et al., 2012). Hence, the proposed model provides estimation results that are consistent with the findings of previous empirical studies on consumers' information search behavior.

Although the estimation results in these empirical studies do not provide significant estimates of the match value, the probability of purchasing the product at the initial stages could be predicted if all parameters in the proposed model show significant results. For example, assume that the initial bias for product j, i.e., ν_{j0}, and the true product match value, Q_j, are both significantly negative, i.e., consumers ascribe a negative value to how much the product matches up to their intrinsic preference or expectation and have lower expectation of the match value compared to the true value before acquiring any information. In addition, assume that the absolute value of the summation of both variables indicating the initial mean perception of the match value is assumed large

enough to cancel out even the positive observed part of utility. In this case, researchers could conclude that consumers are unlikely to buy the product since the expectation of consumers' initial value of expected utility is less than the no purchase option. In other words:

$$E_{\varepsilon}\left[U_{ij0}^{E}\right]=\mathbf{X}_{ij}'\hat{\mathbf{b}}+\hat{\mu}_{ij0}=\mathbf{X}_{ij}'\hat{\mathbf{b}}+\hat{v}_{j0}+\hat{Q}_{j}<0=E_{\varepsilon}\left[U_{io,0}^{E}\right] \cdots\cdots\cdots\cdots\cdots\cdots \text{Eq. (75)}$$

Hence, the estimation results could hint that the product will have failed as the consumers' initial valuation of the product is insufficient to purchase the product and that valuation would have an effect on initial sales. Hence, the estimation results of the proposed model may suggest that the firm should focus on securing the initial sales of the product in this case.

Chapter 7. Conclusion

7.1 Concluding Remarks and Contributions

This study proposes a structural model that explains consumers' behavior of searching for information with learning in the presence of perceived uncertainty. Although exploring the consumer search process appeals to many researchers in the fields of marketing and economics, existing research streams are limited in terms of the description of the actual consumer decision-making process. The consumer search framework, one of major streams of existing studies, has focused on consumers' decision whether or not to search for information based on the assumption that once consumers search for information on the alternative, the related uncertainty is removed perfectly. Hence, the framework has not considered consumers' successive learning process from information search. In the consumer search framework, the consumer's objective of searching is to find the product that is expected to give the best expected utility or the seller who offers the best price to consumers.

The other major research stream, the consumer learning framework, assumes that even if consumers update their belief based on acquired information, the uncertainty of the alternative is not removed perfectly. In the consumer learning framework, consumers successively update their belief regarding the uncertainty under the assumption of persistent exposure to informational signals. In other words, the framework focuses researchers' attention on the consumers' approach to processing acquired information.

However, the pre-purchase stage in consumers' decision-making process comprises consumers' behavior to actively search for information and to process the acquired information, both of which are clearly interconnected. Although a few recent studies consider this point (Branco et al., 2012; Lelis & Howes, 2008), their approach remains restricted to the theoretical field without extension to empirical investigation. Moreover, although the most astonishing change brought about by the Internet in consumers' information search behavior is that consumers may learn of product attributes that are unobservable prior to purchase by obtaining product reviews made by others, researchers have not adequately considered this feature through a structural modeling approach.

Based on these points of view, the contributions of this study may be summarized as follows. Firstly, this study proposes a model to explain consumers' indirect learning about the product from online reviews by combining the consumer search and consumer learning frameworks. Consumer learning framework may be presumed to focus on consumers' information processing behavior when exposed to product information, which could be regarded as passive information search behavior. On the other hand, consumer search framework focuses on explaining consumers' active information search behavior under the assumption that repeatedly searching for information on the same alternative is meaningless. By combining the theoretical background of these two research frameworks, the model proposed in this study enables us to explain consumers' active and repeated search for product information along with the processing of acquired information.

Secondly, this study suggests a framework that allows the prediction of consumers' product preferences using data related to their information search behavior. In the

proposed model, consumers' decision to search for information is based on the product utility function. Hence, the utility of the product could be estimated by observing consumers' decisions to search for information.

Thirdly, this study suggests a basic model to be used for empirical analysis. The framework suggested by this study may be applied to any dataset collected online. In addition, the proposed model could be applied to empirically examine various hypotheses of consumer search behavior.

