Ph. D. Dissertation in Engineering

Development of Discrete Choice Model Considering the Budget Allocation Stage

: Focusing on Multi-stage and Multi-category Cases with Outside goods

소득배분 단계를 고려한 이산선택 모형 개발

August 2012

Jungwoo Shin

Technology Management, Economics and Policy Program
College of Engineering
Seoul National University

Development of Discrete Choice Model Considering the Budget Allocation Stage

: Focusing on Multi-stage and Multi-category Cases with Outside goods

지도교수 이 종 수
이 논문을 공학박사 학위논문으로 제출함

2012 년 8 월

서울대학교 대학원
협동과정 기술경영경제정책전공
신 정 우

신정우의 박사학위논문을 인준함
2012 년 8 월

위 원 장 _____(인)

부위원장 _____(인)

위 원 _____(인)

위 원 _____(인)

위 원 _____(인)

Abstract

Development of Discrete Choice Model Considering the Budget Allocation Stage

: Focusing on Multi-stage and Multi-category Cases with Outside goods

Jungwoo Shin
Technology Management, Economics and Policy Program
College of Engineering
Seoul National University

Due to—and in pursuit of—economic growth in the market, firms tend to release a variety of products and services onto the market. With competition occurring among these products, only a few of them survive and actually make a profit. In addition, while government performs various activities—such as constructing infrastructure, establishing a technology development roadmap, and supporting the diffusion of eco-friendly products, in order to pursue sustainable national development—few of its policies have had a lasting effect on the market. Thus, firms and government each bear limitations in being able to suggest effective product plans and policies, given that market uncertainty means

that consumer preferences vis-à-vis new products or policies cannot be predicted in advance. Therefore, to reduce market uncertainty, research on purchasing behavior is essential to understanding the perspectives inherent in establishing management strategy and policy direction.

Research on consumer purchasing behavior towards various products and services has been conducted in a variety of fields, such as the information technology (IT), marketing, energy, and environmental industries, among others. Additionally, various models have been developed and applied to analyze consumer purchasing behavior. Among these various methodologies, the discrete choice model—one of the more effective methods used to estimate consumer preference—has been developed in a diversity of ways, including the single choice model, the multiple choice model, and the choice model with heterogeneity. However, most of the previous models have focused only on inside goods, and do not consider product prices or attributes in other categories (i.e., outside goods). Economically, outside goods are included in a choice model, to analyze realistic price policy and consumer preference. Moreover, because a consumer's decision-making process contains more than one stage—save for those related to the simplest of decision-making—choice models that consider multiple stages of the consumer decision-making process need to be developed.

From these points, it is clear that the development of a new choice model that considers both outside goods and structure with regard to a consumer's decision-making process is needed. The issue of considering outside goods within a choice model of

demand analysis converges into cases that consider the budget allocation stage of a consumer's decision-making process, from an individual-level perspective. Therefore, the purpose of this dissertation is to develop a choice model that includes the budget allocation stage, in order to consider both outside goods and the consumer's decision-making stage. The proposed model in this dissertation can be used to undertake comprehensive analysis vis-à-vis the impact of the budget structure of outside goods on the product choice of inside goods; it can also be used in the general analysis of consumer preference vis-à-vis product prices and the attributes of inside goods.

The proposed choice models in this dissertation are classified into two cases: one consists of the budget allocation and product choice stages, and the other of the budget allocation, product choice, and product usage stages. Based on these two proposed model types, this dissertation conducts empirical research on representative products from the information communication and technology (ICT), household products, and automobile industries. In the first empirical study, because consumer consumption patterns have been influenced in many ways by the introduction of "smart" devices, consumer purchasing behavior with regard to smart devices is analyzed by considering the budget allocation stage. In the second empirical study, demand analysis with regard to choices in eco-friendly products—with the endpoint of promoting green growth—is conducted by using a choice model with considering the budget allocation stage. In the third empirical study of this dissertation, consumer preference with regard to product choice and usage for smart cars—a next-generation automobile—is analyzed by considering the budget

allocation stage. Through the three empirical studies, the utilization of the proposed models and the implications that derive from proposed models are examined.

In conclusion, based on the proposed models in this dissertation, it is possible to analyze consumer purchasing behavior more accurately by comprehensively analyzing the impact of changes in the allocated budget between inside goods and outside goods on product choice and usage of inside goods. Moreover, the estimation results from the proposed model are expected to provide useful information that can be used to perform accurate demand forecasting vis-à-vis new products, and establish policy suggestions. Additionally, the proposed models are also expected to bear methodological implications with regard to analysis of the consumer decision-making process.

Keywords: Choice model, Consumer purchasing behavior, Budget allocation, Outside goods, Demand analysis, Consumer decision-making process

Student Number: 2007-20888

Contents

List of Tables

List of Figures

Chapter 1. Introduction

1.1 Research Background

1.1.1 Standard Choice Model for Analyzing Consumer Purchasing Behavior

Research on consumer purchasing behavior with regard to a variety of products and services has been conducted in a diversity of ways. Studies of consumer purchasing behavior analyze changes in demand as generated by a product's attributes; generally speaking, they touch on the selection of certain products versus others in competition, and product diffusion, by analyzing consumer choice behavior. The analysis results have been used to establish management strategies, such as those that dictate the direction of product development and corporate decisions regarding price levels. For instance, analysis regarding changes in consumer demand according to the level of product attributes could generate results that provide direction vis-à-vis product development; likewise, analysis regarding changes in consumer demand as a function of price could derive findings that can be used in future product pricing. Furthermore, consumer purchasing behavior has been examined, and the results of such research have been used to establish national policy.

Over the last century, businesses—and the people who run them—have been continuously pursuing rapid economic growth through industry restructuring. Due to

rapid economic growth in the market, plenty of products and services are being released, but only a few of them survive and make a profit. In addition, while the government performs various activities—such as building infrastructure, establishing a technology development roadmap, and supporting the proliferation of environmentally friendly products at the national level, in order to promote sustainable national development—moves to establish an effective policy without giving due consideration to consumer preferences and purchasing behavior have been challenged. Thus, research on consumer purchasing behavior is important not only in academia, but also in making government enterprise management and policy-making more effective.

From the viewpoint of entrepreneurs and the necessity of analyzing consumer purchasing behavior, market uncertainty arises when companies launch new products, because it is difficult to gauge in advance what customers' reactions to new products will be like. For example, market experts expected the iPad not to be successful in the market, and that it would be "weeded out" when it first emerged. The rationale for this expectation is that it featured the same content, at a different size, as the iPhone, which had already spread among customers, and that netbooks—which are lighter than laptops—had already been on the market for quite some time. However, because Apple had identified the consumer desire to use visual multimedia content on a large-sized screen, it has had tremendous success with the iPad (Choi et al., 2012). Thus, in order to reduce market uncertainty, analysis of consumer purchasing behavior in terms of the management strategy for a new product is essential.

Regarding the need to analyze consumer purchasing behavior from a policy perspective, when establishing new policies in a country, proactively identifying consumer reactions to a new policy are useful in terms of making suggestions and considering new policies. For example, when an adequate social infrastructure has been constructed, purchasing subsidies and automobile-related tax subsidies can be used, to good effect, to promote the proliferation of environmentally friendly electric cars; these two instruments can thus be considered means of carrying out national policies. In this case, it is difficult to predict which of the two instruments would be more effective in promoting electric automobiles; thus, by offering both and analyzing consumer purchasing behavior with regard to electric cars and the form of subsidy to which consumers best respond, policy-makers can reduce uncertainty (Shin et al., 2012). Therefore, analysis of consumer purchasing behavior is needed, to establish effective policy direction.

A variety of discrete choice models have been developed and used to analyze consumer purchasing behavior. Discrete choice models are one of the useful methods to estimate consumer preference (McFadden, 2000). Various discrete choice models are classified according to their response type (e.g., single/multiple choice model) and inherent assumptions (e.g., single/multiple choice model with homogeneity/heterogeneity of consumer taste). In the single choice models, the multinomial logit models, the multinomial probit (MNP) models and so on are included. In the multiple choice models, the multivariate probit (MVP) models, the multiple discrete-continuous extreme value

(MDCEV) models, the latter of which also considers product usage, and so on are included. In the single choice models that consider consumer heterogeneity, random coefficient logit (or mixed logit) models and so on are included. Discrete choice models have been used in various industries, such as energy and the environment (Matsukawa and Ito, 1998; Lee and Kwak, 2007), IT (Kim et al., 2006; Mueller et al., 2006), and marketing (Chintagunta, 1992; Allenby et al., 2004a).

Due to the increased availability of consumer information, individual and market-level data have accumulated more than ever; this has enabled researchers to analyze consumer purchasing behavior in a variety of fields. However, despite representing a wealth of information on the purchasing behavior of consumers, a choice model still needs to be developed to analyze purchasing behavior effectively. In addition, when there is a new type of product that does not resemble already-existing products, the development of a choice model and survey techniques—along with survey data collection methods for analyzing and forecasting demand—is especially important.

In summary, the development of both choice models and survey techniques are important to the analysis of consumer purchasing behavior. However, this dissertation focuses on the development of new choice models for use in conducting more accurate demand analysis and forecasting, rather than the development of survey techniques. Thus, because choice models play an important role in analyzing and forecasting demand, management strategy and policy proposals for entrepreneurs and policy-makers are expected to provide important information regarding consumer purchasing behavior, as

seen via the choice models proposed in this dissertation.

1.1.2 Budget Allocation (Inside goods v.s. Outside goods) Stage in the Choice Model

Analysis of consumer purchasing behavior on the basis of a choice model is used to identify the key factors that affect consumer choice. In the analysis, explanatory variables include product attributes, consumer demographic variables, and social/environmental variables. As dependent variables, discrete data (i.e., single choice data and multiple choice data) or continuous data (i.e., singe usage data and multiple usage data) are used. To date, most of the previous literature has analyzed the effect of explanatory variables on the choice of a specific product within a single product category, based on revealed preference (RP) data from actual purchase information or a stated preference (SP) data from a survey. This research assumes that the consumer decision-making process, in a single stage, reveals that consumers choose the most preferred alternative, compared to competitive products. In addition, single-category and multi-category cases from the previous literature can be divided, depending on the alternatives typology. The choice models for the single-stage plus single-category cases are the multinomial logit model, the MNP model, and the mixed logit model, among others; the choice models for the single-stage plus multi-category cases are MVP model and so on.

When consumers decide to buy a product, the consumer decision-making

process follows a multiple-stage decision-making process more frequently than a single-stage one (Bettman, 1979; Shocker et al. 1991). Deaton and Muelbauer (1980) mention that consumers go through two stages in the decision-making process. In the first stage, they decide how to allocate their total income, y, to inside goods and outside goods, whereas in the second stage they determine the demand of each category based on the budget allocation results for each category. Various areas of the literature consider the consumer's two-stage decision-making process (Blundell et al., 1993; Anders and Moser, 2010): when consumers choose a product, they consider not only alternatives in a focused category, but also the total allocated budget in a focused category. Therefore, consumers choose to maximize their utility under their total allocated budget in a focused category, as a budget constraint. For instance, to purchase a new television, a consumer visits the electronics store; under budget—which is allocated for buying a new television from total income Y—he or she chooses the most preferred television.

In addition, some previous studies propose that the consumer decision-making process has more than two stages. For instance, Jaffe and Senft (1966) suggest five decision-making stages: information-seeking, the pre-purchase stage (including budget allocation), the buying stage, and the post-purchase stage. Moreover, Du and Kamakura (2008) divide the budget allocation stage into two stages: the choice of consumption expenditure category, and the decision vis-à-vis expenditure size. From these viewpoints, researchers have sought to develop various models that consider a multi-stage decision-making process.

Choice models for the multi-stage decision-making process could be classified into single-category and multi-category cases, according to the types of alternatives considered in the model. For all multi-stage cases, the impacts of different stages on product choice are analyzed simultaneously; if researchers do not consider simultaneity between choice stages and analyze only product choice in the final stage, endogeneity could occur and the estimation results could be biased. Therefore, to consider a simultaneous situation in which a previous choice affects the subsequent choice, simultaneous equation models that resolve the endogeneity problem have been developed.

On the other hand, the consideration of the budget allocation stage in the choice model is similar to the simultaneous consideration of inside and outside goods in the choice model. In other words, the consideration of outside goods in the choice model, from the consumer purchasing perspective, converges with the budget allocation problem among product categories. Berry, Leveinshon, and Pakes (BLP) (1995) and Nevo (2001) each mention that if outside goods are not considered in the choice model, then this choice model has converged with the choice problem, which includes only inside goods. In this choice model, when prices for all inside goods increase together, the aggregate output does not change. Thus, the choice model—which considers only inside goods—has limitations when it comes to reflecting any demand change that is dependent on price changes for all inside goods. In real life, if the prices for all inside goods increase, the total demand for the category (including inside goods) is reduced. Therefore, to revise an unrealistic situation in the choice model, generally, outside goods are included with the

consumers' utility function in economics. For example, models that consider outside goods have been developed, such as the MDCEV model, BLP model and so on.

Thus, consideration of the budget allocation stage—which brings outside goods to the choice model—is necessary to analyzing consumer purchasing behavior accurately. In addition, by considering the budget allocation stage in the choice model from the perspectives of the multi-stage and multi-category models, a general choice model should be developed (Chintagunta and Nair, 2010). Therefore, the proposed model—which considers the budget allocation stage from the multi-stage and multi-category perspectives—is expected to assist in analyzing product demand, and forecasting, accurately.

1.2 Research Motivation and Objectives

Through the analysis of demand forecasting and product diffusion for new products, studies about policy and product strategy have been conducted by a number of researchers. Such analysis is ultimately used to examine consumer purchasing behavior. From the perspective of consumer purchasing behavior, the collection of pertinent data and the use of analysis models are needed to accurately analyze demand forecasting and the product diffusion of newly released products. Therefore, the development of sophisticated data collection methods to capture information vis-à-vis consumer purchasing behavior and the development of discrete choice models to analyze consumer

preferences are important issues from the perspectives of product strategy and policy research (Allenby et al., 2005).

Data on consumer purchasing behavior can be categorized into two different types: individual-level data and market-level data. Individual-level data are further classified into two classes: stated preference (SP) data, which are collected through surveys, and revealed preference (RP) data, which are collected by purchasing real-life information. SP data are mainly used to analyze future products that have not yet been released to the real market. For instance, SP data are used to analyze consumer preferences vis-à-vis electric cars and hydrogen cars, because such information does not currently exist in the market (Ewing, 2000; Brownstone, 1999; Kim, 2007; Ahn, 2008; Lee and Cho, 2009; Bhat and Sen, 2006). Because SP data are collected through surveys, consumer preference could be overestimated through an over-reliance on them. Thus, to remedy SP data shortcomings, some studies analyze consumer purchasing behavior by using RP data and SP data concurrently (Bhat and Casterlar, 2002; Brownstone, 2000).

Based on McFadden's (1974) random utility theory, methodologies with regard to consumer preference can developed in a variety of ways: the multinomial logit model, the MNP model, the random coefficient logit (or mixed logit model), the MVP model, and the MDCEV model are a few examples. In addition, these consumer preference research methodologies have been used to analyze consumer purchasing behavior in fields as diverse as energy, IT, environment, and education (Ewing and Sarigollu, 2000; Wei et al., 2005; Jepsen, 2008; Allenby et al., 2004a). However, previously used

methodologies in consumer preference research have focused solely on inside goods, and care little about product price or the attributes of other categories (i.e., outside goods) (Chintagunta and Nair, 2010).

To consider outside goods in the choice model, the main purpose of this dissertation is to develop a choice model that includes a budget allocation stage, to analyze consumer purchasing behavior from the perspectives of the multi-stage and multi-category model. According to Chintagunta and Nair (2010), an economic consideration of outside goods in the choice model is important, as it helps one analyze consumer behavior. In other words, alternatives from multiple categories that include both inside and outside goods are considered in the proposed choice model. Allenby et al. (2004b) mention that if researchers would like to analyze realistic pricing policy, outside goods should be included in their choice model. Because a price change for either outside or inside goods affects consumer preference for inside goods, both are considered in the choice model, in order to facilitate the analysis of accurate product demand, product forecasting, and consumer willingness to pay. Therefore, the development of a new choice model that considers outside goods in the form of the budget allocation stage has the advantage not only of estimating consumer purchasing behavior more accurately; it also has the advantage that it utilizes the estimation results by establishing a product strategy to diffuse new products, which can lead to more efficient policy.

Moreover, to analyze consumer purchasing behavior in-depth, the choice model should consider the perspectives of a multi-stage model, as well as those of a multi-

category model. From those perspectives, the impact of a change to allocated budgets between categories of outside and inside goods on choices among inside goods should be analyzed by considering the budget allocation stage between outside and inside goods. Thus, this dissertation proposes a choice model that considers the budget allocation stage for outside goods, under the perspectives of the multi-category and multi-stage models. The choice model proposed in this dissertation follows the consumer decision structure of Deaton and Muelbauer (1980); according to them, this structure is divided into two stages. In the first stage, one decides how to allocate the total income, y, to inside and outside goods. In the second stage, one chooses the product that one is willing to buy in the focus category, under the results of the budget allocation in the first stage. Therefore, by considering the budget allocation stage in the choice model, consumer outside goods are considered in the proposed model.

In summary, the structure of the proposed model in this dissertation—which consists of the budget allocation stage and the product choice stage—is called the multi-stage and multi-category discrete-continuous choice model, a schematic of which is shown in Figure 1.