Lastly, the proposed model in this study provides more practical implications than the conventional discrete choice model by providing information about the perception bias, signal noise, and match value. By definition, the perception bias of consumers provides marketers or firms with information on how consumers initially perceive the product. For example, a negative perception bias indicates that consumers underestimate the product in the absence of additional information. In this case, firms could conduct major campaigns to increase the chances of product exposure in order to change consumer perception, or could encourage consumers to search for more product information, which may positively affect consumer's probability of purchase. If the perception bias is positive, firms are recommended to investigate why consumers overestimate the product and utilize this information to fill the perception gap. In this way, the research framework suggested in this study is useful to build a marketing strategy from a practical standpoint.

7.2 Limitations and Future Research Topics

Although this study proposes a structural model to explain consumer behavior in the pre-purchase stage with empirical investigation, some limitations remain with respect to describing consumers' information search behavior. This section discusses the limitations of this study and recommends future research topics related to these limitations.

Firstly, this study does not consider the effect of information sources. It is well known that consumers' information search behavior varies depending on the characteristics of their information sources, including credibility, search cost, or consumer involvement. In this respect, a consumer's decision to search for information inherently contains the problem of choosing the information source. Hence, future research may consider including this choice task in the framework of consumers' information search process. To achieve this, the variance of signal noise and search cost may be considered information source specific variables in the proposed model.

Second, this study suggested an algorithm for structural estimation, restricted to the binary choice situation for simplicity. However, extending to the multinomial choice situation is more plausible. Future research may deal with this by adopting the same estimation scheme, Bayesian MCMC with the kernel smoothing method for calculating expected reward. In addition, future research could construct estimation algorithms that allow heterogeneous consumer preferences in order to release the restriction in the proposed model.

Thirdly, the heterogeneity of parameters in the learning process may be specified in a hierarchical structure. For example, consumers' initial perception bias could be affected

by prior knowledge about the product or usage experience of similar products. Hence, implementing a hierarchy in the parameters may explain many additional aspects of consumers' information search behavior.

Lastly, this study considers the objective of information search to be limited based on the product level by assuming that the unobserved part of consumers' product utility is captured by the aggregate match value. However, consumers may search for information on attribute values that they are uncertain of or may acquire information in order to revise their attribute valuations. Future research may make a distinction between these two scenarios of information search and adapt the proposed model for each case.

Bibliography

Ackerberg, D. (2003). Advertising, learning, and consumer choice in experienced goods markets: An empirical examination. *International Economic Review*, 44, 1007-1040.

Aguirregabiria, V. & Mira, P. (2010). Dynamic discrete choice structural models: A survey. *Journal of Econometrics*, 156, 38-67.

Alloy, L.B. & Tabachnik, N. (1984). Assessment of covariation by humans and animals: The joint influence of prior expectations and current situational information. *Psychological Review*, 91(1), 112-149.

Assunção, J.L. & Meyer, R.J. (1993). The rational effect of price promotions on sales and consumption. *Management Science*, 39(5), 517-535.

Awasthy, D., Banerjee, A., & Banerjee, B. (2012). Understanding the role of prior product knowledge to information search: An application of process theory to the Indian market. *Asia Pacific Journal of Marketing and Logistics*, 24(2), 257-287.

Bather, J. (2000). *Decision theory: An introduction to dynamic programming and sequential decisions*. Chichester: John Wiley & Sons, Ltd.

Bei, L.-T., Chen, E.Y.I., & Widdows, R. (2004). Consumers' online information search behavior and the phenomenon of search vs. experience products. *Journal of Family and Economic Issues*, 25(4), 449-467.

Bhatnargar, A. & Ghose, S. (2004). Online information search termination patterns across

product categories and consumer demographics. *Journal of Retailing*, 80, 221-228.

Bickart, B. & Schindler, R.M. (2001). Internet forums as influential sources of consumer information. *Journal of Interactive Marketing*, 15(3), 31-40.

Branco, F., Sun, M., & Villas-Boas, J.M., (2012). Optimal search for product information. *Management Science*, 58(11), 2037-2056.

Camerer, C.F., Ho, T.-H., & Chong, J.-K. (2004). A cognitive hierarchy model of games. *Quarterly Journal of Economics*. 119(3), 861-898.