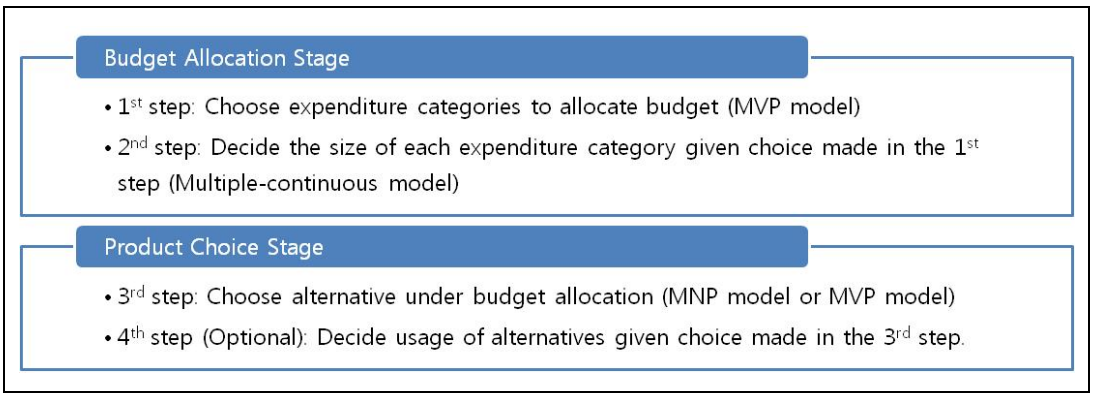

Figure 1. Conceptual Framework of the Consumer Decision Process

Under the multi-stage and multi-category case suggested in this dissertation, the proposed choice model—which also considers the budget allocation for outside goods—could be applied to various products from the IT or energy industries, among others. By considering the budget allocation for outside goods in the choice model, the proposed model could possibly assist in analyzing five factors in addition to consumer preference vis-à-vis product price and the attributes of inside goods, both of which are normally analyzed by the previous model: (1) the relationship between the choice of budget category for outside or inside goods in the budget allocation stage and the product choice for inside goods in the product choice stage, (2) the relationship between the choice of budget category for outside or inside goods in the budget allocation stage and the product usage for inside goods in the product choice stage, (3) the relationship between the choice of the budget size for outside or inside goods in the budget allocation stage and the product choice for inside goods in the product choice stage, (4) the relationship between the choice of the budget size for outside or inside goods in the budget allocation stage and

the product usage for inside goods in the product choice stage, and (5) the marginal rate of substitution (MRS) among the budget categories.

Therefore, by way of the estimation results derived from the proposed model, it is possible for consumer purchasing behavior to be analyzed more accurately, as one can comprehensively analyze the impact of the change of the allocated budget between inside goods and outside goods on product choice for inside goods. Moreover, the estimation results could be used to underscore policy implications vis-à-vis pricing strategy, product strategy, and the like. An outline of the proposed model for this dissertation is shown in Figure 2.

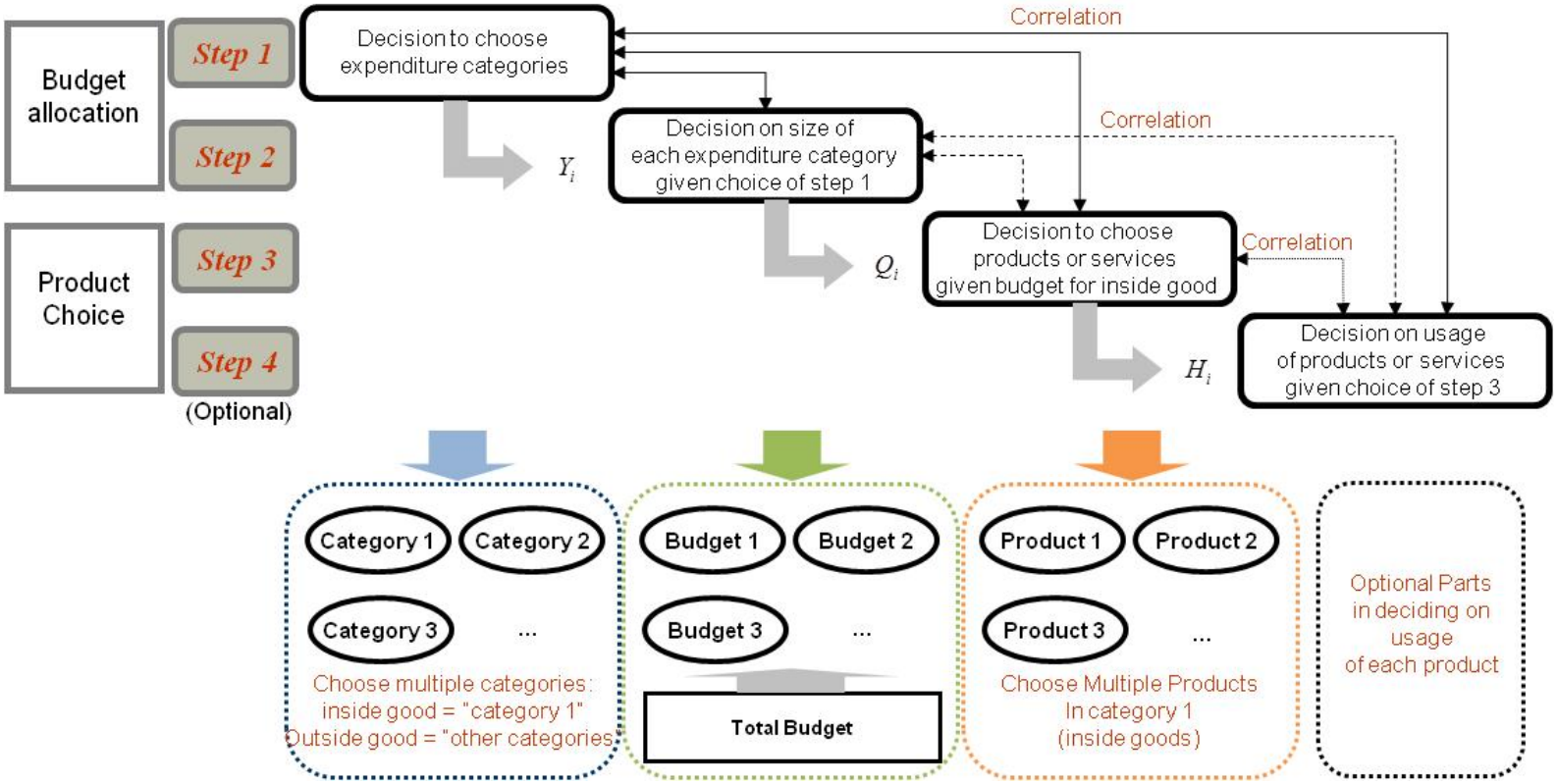

Figure 2. Outline of Proposed Model in the Current Study

The composition of this dissertation is as follows. From the perspectives of the multi-stage and multi-category models, chapter 2 reviews the previous literature concerning the choice model in order to develop a new choice model that considers the budget allocation for outside goods. By examining previous models from the perspectives of the multi-stage and multi-category models, the limitations of those previous models, as well as the direction for the development of a new choice model, are examined. Chapter 3 explains the models proposed in this dissertation. First, the research subject is summarized in section 3.1; in section 3.2, the recursive model—which has a structure similar to that in the model proposed in this dissertation—is explained, and the structures of the proposed models are then explained. Finally, a simulation study is conducted to verify the proposed models, and the implications are explained. Chapter 4 conducts empirical studies based on the choice models proposed in this dissertation; in section 4.1, the impact of the budget allocation and the product choice stages on the choice of products from the information and communications technology (ICT) industry is analyzed by using the proposed model. In section 4.2, the impact of the budget allocation and product choice stages on the choice of products from the household products industry is analyzed by using the proposed model. In section 4.3, under the assumption of a single choice and single usage situation, the effects of budget allocation, product choice, and product usage in the automobile industry are analyzed comprehensively by using the proposed model. Chapter 5 summarizes the results of this dissertation, and explains the managerial and policy-related implications from the perspective of the proposed models.

Chapter 2. Previous Literature

The discrete choice model is a useful method of estimating consumer preference (McFadden, 2000). To improve the analysis of on-demand forecasting, this dissertation suggests a new choice model that considers the budget allocation stage with outside goods. Before that, this chapter reviews the previous literature in two parts: in terms of the single/multi-stage choice model, and in terms of the single/multi-category model. First, this chapter reviews four cases of choice models (i.e., multinomial logit, MNP, MVP, and mixed logit models) in a single-stage choice situation that involves no outside goods and three cases of choice models (i.e., the two-stage choice model, consumer decision tree, and recursive model) in a multi-stage situation with no outside goods; these are used in various areas to analyze consumer preference. Second, two cases of choice models (i.e., a MDCEV model with outside goods and the BLP model) in a single category with outside goods, plus one case of a choice model (i.e., the almost ideal demand system (AIDS) model) in a multi-category with outside goods, are examined in this chapter. In addition, the limitations of previous models, as well as how those models can be improved upon, are presented. Finally, recent issues concerning the budget allocation stage within choice model are reviewed.

2.1 Choice Models with a Single/Multi-stage Model

2.1.1 Single-stage Model

Choice models from among single-stage models without outside goods describe the choice situations of decision-makers in a certain category, and analyze their preferences. Therefore, these models are used with a variety of consumer choice data, including single-choice, ranking-choice, rating-choice, and multiple-choice data. Depending on the consumer choice data used, different choice models can be used. Among the various choice models used in previous research, this section mainly considers the multinomial logit model, MNP model, random coefficient logit model (or mixed logit model), and MVP model.

The choice models considered in this section are based on the random utility model and are derived under the assumption that decision-makers choose alternatives in order to maximize their utility. When a decision-maker *n* chooses the *j* alternative from among *J* alternatives, the utility of decision-maker *n* is described as the following Eq. (1) (McFadden, 1994; Train, 2003):

(1) $U_{nj} = V_{nj} + \varepsilon_{nj}$

The first part of the utility function, V_{nj}, represents the deterministic part of the utility; the second part of the utility function, ε_{nj}, represents the stochastic part of the

utility. The stochastic part cannot be observed by the researchers. Under the assumption of utility-maximizing behavior, the choice probability that decision-maker *n* chooses the *j* alternative from among the *J* alternatives is derived as follows in Eq. (2)

$$
\begin{aligned}
P_{nj} &= \Pr ob(U_{nj} > U_{nk}, \forall k \neq j) \\
&= \Pr ob(V_{nj} + \varepsilon_{nj} > V_{nk} + \varepsilon_{nk}, \forall k \neq j) \\
&= \Pr ob(\varepsilon_{nk} - \varepsilon_{nj} < V_{nj} - V_{nk}, \forall k \neq j)
\end{aligned} \tag{2}
$$

When the joint density of the stochastic part of the utility is defined as $f(\varepsilon_n)$, the choice probability is derived by the integration over the density of the unobserved part of the utility, $f(\varepsilon_n)$. Therefore, depending on the specification of the density function of the unobserved part, a variety of choice models can be derived (Train, 2003).

Since the choice models considered in this section includes alternatives only with respect to single-stage choice situations that feature inside goods—i.e., these models do not include multiple stages of choice and outside goods—these models cannot address the effect of a characteristic change in outside goods on demand for inside goods (Chintagunta and Nair, 2010). In addition, because demand forecasting for new products normally use SP data by conducting conjoint surveys—and, furthermore, consider only single-stage choice situations that lack outside goods—consumer willingness to pay for alternatives could be overestimated (Banfi et al., 2008). Therefore, in the absence of outside goods in the model, the estimation results of consumer preference for new products could be biased.

Multinomial Logit Model

When decision-makers choose one alternative among *J* alternatives, the multinomial logit model or MNP model can be used to analyze their preferences. In the former model, the influence of each attribute on alternatives is assumed to be homogenous, and the stochastic part of the utility is assumed to have an independently and identically distributed (iid) type-I extreme value distribution. Due to the assumption relating to the stochastic part of the utility, the choice probability becomes a simple closed form, as shown as in Eq. (3):

$$(3)\quad P_{nj} = \frac{\exp(V_{nj})}{\sum_i \exp(V_{ni})} = \frac{\exp(\beta' X_{nj})}{\sum_i \exp(\beta' X_{ni})}$$

where X_{nj} is the explanatory variables observed and β is the degree of influence of each attribute on utility. Being a simple closed form, the multinomial logit model is easy to estimate. However, due to the iid assumption, the “independence from irrelevant alternatives” (IIA) property is exhibited by the multinomial logit model. Therefore, it is difficult to reflect the realistic substitution pattern. Moreover, the multinomial logit model has an unrealistic assumption: that all consumers have the same preference for a certain product. Nevertheless, the multinomial logit model is still widely used to analyze consumer preference (Matsukawa and Ito, 1998; Ewing and Sarigollu,

2000; Kim et al., 2006; Wei et al., 2005).

Multinomial Probit Model

Like the multinomial logit model, the MNP model assumes that the influence of an attribute on various alternatives is homogenous, but that the stochastic part of the utility is assumed to have a normal distribution with covariance matrix Ω. Given this assumption vis-à-vis the stochastic part of the utility in the MNP model, this model can reflect how the correlation relationship between alternatives and the IIA property can be avoided (Train, 2003). Under the assumption regarding the stochastic part of utility, the choice probability is shown as in Eq. (4):

$$(4)\quad P_{nj} = \int I(\varepsilon_{nk} - \varepsilon_{nj} < V_{nj} - V_{nk}, \forall k \neq j)\phi(\varepsilon_n)d\varepsilon_n$$

where $\phi(\varepsilon_n)$ is defined as the density function of ε_n. Unlike with the multinomial logit model, the choice probability of the MNP model is not a simple closed form; therefore, the simulation approach or some other estimation method is needed to estimate the MNP model. Recently, the simulation approach and the Bayesian estimation method have been variously used with the MNP model. Although the MNP model has a computational burden that needs to be estimated, various research fields are using it to analyze consumer preferences (Lee and Kwak, 2007; Garrido and Mahmassani, 2000;

Kim et al., 2003; Jepsen, 2008; Chintagunta, 1992).

Random Coefficient Logit Model or Mixed Logit Model

Although the MNP model could reflect the IIA property, it cannot reflect the random taste variation, which is also referred to as "heterogeneity in consumer preference." The random coefficient discrete choice model, also called the mixed logit model, can accommodate the heterogeneity of consumer preference by incorporating a stochastic term into each coefficient. In the mixed logit model, based on random utility theory, it is assumed that each consumer *i* has his or her own utility function for each product or service *j* in choice set *t*. The utility function is stated as follows in Eq. (5):

(5) $U_{ijt} = V_{ijt} + \varepsilon_{ijt} = \sum_k \beta'_{ik} X_{jkt} + \varepsilon_{ijt}$

As shown in Eq. (5), the utility function distinguishes the effects of deterministic factors, V_{ijt}, from those of random factors, ε_{ijt}. The deterministic part consists of the marginal utility of each attribute k of the alternative j, β_{ik}, and the vector of each attribute, X_{jkt}. Unlike with Eq. (1), the mixed logit model allows for the stochastic nature of marginal utility, β, so as to assume the heterogeneity of each consumer preference. In this model, β_{ik} is set to a vector that follows the multivariate normal

distribution with the meaning of b_k and the covariance matrix of Σ_k —i.e., $\beta_{ik} \sim N(b_k, \Sigma_k)$—and ε_{ijt} is assumed to be a random disturbance following a type-I extreme value distribution.

The mixed logit model has an advantage: it can set up the distribution of each attribute's coefficient, based on the impact of certain attributes on consumers (Train, 2003). To reflect consumer heterogeneity, each attribute's coefficient is generally assumed to be in a normal distribution. However, because the normal distribution is not suitable for certain situations—i.e., when all consumers have the same direction of preference—the distribution of these coefficients needs to assume another distribution (Train and Sonnier, 2005). When each coefficient is allowed to have a specific distribution that reflects consumers' general preference structures, consumer preference can be analyzed more accurately. The likelihood function of consumer *n* who has an *n*-times choice situation can be denoted as follows, in Eq. (6):

$$(6) \quad L(d_n | \beta_n) = \left(\prod_{t=1}^{T} \frac{e^{C(\beta_n)' x_{jt}}}{\sum_{k=1}^{J} e^{C(\beta_n)' x_{kt}}} \right)$$

where d_n indicates that each consumer *n* chooses *T* times among a total of $T \times J$ alternatives in each choice set. If β_n has the density function $f(\beta)$, the choice probability in the mixed logit model is that Eq. (6) is integrated over the density function

of β, $f(\beta)$. Because the mixed logit model does not have a simple closed form that is similar to that of the MNP model, it is too complicated to estimate each parameter by classical methods, e.g., maximum likelihood estimation. Therefore, a simulation-based estimation approach is needed (Brownstone and Train, 1999; Calfee et al., 2001; Layton, 2000). Otherwise, a Bayesian estimation method is used, because Bayesian estimation bears some advantages: it does not require the complicated calculation of the integration of multivariate density function, and it overcomes both the initial point problem and the global optimal solution problem (Edwards and Allenby, 2003; Allenby and Rossi, 1999; Huber and Train, 2001). Moreover, the result of Bayesian estimation can also be converted to a classical estimation result (Train and Sonnier, 2005).

Multivariate Probit Model

In many situations, consumers choose multiple alternatives concurrently. The multivariate discrete choice model have been developed to analyze such cases; the MVP model is one such model (Baltas, 2004; Edwards and Allenby, 2003; Chib and Greenberg, 1998). Although they can be applied to multiple-choice data (Hausman and McFadden, 1984; Boztug and Hildebrandt, 2006), multivariate logit models cannot derive the complementary or substitutive relationships among and within product categories, on account of the IIA property. As the MVP model does not assume the presence of the IIA property, the analysis of complementary or substitutive relationships among various

alternatives proves useful. Therefore, the utility function of consumer *i* for alternative *j* in this model is as follows:

$$(7)\quad \begin{aligned} U_{ij} &= \beta' X + \varepsilon_{ij} = \gamma_j + \sum_d \beta'_{jd} S_{id} + \varepsilon_{ij} \\ Y_{ij} &= \begin{cases} 1 \ if \ \ U_{ij} > 0 \\ 0 \ if \ \ o.w \end{cases} \end{aligned}$$

where γ_j is called the alternative specific constant (ASC) of each alternative *j*. S_{id} is the socio-demographic variable for each socio-demographic indicator *d*, such as gender, age, and income level. In the MVP model, the disturbance ε_{ij} is assumed to follow the multivariate normal distribution with a zero mean and the variance–covariance matrix Ω. The choice probability that consumer *i* chooses the multiple alternatives is as follows:

$$(8)\quad P(Y_i \,|\, \beta, \Omega) = \int \ldots \int \phi_J \left(\varepsilon_{i1}, \ldots, \varepsilon_{iJ} \,|\, 0, \Omega \right) d\varepsilon_{i1}, \ldots, d\varepsilon_{iJ}$$

where $\phi_J \left(\varepsilon_{i1}, \ldots, \varepsilon_{iJ} \,|\, 0, \Omega \right)$ is the *J*-variate normal density function with a zero mean and the variance–covariance matrix Ω. Because the choice probability of the MVP model does not allow a simple closed form, a simulation-based estimation method or Bayesian estimation method is needed to estimate consumer preference. Given the

advantage of the MVP model, it has been used in a variety of fields (Edwards and Allenby, 2003; Manchanda et al., 1999; Rao and Winter, 1978; Seetharaman et al., 2005).