Chang, L.-Y., Lee, Y.-J., & Huang, C.-L. (2012). The influence of e-word-of-mouth on the consumer's purchase decision: A case of body care products. *Journal of Global Business Management*, 6, 1-7.

Cheng, R. (2013). Sequential search beats best-of-n search. *Sequential Analysis*, 32, 183-195.

Ching, A.T., Erdem, T., & Keane, M.P. (2013). Learning models: An assessment of progress, challenges, and new developments. *Marketing Science*, 32(6), 913-938.

Ching, A.T., Imai, S., Ishihara, M., & Jain, N. (2012). A practitioner's guide to Bayesian estimation of discrete choice dynamic programming models. *Quantitative Marketing & Economics*, 10(2), 151-196.

Chorus, C.G. & Timmermans, H.J.P. (2008). Revealing consumer preferences by observing information search. *Journal of Choice Modelling*, 1(1), 3-25.

Chylinski, M.B., Roberts, J.H., & Hardie, B.G.S. (2012). Consumer learning of new

binary attribute importance accounting for priors, bias, and order effects. *Marketing Science*, 31(4), 549-566.

Daurer, S., Molitor, D., & Spann, M. (2012). Measuring individual search costs on the mobile Internet. *Proceedings of the 20th European Conference on Information Systems* (ECIS). AIS

Daurer, S., Molitor, D., Spann, M., & Manchanda, P. (2015). Consumer Search Behavior on Mobile Internet: An empirical analysis (Ross School of Business Working Paper No.1275). Ann Arbor, MI: University of Michigan. Retrieved September 25, 2015, from http://www.researchgate.net/publication/277164466\nhttp://www.researchgate.net/publication/277164466

Degroot, M.H. (2004). *Optimal statistical decisions* (Wiley Classics Library ed.). New Jersey: John Wiley & Sons, Inc.

De los Santos, B., Hortaçsu, A., & Wildebeest, M.R. (2012). Testing models of consumer search using data on web browsing and purchasing behavior. *American Economic Review*, 102(6), 2955-2980.

Detlor, B., Sproule, S., & Gupta, C. (2003). Pre-purchase online information seeking: Search versus browse. *Journal of Electronic Commerce Research*, 4(2), 72-84.

Erdem, T. & Keane, M.P. (1996). Decision-making under uncertainty: Capturing dynamic brand choice processes in turbulent consumer goods markets. *Marketing Science*, 15(1), 1-20.

Erdem, T., Keane, M.P., Öncü, T.S., & Strebel, J. (2005). Learning about computers: An anaylsis of information search and technology choice. Quantitative Marketing and Economics, 3(3), 207-246.

Gabaix, X., Laibson, D., Moloche, G., & Weinberg, S. (2006). Costly information acquisition: Experimental analysis of a boundedly rational model. *American Economic Review*, 96(4), 1043-1068.

Gamerman, D. & Lopes, H. (2006). *Markov Chain Monte Carlo: Stochastic simulation for Bayesian inference* (2nd ed.). New York: Chapman & Hall/CRC.

Ghose, A., Ipeirotis, P.G., & Li, B. (2012). Search less, find more? Examining limited consumer search with social media and product search engines. *Proceedings of the 33rd International Conference on Information Systems* (pp. 1-19). Orlando: Association for Information Systems.

Ghose, A., Goldfarb, A., & Han, S.P. (2013). How is the mobile Internet different? Search costs and local activities. Information Systems Research, 24(3), 613-631.

Goh, K.-Y., Chu, J., & Wu, J. (2015). Mobile advertising: An empirical study of temporal and spatial differences in search behavior and advertising response. *Journal of Interactive Marketing*, 30, 34-45.

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

Häubl, G., Dellaert, B.G.C., & Donkers, B. (2010). Tunnel vision: Local behavioral

influences on consumer decisions in product search. *Marketing Science*, 29 (3), 438-455.

Hoeffler, S., Ariely, D., & West, P. (2006). Path dependent preferences: The role of early experience and biased search in preference development. *Organizational Behavior and Human Decision Processes*, 101, 215-229.

Holloway, D. (2012, April 6). Survey: 85% of consumers use the Internet to find local businesses. Message posted to https://biz.yelp.com/blog/survey-85-of-consumers-use-the-internet-to-find-local-businesses.