2.1.2 Multi-stage Model

2.1.2.1 Two-stage Choice Model

When consumers encounter the decision-making process in the process of buying products, they make a choice among several alternatives; indeed, their decision-making process has more than a single stage (Bettman, 1979). The early two-stage choice model of Gensch (1987) assumes that consumers undergo a two-stage decision process, as follows. In the first stage—the elimination stage—the consumer reduces all feasible alternatives to only several alternatives. In the second stage, or the decision stage, the consumer chooses the most preferred alternative among the several alternatives seen in the first stage. These two-stage decision processes have been explained through the use of trading methods that involve information display boards, verbal protocols, and other instruments (Bettman and Park, 1980; Lussier and Olshavsky, 1979; Olshavsky, 1979; Payme, 1976). Gensch (1987) empirically shows that previous purchase behavior theory vis-à-vis a consumer's two-stage choice-making process is valid.

The consumer's sequential decision-making process has been variously outlined, because the multi-stage choice process provides more information for practitioners of management and policy than does the single-stage choice process. In particular, Bettman

and Park (1980) mention that it is important to consider prior cognition and experience in the choice process, to provide information to marketers and policy-makers at various levels. In other words, results from the multi-stage choice-making process show how consumer knowledge and experience affect each decision stage, and they efficiently identify which information is the most important in selecting a purchase alternative at the choice stage. Shocker et al. (1991) mention that if consumers face complex decision-making situations, they use processes that feature more than two stages. From this perspective, a sequential decision-making process can be elucidated from the two-stage choice process to a variety of models, such as the three-stage choice process.

The consumer sequential decision-making process starts from the consumer perspective, i.e., that the consumer's choice set changes as a function of the choice situation (Hauser and Wernerfelt 1989, 1990; Lehmann and Pan, 1991; Nedungadi, 1990; Ratneshwar and Shocker, 1991; Robers and Lattin, 1991; Shocker et al, 1991; Simonson and Tversky, 1992). This means that depending on the choice situation at hand, with a complex decision-making process, the consumer's choice set changes so as to reduce the complexity of the choice situation. Kardes et al. (1993) consider three-stage decision-making processes as sequential decision-making processes, rather than two-stage processes. In the first stage, one reduces a universal set that includes all available alternatives to a marketplace purchase to a retrieval set that includes only possible alternatives to that purchase; this means that consumers are able to directly access alternatives, guided by prior cognition or memory. In the second stage, one reduces the

retrieval set to a consideration set that includes several alternatives, from which one chooses the most preferable alternative. In the third stage, which involves the decision-making stage, one chooses a final purchase alternative among the consideration set. They suggest a multi-stage decision process by using a sequential logit model. Based on the three-stage decision processes suggested by Kardes et al. (1993), Grewal et al. (2003) analyze the consumer decision-making process in terms of brand, by considering word-of-mouth and similarity.

With the development of the internet, the impact of consumer information searches on consumer purchase behavior has increased, especially with regard to online shopping. From this viewpoint, Teo and Yeong (2003) analyze consumer purchase behavior in an online shopping environment in Singapore. In addition, unlike the aforementioned literature—which analyzes the multi-stage decision process while focusing on choice set—they define the multi-stage decision process by using the Engel, Blackwell, and Miniard (EBM)[1] model. In their model, the consumer decision-making process is divided into three stages: the information search stage, the alternative evaluation stage, and the purchase stage. Based on a three-stage consumer decision process, consumer purchase behavior in the online shopping environment is analyzed empirically by using a structure equation. Additionally, Kohli et al. (2004) analyze the consumer decision-making process in the online shopping market by adopting Simon's

[1] Based on the Engel, Kollat, and Blackwell (EKB) model, which was first proposed by Engel et al. (1978), the EBM model was developed by Engel et al. (1995). The EBM model divides the consumer decision-making process into five stages as follows: need recognition, information search, alternative evaluation, purchase, and after-purchase evaluation.

decision-making model;[2] moreover, they add a time-saving/cost-saving and satisfaction stage to the decision-making process, and then undertake empirical analysis through the use of a structure equation.

Except for simple choice situations, consumers decide to purchase products after going through a multi-stage decision process, rather than a single-stage decision process. Thus, a multi-stage choice model that considers a consumer's sequential decision-making processes provides more information about consumer purchasing behavior than a single-stage choice model that assumes only a simple decision process. Moreover, the results of using a multi-stage decision process can inform various managerial strategies and policy suggestions. Therefore, because more information about consumer purchasing behavior is provided through the consideration of a multi-stage decision process, it is necessary to develop a new methodology with regard to consumer choice models that consider the multi-stage decision process.

2.1.2.2 Consumer Decision Tree

To consider the consumer multi-stage decision process, some previous studies have suggested hierarchical choice models (Lehmann and Moore, 1986). A hierarchical choice model is defined as a consumer decision-making process with a decision tree structure that assumes that consumers compare the attributes of products sequentially. Consumer

[2] Simon's decision-making model divides the consumer decision process into three stages: the intelligence, design, and choice stages.

choice behavior is analyzed in line with this consumer decision tree, which is assumed by researchers. In other words, consumers compare the attributes of products from the upper level to the lower level of the decision tree, and then they decide which products they will purchase. Thus, consumers make choices by using a sequential decision-making process from the upper level to the lower level of decision tree, while eliminating unsatisfactory attributes as per the tree structure (Currim, 1982; Dubin, 1986; Lehmann and Moore, 1986). Representative models that analyze the consumer decision-making process by using a decision tree structure are Tversky and Sattah's (1979) PRETREE model and McFadden's (1986) nested logit model.

However, most hierarchical choice models that adopt a decision tree structure have a limitation in how they define how individual consumers decide brand choice (Currime et al., 1988). In other words, the consumer decision-making process is analyzed while assuming that all consumers bear the same decision tree structure. Because most of the previous hierarchical choice models are established under the unrealistic assumption that there is consumer decision tree homogeneity, it is desirable to establish a new methodology that suggests how a hierarchical decision tree is constructed.

From this viewpoint, based on data mining methodology, research on how to construct a consumer decision tree can be conducted. To analyze the structure of decision trees, various algorithms can be developed to automatically classify consumers as a homothetic group, based on individual-level data. For instance, representative algorithms include automatic interaction detection (AID; Sonquist et al., 1971), chi-square AID

(CHAID; Kass, 1980), classification and regression trees (CART; Breiman et al., 1984), ID3 (Quinlan, 1986), and the concept learning system (CLS; Currim et al., 1988), among others. These decision tree methods are used to create target marketing strategies, by classifying a group of consumers based on their consumer choice data. A variety of previous methodologies that classify consumer groups can be used to determine the structure of a hierarchical choice model (Currim et al., 1988).

According to Roberts and Nedungadi (1995)—who review the literature on the consumer decision-making process—a multi-stage decision process should be considered to understand the consumer behavior process more accurately. In addition, it is important to consider a multi-stage decision process when defining the consumer consideration set and analyzing consumer purchasing behavior. Therefore, to better understand managerial and policy implications—by first understanding the structure of the consumer decision-making process—it is necessary to develop a model development from the perspective of the hierarchical choice model.

2.1.2.3 Recursive Model

When consumers decide to buy products or services, consumer decision-making undergoes a multi-stage process. Therefore, an analysis of consumer choice probability through a single equation that considers only final choice stage is limited in its ability to provide highly accurate demand forecasts. For instance, Gruber and Owings (1996) analyze the choice probability of caesarean section delivery by using a single equation,

but Febbri et al. (2004) suggest that results from a multi-stage process that includes both hospital type as a preceding choice and caesarean section delivery as a second choice are more accurate than the results from Gruber and Owings (1996). Because estimation results could be biased and inconsistent when one does not consider the preceding choice stage of the consumer decision-making process, the preceding choice stage should be included to ensure more accurate analysis (Marra and Radice, 2011).

From this perspective, simultaneous equation models with both continuous and discrete endogenous variables are introduced by Heckman (1978) and Amemiya (1978).[3] The bivariate probit model with an endogenous dummy (or recursive bivariate probit model), and simultaneous equation models with limited dependent variables (SLDV) are representative models. Both the recursive bivariate probit model and the SLDV model consider the one-way causal relationship between preceding and subsequent choices (Ye et al., 2007).

A recursive bivariate probit model uses the result of the first decision in the bivariate probit model as an endogenous dummy variable in the second equation of the bivariate probit model. Because the consumer decision-making process is affected by the consumer's preceding choice, a recursive bivariate probit model considers both preceding and subsequent choices, in tandem. A recursive bivariate probit model has been applied in various fields to analyze choice probability more accurately; for instance, Baslevent and

[3] Simultaneous equation models with a limited dependent variable is also introduced by Blundell and Smith (1989, 1994), Nelson and Olson (1978), Rivers and Vuong (1988), Smith and Blundell (1986), and Vella (1993).