Huang, P., Lurie, N.H., & Mitra, S. (2009). Searching for experience on the web: An empirical examination of consumer behavior for search and experience goods. *Journal of Marketing*, 73(2), 55-69.

Hutchinson, J.W. & Meyer, R.J. (1994). Dynamic decision making: Optimal policies and actual behavior in sequential choice problems. *Marketing Letters*, 5(4), 369-382.

Imai, S., Jain, N., & Ching, A. (2009). Bayesian estimation of dynamic discrete choice models. *Econometrica*, 77(6), 1865-1899.

Kaye, L. (2014, October 28). 94% of B2B buyers research online for purchase decisions. Message posted to http://www.brafton.com/news/94-percent-b2b-buyers-research-online-purchase-decisions.

Kim, J.B., Albuquerque, P., & Bronnenberg, B.J. (2010). Online demand under limited consumer search. *Marketing Science*, 29(6), 1001-1023.

Klein, L.R. (1998). Evaluating the potential of interactive media through a new lens:

Search versus experience goods. *Journal of Business Research*, 41(3), 195–203.

Klein, L.R. & Ford, G.T. (2003). Consumer search for information in the digital age: An empirical study of repurchase search for automobiles. *Journal of Interactive Marketing*, 17(3), 29-49.

Koulayev, S. (2013). Search with dirichlet priors: Estimation and implications for consumer demand. *Journal of Business & Economic Statistics*, 31(2), 226-239.

Kundisch, D. (2000). Buyer search behavior on an electronic commodity market: Consumer's decision for a sequential or simultaneous search method. In S.H. Kim, O. Sheng, J.K. Lee, A.B. Whinston, & B. Schmid (Eds.), *Proceedings of the International Conference on Electronic Commerce (ICEC) 2000* (pp. 88-93). Seoul: International Center for Electronic Commerce.

Lee, C., Kim, J., & Chan-Olmsted, S.M. (2011). Branded product information search on the Web: The role of brand trust and credibility of online information sources. *Journal of Marketing Communications*, 17(5), 355–374.

Lelis, S. & Howes, A. (2008). A Bayesian model of how people search online consumer reviews. In B. C. Love, K. McRae, & V. M. Sloutsky (Eds.), *Proceedings of the 30th Annual Conference of the Cognitive Science Society* (pp. 553-559). Austin, TX: Cognitive Science Society.

Levin, I.P., Huneke, M.E., & Jasper, J.D. (2000). Information processing at successive stages of decision making: need for cognition and inclusion–exclusion effects. *Organizational Behavior and Human Decision Processes*, 82(2), 171-193.

Liu, C., Rau, P.-L., & Gao, F. (2010). Mobile information search for location-based information. *Computers in Industry*, 61(4), 364–371.

McCall, J.J. (1970). Economics of information and job search. *Quarterly Journal of Economics*, 84(1), 113–26.

Mohr, J., Sengupta, S., & Slater, S. (2010). *Marketing of high-technology products and innovations* (3rd ed.). New Jersey: Pearson Education, Inc.

Morrison, K. (2014, November 28). 81% of shoppers conduct online research before buying [infografic]. Message posted to http://www.adweek.com/socialtimes./81-shoppers-conduct-online-research-making-purchase-infographic/208527.

Narayanan, S. & Manchanda, P. (2009). Heterogeneous learning and the targeting of marketing communications for new products. *Marketing Science*, 28(3), 424-441.

Nelson, P. (1970). Information and consumer behavior. *Journal of Political Economy*, 78(2), 311-329.

Norets, A. (2009). Inference in dynamic discrete choice models with serially correlated unobserved state variables. *Econometrica*, 77(5), 1665-1682.

Ratchfold, B.T. (1982). Cost-benefit models for explaining consumer choice and information seeking behavior. *Management Science*, 28(2), 197-212.

Richard, M.-O., Chebat, J.-C., Yang, Z., & Putrevu, S. (2010). A proposed model of online consumer behavior: Assessing the role of gender. *Journal of Business Research*, 63, 926-934.

Rust, J. (1987). Optimal replacement of GMC bus engines: An empirical model of Harold

Zurcher. *Econometrica*, 55, 999-1033.