El-hamidi (2009) use a recursive bivariate probit model to analyze the relationship between early retirement as a preceding choice and the post-retirement employment decision as a subsequent choice, in the Egyptian economy. Their results show that people who would like to work after retirement have a greater tendency to choose early retirement. In addition, Marra and Radice (2011) analyze the relationship between a woman's education as a preceding choice and fertility as a subsequent choice, through the use of a semiparametric inference method.

Additionally, health economy studies have used a recursive bivariate probit model to analyze the impact of supplemental insurance ownership on healthcare demand (Holly et al., 1998; Buchmueller et al., 2005) and the impact of self-reported disability measurements on social benefits (Benitez-Silva et al., 2004), *inter alia*. In addition, Latif (2009), Kawatkar and Nichol (2009), Hewett et al. (2008), Goldman et al. (2001), and others use a recursive bivariate probit model to undertake research in health economy, while researchers in law and economics (Deadman and MacDonald, 2004) use this model to analyze the relationship between criminal behavior as a preceding choice and victimization as a subsequent choice.

Similar to the recursive bivariate probit model, the SLDV model is derived by using the first choice from the bivariate probit model as an endogenous dummy variable in the second equation of the tobit model; as such, the SLDV model considers the impact of the preceding choice on the subsequent choice. In addition, the SLDV model has been used to consider the consumer multi-stage choice situation, in a variety of fields. Along

with various application studies, research on estimation methods that use the SLDV model has been conducted. Blundell and Smith (1989) suggest a two-stage algorithm to estimate the SLDV model. Li (1998) mentions that a full information maximum likelihood (FIML) estimator incurs computational difficulties, unlike the two-stage algorithm suggested by Blundell and Smith (1989), but that an FIML estimator is more efficient than a two-stage algorithm, as an SLDV model-based estimation method. However, Li (1998) suggests a Bayesian estimation method to overcome such computational difficulties of FIML estimator.

Blundell and Smith (1989) explain that consumers are faced with a joint problem vis-à-vis the simultaneous choice situation: consumers simultaneously choose several dependent variables in an unobserved and sequential decision-making process. Therefore, consideration of the impact of a preceding choice on a subsequent choice is important to analyzing consumer choice probability within a consumer choice problem more accurately. From this viewpoint, model development with considering a multi-stage structure is essential to demand analysis and product forecasting.

2.2 Choice Models with a Single/Multi-category Model

Economically, outside goods constitute an important factor in the demand function, when determining changes to a total category of demand as a function of net price changes (Chintagunta and Nair, 2010). From this viewpoint, some researchers have sought to

consider outside goods in their single/multi-category choice models. In this section, I first review two cases of choice models that include outside goods in a single-category model: (1) MDCEV model with outside goods, and (2) the BLP model. Second, I also consider one case that includes outside goods in a multi-category model. In addition, limitations therein and what elements could be improved are examined.

2.2.1 Single-category Model

2.2.1.1 Multiple Discrete-Continuous Extreme Value Model

The MDCEV model is first proposed by Kim et al. (2002), and further developed by Bhat (2005, 2008) and Bhat and Sen (2006). The MDCEV model is advantageous when considering multiple-choice behavior and product usage simultaneously. There are multivariate logit models and MVP models (Baltas, 2004; Edwards and Allenby, 2003), both of which are also used to consider multiple-choice situations. However, these models bear several limitations: both the multivariate logit and MVP models consider only multiple-choice behavior, and the multivariate logit model has a limitation relating to the IIA assumption.

In addition, the MDCEV model can be used not only to consider consumer choice behavior, but also to analyze additional utility arising from product use. Additional utility derived from product use follows the law of diminishing marginal utility of consumption, which states that additional utility gradually decreases as the usage

increases; the MDCEV model, meanwhile, considers additional utility derived from product use under the law of diminishing marginal utility, as well as utility from choice and multiple-choice situations.

The MDCEV model first proposed by Bhat (2005, 2008) converts the choice probability into a closed form by assuming the disturbance term to be the standard extreme-value distribution. In addition, because consumer preferences differ depending on individual consumer characteristics, consumer heterogeneity should be considered in the model. A mixed MDCEV model can be used to consider consumer heterogeneity by assuming a distribution for each parameter in the MDCEV model (Ahn et al., 2008; Bhat, 2005, 2008).

Under the MDCEV model, if the ith consumer chooses j alternatives among K alternatives, and if i^{th} consumer uses m_j for each J alternative and if there are no outside goods, the utility of the ith consumer is shown as follows:

$$(9)\quad U_i(m_1,...m_J,0,...,0) = \sum_{j=1}^{K} \Psi(x_j)(m_j+\gamma)^{\alpha_j}$$

In Eq. (9), K represents the number of alternatives that exist in an alternatives set. $\Psi(x_j)$ represents the baseline utility from selecting the jth alternative. x_j implies the attributes that comprise the jth alternative, and m_j implies the usage of the jth alternative. γ is a parameter used to determine whether an interior or corner solution will be found;

if the condition satisfies $\gamma \neq 0$, a corner solution can exist, because the jth alternative may not be used. However, if the condition satisfies $\gamma = 0$, an interior solution always exists, because the usage of all alternatives is greater than 0 (Bhat, 2005, 2008; Kim et al., 2002). α_j is a satiation parameter that affects the degree of diminishing marginal utility; therefore, to satisfy the law of diminishing marginal utility, α_j has a value ranging from 0 to 1. For this reason, α_j is defined as $\alpha_j = 1/(1+\exp(-\delta_j))$ (Bhat, 2005).

The baseline utility, $\Psi(x_j)$, is defined as an exponential function, because it always has a positive value. Consumers choose optimal alternatives, so as to maximize their utility under budget constraints. Therefore, budget constraints must be considered in any model used to analyze consumer optimal behavior. If the budget constraint is about rate, m_j represents the usage fee for the jth alternative, and M is the total budget amount. Consequently, the utility maximization problem—based on the budget constraint shown—can be solved through the use of the Lagrangian method and the Kuhn–Tucker condition. The result shows consumer optimal choice and usage as in Eq. (10). At this point, Bhat (2005) expresses the closed form for choice probability, by assuming the disturbance term is the standard extreme-value distribution.

$$(10)\quad P\left(m_2^*, m_3^*, ..., m_J^*, 0, 0, ..., 0\right) = \left[\prod_{i=1}^{J} c_i\right]\left[\sum_{i=1}^{J} \frac{1}{c_i}\right]\left[\prod_{i=1}^{J} e^{V_i} \Big/ \left(\sum_{j=1}^{K} e^{V_j}\right)^J\right](J-1)!$$

where, $c_i = \left(1 - \alpha_i / m_i^* + \gamma\right)$, $V_j = \beta' x_j + \ln \alpha_j + (\alpha_j - 1)\ln(m_j^* + \gamma)$

In Eq. (10), J represents the number of alternatives. If a consumer chooses only one alternative (J = 1), Eq. (10) converges to the multinomial logit model. If the MDCEV model assumes the distribution in each parameter—as is done to consider the heterogeneity of consumer preference—the mixed MDCEV model can be shown (Ahn et al., 2008). The probability of a mixed MDCEV model is shown as follows:

$$(11)\quad \tilde{P}_n = \int \left\{ \left[\prod_{i=1}^{I} c_i \right] \left[\sum_{i=1}^{I} \frac{1}{c_i} \right] \left[\prod_{i=1}^{I} e^{V_i} \Big/ \left(\sum_{i=1}^{I} e^{V_j} \right)^{I} \right] (I-1)! \right\} f\left(v_n \mid \theta\right) dv_n$$

To incorporate outside goods into the model, Bhat (2008) assumes that outside goods have a unit price as a first good and are identified as $\Psi(x_1, \varepsilon_1) = \exp(\varepsilon_1)$. Then, additional utility from outside goods is added to the utility functional form in Eq. (9). The choice probability of the MDCEV model with outside goods is similar to that of the MDCEV model without outside goods. Bhat (2008) analyzes only the difference in the MDCEV model with outside goods, between when the satiation parameter has a restriction and when it does not. In addition, Bhat (2008) considers only the single category in multiple-choice situations; therefore, there is no significant difference between the models with and without outside goods. However, the MDCEV model has been used in many fields to analyze both choice and usage (Bhat and Sen, 2006; Ahn et

al., 2008).

2.2.1.2 BLP Model

Most studies on consumer preference have been conducted through the use of individual-level data, but the BLP model (Berry, Levinshon, and Pakes, 1995) uses market share data to analyze consumer preference structure at the market level. In particular, the BLP model shows a realistic substitution pattern among alternatives by considering outside goods in a single category. In addition, endogeneity with regard to price creates bias when estimating the price response parameter (Besanko et al., 1998); however, Berry, Levinshon, and Pakes (1995) suggest a method of resolving this endogeneity problem—a problem that arises from the correlation between price and the unobserved attributes of products. In other words, they use instrument variables to determine the nature of the endogeneity problem. Moreover, when only the distributions of consumer demographics are accessible, Nevo (2001) reflects this information in the model structure to analyze the effect of consumer demographic variables.