Shim, S., Eastlick, M.A., Lotz, S.L., & Washington, P. (2001). An online pre-purchase intentions model: The role of intention to search. *Journal of Retailing*, 77, 397-416.

Shin, S., Misra, S., & Horsky, D. (2012). Disentangling preferences and learning in brand choice models. *Marketing Science*, 31(1), 115-137.

Signorino, C.S. (2003). Structure and uncertainty in discrete choice models. *Political Analysis*, 11(4), 316-344.

Simon, H.A. (1955). A behavior model of rational choice. *Quarterly Journal of Economics*, 69(1), 99-118.

Stigler, G.J. (1961). The economics of information. *Journal of Political Economy*, 69(3), 213-225.

Su, B. (2008). Characteristics of consumer search on-line: How much do we search? *International Journal of Electronic Commerce*, 13(1), 109-129.

Train, K. (2009). *Discrete choice method with simulation* (2nd ed.). New York: Cambridge University Press.

Wan, Y., Nakayama, M., & Sutcliffe, N. (2012). The impact of age and shopping experiences on the classification of search, experience, and credence goods in online shopping. *Information Systems and e-Business Management*, 10(1), 135-148.

Weitzman, M.L. (1979). Optimal search for the best alternative. *Econometrica*, 47(3),

641-654.

Wiegmann, D.D., Seubert, S.M., & Wade, G.A. (2010). Mate choice and optimal search behavior: Fitness returns under the fixed sample and sequential search strategies. *Journal of Theoretical Biology*, 262, 596–600.

Yang, L.C., Toubia, O., & De Jong, .M.G. (2015). A bounded rationality model of information search and choice in preference measurement. *Journal of Marketing Research*, 52(6), 166-183.

Zhao, Y., Yang, S., Narayan, V., & Zhao, Y. (2013). Modeling consumer learning from online product reviews. *Marketing Science*, 32(1), 153-169.

Abstract (Korean)

최근 인터넷을 통한 정보탐색의 비중이 증가하면서, 소비자들은 제품에 대해 느끼는 불확실성이 충분히 줄어들 때까지 온라인 리뷰 정보를 반복적으로 탐색하고 효용을 업데이트하는 경향을 보인다. 그러나 소비자의 의사결정 행위에 대해 다룬 기존 연구들은 전체 의사결정 과정이 아닌 일부 단계만을 포함하고 있을 뿐 아니라, 실증분석을 통해 의사결정 구조를 명확히 추정하는 대신 이론적 모형 제시에만 그치고 있다는 한계점을 가지고 있다. 따라서 본 연구는 소비자의 순차적 정보탐색 행위 및 학습과정을 설명할 수 있는 구조적 모형을 제안하고 실증분석을 수행함으로써 기존 문헌이 가지는 한계를 극복하고자 하였다. 본 연구는 소비자의 정보탐색 행위를 설명하기 위한 동적 이산선택모형을 구성하고 베이지안 학습 매커니즘을 도입해 소비자의 효용 학습과정을 모형에 반영하였다. 이와 더불어 본 연구는 소비자의 정보탐색 행위에 대한 관측치를 바탕으로 제품에 대한 소비자의 선호를 추정할 수 있는 연구 프레임워크를 제안하였다.

본 연구는 제안모형의 추정을 위해 베이지안 추정 방법론을 활용하였다. 기존 동적 이산선택모형 추정방식인 동적 계획법의 계산상 부담을 피하기 위해, 본 연구는 변형 메트로폴리스-헤이스팅스 알고리즘의 일종인 IJC 알고리즘을 적용하여 예측치를 추출하였다. 마지막으로 본 연구는 선택실험을 통해 수집

한 데이터를 활용하여 실증분석을 수행함으로써, 제안모형의 실증적 타당성을 확인하였다. 실증분석 결과, 제안모형은 소비자가 최초에 가지는 인식 편향 및 불확실성 정도 등과 같이 불확실성에 대한 정보를 제공해줌으로써, 기존 이산선택 모형에 비해 더욱 실질적인 함의를 제공할 수 있을 것으로 기대된다.

주요어 : 정보 탐색; 순차적 탐색; 소비자 학습; 동적 이산선택모형; 베이지안 추정
학 번 : 2010-21082