This section reviews the random coefficient model by demand which uses market-level data. The random coefficient model is a type of multinomial choice model that involves more than two alternatives and assumes a single-choice situation. The alternatives are interchangeable, and outside goods are included as available alternatives. Outside goods come into play when consumers do not choose one of the available

alternatives from the inside goods.

Based on the random utility model, if the ith consumer chooses j alternatives in market t, the utility of the ith consumer, u_{ijt}, is shown as follows:

$$(12)\quad \begin{aligned} & u_{ijt} = \alpha_i (y_i - p_{jt}) + x_{jt}\beta_i + \eta_{jt} + \varepsilon_{ijt} \\ & i = 1,...,I, \quad j = 0,...,J, \quad and \quad t = 1,...,T, \quad respectively \end{aligned}$$

where "j = 0" indicates that the consumer has chosen outside goods. In Eq. (12), x_{jt} represents the observed attributes of the alternatives and η_{jt} represents the unobserved attributes of the alternatives. y_i is the income of the ith consumer and p_{jt} is the price of alternative j in market t. Finally, ε_{ijt} represents a stochastic term with zero mean and follows an identical independently distributed assumption.

To reflect consumer heterogeneity in each coefficient of product attributes, the distribution of each coefficient is assumed to have a multivariate normal distribution as follows:

$$(13)\quad \begin{pmatrix} \alpha_i \\ \beta_i \end{pmatrix} = \begin{pmatrix} \alpha \\ \beta \end{pmatrix} + \begin{pmatrix} \Sigma_\alpha \\ \Sigma_\beta \end{pmatrix} (v_{i\alpha}, v_{i\beta}), \quad v_i \sim P_v^*(v)$$

where v_i represents the random taste variable. After substituting Eq. (13) into

Eq. (12), the utility function u_{ijt} is divided into two parts, as seen in Eq. (14): one is the mean utility from alternatives, and the other is heterogeneous utility that depends on consumer taste.

$$
\begin{aligned}
u_{ijt} &= \alpha_i (y_i - p_{jt}) + x_{jt}\beta_i + \eta_{jt} + \varepsilon_{ijt} \\
&= \alpha_i y_i + (-\alpha p_{jt} + x_{jt}\beta + \eta_{jt}) + (p_{jt}, x_{jt})(\mathbf{\Sigma} v_i) + \varepsilon_{ijt} \qquad (14) \\
&= \alpha_i y_i + \delta_{jt} + u_{ijt} + \varepsilon_{ijt}
\end{aligned}
$$

where δ_{jt} represents mean utility, which is the same for all consumers. By following the structure of Eq. (14), if consumers choose outside goods, the utility function of consumers is as follows in Eq. (15):

$$u_{iot} = \alpha_i y_i + \eta_{ot} + \mathbf{\Sigma}_0 v_{i0} + \varepsilon_{i0t} \qquad (15)$$

Because the multinomial choice model assumes that consumer choice is determined by the utility difference among alternatives, the utility of outside goods is normalized to zero, for the sake of simplicity (Berry, Levinshon, and Pakes, 1995). When the distribution of the stochastic term is assumed to follow a type-I extreme value distribution, the market share of product *j* in market *t* is as follows:

$$(16)\quad s_j = \int_t \int_v \left[\frac{\exp(\delta_{jt} + \mu_{ijt})}{1 + \sum_{k=1}^{J} \exp(\delta_{kt} + \mu_{ikt})} \right] dP_v^* dt$$

The estimation method for the BLP model is a generalized method of moments (GMM) that uses instrument variables to solve an endogeneity problem. The instrument variable *Z* is used because price correlates with unobserved parts on the manufacturer's side; for example, various attributes affect demand for a product, but researchers cannot observe or measure them. On the other hand, producers can observe and know of all the product attributes, including those unobservable from the researcher's viewpoint. Therefore, producers reflect the cost of all price attributes, but researchers tend only to consider several certain attributes, with the remaining attributes being included in the error term. Thus, researchers tend to adopt an instrument variable to classify the endogenous effect between unobserved parts and price (Berry, 1994). On the other hand, Nevo (2001) argues that although a simultaneous consideration of both the demand and supply sides shows highly efficient estimation results, these methods increase computational and programming complexity. Therefore, Nevo (2001) points out that a consideration of only the demand side is also sufficient in estimating reasonable parameters.

Although the BLP model shows realistic substitution patterns among alternatives by considering outside goods, this model still has a limitation resulting from the inclusion

of only outside goods within a single category. If the expenditure of outside categories is increased by changing the budget allocation, the demand for inside goods changes. Because this situation does occur in real life, model development is needed to consider outside goods in a multi-category situation. In any case, the BLP model has been used in a variety of fields (Nevo, 2001; Sudhir, 2001; Train and Winston, 2007).

2.2.2 Multi-category Model

2.2.2.1 Almost Ideal Demand System Model

When consumers decide to buy a product, they consider not only single-category but also multi-category choice situations. From this perspective, models for use in analyzing the demand of consumption categories on the macroeconomic side have been developed and considered. In other words, a change in product demand as a function of the price changes within each category has been analyzed through the use of aggregate consumption data and demand equations that derive from consumer theory. In the early research, the linear expenditure system (LES) model has been used to conduct empirical analysis vis-à-vis demand theory (Goldberber and Gamaletsos, 1970; Parks, 1969; Pollak and Wales, 1969). However, due to a limitation of the LES model—i.e., it excludes inferior goods from its explanation—the Rotterdam model, the translog model, and the AIDS model, among others, are suggested as new models.

Among the various models that analyze demand structure by considering a

multi-category situation, the AIDS model (Deaton and Muellbauer, 1980) has been applied in a variety of industries. For instance, by using the AIDS model, Chang et al. (2002) analyze the demand for wine; Rolle (1997), the demand for railroad travel; Edgerton (1997), the demand related to the consumption of nondurable consumer goods; and Syriopoulos (2002), the demand for financial instruments such as stocks, bonds, and mutual funds. In addition, Verbeke and Ward (2001) analyze the impact of advertisements via mass media on the consumption of fresh meat; and Duffy (2003), the impact of advertisements on the consumption of beer, wine, cigarettes, and other goods, by using the AIDS model. Cotterill and Putsis (2000) use the AIDS model to analyze the impact of brand on consumption, while Tiezzi (2002) uses it to analyze the impact of the environmental protection fee—a kind of tax—on the structure of consumer consumption. Clearly, the AIDS model has been applied to the analysis of demand in multi-category situations in a variety of fields.

The AIDS model as suggested by Deaton and Muellbauer (1980) has a theoretical advantage over the Rotterdam model and the translog model; it also has an advantage over the LES model, in that it is easy to use. The AIDS model assumes that consumers take a two-stage decision process, as follows. In the first stage—i.e., the budget allocation stage—total income, y, is allocated to inside and outside goods, in that order. In the second stage—i.e., the decision stage—the consumer decides the demand of each category, under the budget constraint for each category. Deaton and Muellbauer (1980) assume that only an income effect and price elasticity impact the price change in

outside goods on the demand for inside goods. Moreover, all alternatives in outside goods are assumed to have a substitution or complementary relationship with inside goods. Under these assumptions, the AIDS model is not a single-equation demand model in each category; it has the structure of a complete demand system, which explains not only the relationships among demand items but also the expenditure allocation problem among all items in consumption categories, by compensating for the weak point of a single equation.

The AIDS model derives from the cost minimization problem of the consumer cost or expenditure function. By using the shepherd lemma—with a partial derivative of the expenditure function with respect to the price of each alternative equaling the amount of expenditure for each alternative—the demand model is shown as follows:

$$w_i = \alpha_i + \sum_{j=1}^{n} \gamma_{ij} \ln p_j + \beta_i \ln(y / P^*) \tag{17}$$

$$\ln P^* = \sum_{k=1}^{n} w_k \ln p_k$$

where, w_i is the expenditure share of the ith category,

p_i is the price of the ith category,

y is the total expenditure, and P is the price index.

Based on the estimation results from Eq. (17), the price elasticity, cross-price elasticity, and income effect can be analyzed. Carpentier and Guyomard (2001) assume

that consumers make a decision to maximize their utility under two-stage budgeting,[4] which resembles the consumer decision-making structure of Deaton and Muellbauer (1980). Under two-stage budgeting—an extension of the AIDS model—they analyze the income effect and cross-price elasticity between budget allocation in the first stage and expenditure in the second stage by using a true cost-of-living (TCOL) price index and a quantity index for budgeting the broad group in the first stage.

Thus, while the AIDS model has been applied to analysis of the consumer demand structure in a variety of fields, at an aggregate level, it analyzes only price elasticity and income effect among income categories. Therefore, a consumer decision model that simultaneously considers both a multi-category choice situation and the attributes of products other than price should be developed. However, the AIDS model continues to be used in a variety of industries to analyze the consumer demand structure from a macroscopic viewpoint.

2.3 Limitation and Recent Issues considering the Budget Allocation Stage in the Choice Model

A consideration of the consumer budget allocation stage in the choice model is similar to the two-stage budgeting suggested by Deaton and Muelbauer (1980), which allocates a budget to each category in the first stage and chooses an alternative under budget

[4] In the first stage, the total expenditure is allocated to the broad item-groups. In the second stage, the expenditure of the broad item-groups is allocated to the elementary commodities.

allocation limitations in the second stage. In other words, a choice model that considers the budget allocation stage includes both multi-stage and multi-category perspectives. From those perspectives, various models have been developed to reflect the consumer choice-making process. Especially for the multi-stage perspective, Gensch (1987) mentions that estimation results from a multi-stage choice model provide greater managerial insight than most estimation results derived from a single-stage approach. From this viewpoint, a two-stage choice model, a consumer decision tree model, recursive models, and the like have been developed and are being applied in a variety of fields. However, choice models from multi-stage perspectives consider only alternatives within the target category and are unconcerned with outside categories.

The inclusion of an outside category in a choice model is similar to a consideration of the budget allocation stage from the consumer side of the consumer decision-making process. This means that a consideration of outside goods is similar to a consideration of the budget allocation stage, from an individual-level perspective. Previous models that include outside goods have been developed fragmentarily under single-category perspectives. Although most existing studies actually recognize the importance[5] of a consideration of outside goods in a choice model, their empirical models consider only inside goods, and they analyze consumer preference through the use

[5] Berry, Levinshon, and Pakes (1995) and Nevo (2001) mention that if outside goods aren't included in the choice model, then the choice problem becomes a simple choice problem for inside goods. However, generally, when price increases within a choice model that includes only inside goods, the problem occurs whereupon aggregate output is not reduced. Chintagunta and Nair (2010) explain that an economic specification of outside goods in the demand function is important to analyzing the change in the total category demand, depending on the price change.

of these models. Most of the previous literature that considers outside goods in discrete choice models defines “outside goods” as no purchase (Ackerberg, 2003; Goolsbee and Petrin, 2004; Berry and Haile, 2009; Bhat, 2008). For instance, in the MDCEV model suggested by Bhat (2008), the utility function includes outside goods, each of which is assumed to have a unit price. However, a choice model that includes outside goods cannot analyze the relationship between outside and inside goods, and so analysis is limited to a single category. Thus, a choice model that defines “outside goods” as embodying a no-purchase option is limited in its ability to analyze the impact of price changes in outside goods and the impact of changes at the attribute level on the demand of inside goods.

On the other hand, some previous studies define “outside goods” as a gap between total market share and the market share of inside goods (Kim et al., 2005; Berry, Levinsohn, and Pakes, 1995). In other words, outside goods represent “(Market share of outside goods) = 1 – (Sum of market share of inside goods).” For example, based on market-level data, Berry, Levinshon, and Pakes (1995) consider outside goods in their choice model and analyze a realistic substitution pattern among choice alternatives. However, Berry, Levinshon, and Pakes (1995) assume that all alternatives are interchangeable within a single category, and they define “outside goods” as embodying a no-choice option; this means that consumers do not choose an alternative among the inside goods. Therefore, when the attribute level in another category changes, the impact on a particular category (including inside goods) is not analyzed. This means that an analysis of consumer purchase behavior is needed, based on a choice model with

considering outside goods under multi-category perspectives. Allenby et al. (2004b) define “inside goods” as a product category from which consumers are willing to buy, and “outside goods” as a product category that has a substitution or complementary relationship with inside goods. For instance, on larger shopping trips, products typically associated with small shopping trips (e.g., snacks) are defined as “outside goods”; in such cases, empirical analysis shows that there is a substitution relationship between inside goods and outside goods. Although Allenby et al. (2004b) consider somewhat the multi-category situation, they are limited in their definition of “outside goods,” because researchers arbitrarily define them based on their ability to have substitution/complementary relationships with inside goods.

According to Chintagunta and Nair (2010), research that analyzes the multi-category demand system through the use of a budget allocation model is essential. Outside goods in a single category are defined as alternatives that are not included in analyses and are not choice alternatives within a category; however, outside goods in a multi-category are defined as those from all the remaining categories, except the category that includes inside goods. From a macroscopic perspective, the AIDS model—as suggested by Deaton and Muellbauer (1980) and which considers multi-category situations—analyzes price elasticity and income effect among various consumption categories. In fact, Deaton and Muellbauer (1980) analyze the relationships among consumption categories, but not in single-category situations; they analyze only price elasticity among categories, based on the results of budget allocation at the aggregate

level. Although price is often the most important factor in choosing a product, there exists a number of product attributes that affect consumer decision-making. Therefore, the AIDS model bears a limitation, in that it does not consider attributes other than price and income.

Therefore, the purpose of this dissertation is to suggest a choice model that takes into consideration both multi-stage and multi-category perspectives, in order to reflect the budget allocation stage of the choice model. The multi-stage and multi-category choice model proposed in this dissertation could not only assist in the analysis of MRS among consumption categories; it is also expected that the results of demand analysis are rendered more accurate by including the budget allocation stage in the choice model. Thus, the estimation results derived from the proposed model provide greater managerial and policy implications to both policy-makers and marketers.

Chapter 3. Models

This chapter proposes a choice model that considers the budget allocation stage. The choice models proposed in this dissertation are considered from both multi-stage and multi-category perspectives, and they involve simultaneous problems that occur during the consumer decision-making process. In addition, by considering the budget allocation stage as a preceding stage in the choice model, the proposed models resolve the endogeneity problem that otherwise occurs if a consumer's preceding choice is not considered in the consumer decision-making process. Before introducing the proposed models in this chapter, section 3.1 summarizes the research motivation and purpose, and section 3.2 introduces the model proposed in this dissertation and conducts a simulation study that acts as a validation test. Before developing the choice model, section 3.2.1 reviews simultaneous equation models with both continuous and discrete endogenous variables that are representative models of the multi-stage and single-category model. Section 3.2.2 develops a multi-stage and multi-category discrete choice model that considers the budget allocation and product choice stages, as well as multi-stage and multi-category discrete-continuous choice models that consider the budget allocation, product choice, and product usage stages. Moreover, extended models are also introduced. Section 3.2.3 explains identification issues and the estimation process inherent in the proposed models, and section 3.2.4 conducts validation tests for the proposed models.

Section 3.2.5 explains the implications of the proposed models.

3.1 Research Subjects

An economics-oriented consideration of outside goods within the choice model is important to analyzing accurate product demand and to forecasting the consumer choice problem (Chintagunta and Nair, 2010). Moreover, to suggest a realistic pricing policy, choice models that consider outside goods are used in analyses (Allenby et al., 2004b). The issue of considering outside goods within a choice model converges with the consideration of the budget allocation stage, all the way to the final consumer decision-making process. Actually, by tacitly assuming that previous models include the budget allocation stage, the previous literature can be seen as analyzing the results of consumer choice. In other words, previous studies assume that consumer choice is a result of budget constraints, and so they analyze consumer preference by using a single equation. However, according to Febbri et al. (2004), an empirical model that considers both the choice stage and the stage that precedes it could analyze choice probability more accurately than the empirical model, which considers only a single equation at the choice stage. This means that when the consumer choice problem is handled, it is possible to derive more accurate consumer demand analysis by considering the budget allocation stage as a preceding stage toward an empirical model.

In this dissertation, to improve the accuracy of demand forecasting, a choice

model that considers the budget allocation stage is suggested, from both the multi-stage and multi-category perspectives. The knowledge gleaned from a review of the previous literature discussed in chapter 2 indicates that previous models have been developed from the multi-stage/multi-category or multi-stage and single-category perspectives, but the development of a choice model under the multi-stage and multi-category perspectives has not been sufficient in itself. There is an AIDS model that considers the multi-stage and multi-category situation, but it analyzes only price elasticity and the income effect among an alternative group and from the macroscopic perspective. In other words, a limitation of the AIDS model is that it does not analyze in detail consumer preferences vis-à-vis the attributes of alternatives. In addition, consumer choice is heavily affected by budget, besides the product price, as well as the attributes within the consumer choice problem. Moreover, the consumer budget size and the allocation choice for each category does affect consumer choice. Therefore, it is appropriate to consider the budget allocation stage after breaking it out into the budget size and allocation choice for each category.

From this viewpoint, this dissertation proposes a choice model that consists of two parts: one is a consideration of budget allocation, and the other is of product choice. The budget allocation stage is divided into two parts: the selection of the budget category, and the decision regarding the consumer budget size for each category. Meanwhile, the product choice stage is also divided into two parts: the selection of alternatives, and the decision regarding the usage of alternatives. Of particular note is the fact that the budget allocation stage is defined as a shock effect; it is a change in consumption-expenditure

structure related to the purchase of specific alternatives that is highlighted in this dissertation. In other words, in the selection stage of the budget category, categories are chosen so as to change the expenditure structure related to the purchase of specific alternatives; also, the change in the budget size for each category is decided in order to purchase a specific alternative, and this occurs in the decision stage for the consumer budget size for each category. Thus, to identify the budget allocation stage—which consists of the selection of the budget category and the decision regarding budget size—the change to the consumption-expenditure structure related to the purchase of products is used in the proposed models, and in this dissertation is defined as a shock effect. Figure 3 summarizes the structure of the proposed model, in the form of a schematic diagram of the proposed model used in this dissertation.

This dissertation suggests the use of two kinds of choice models: the multi-stage and multi-category discrete choice model (case 1), which considers the budget allocation and product choice stages, and the multi-stage and multi-category discrete-continuous choice model (case 2), which adds product usage to the case 1 model. Moreover, extended models are proposed to consider multiple choices and multiple uses, from both single-choice and single-usage perspectives.

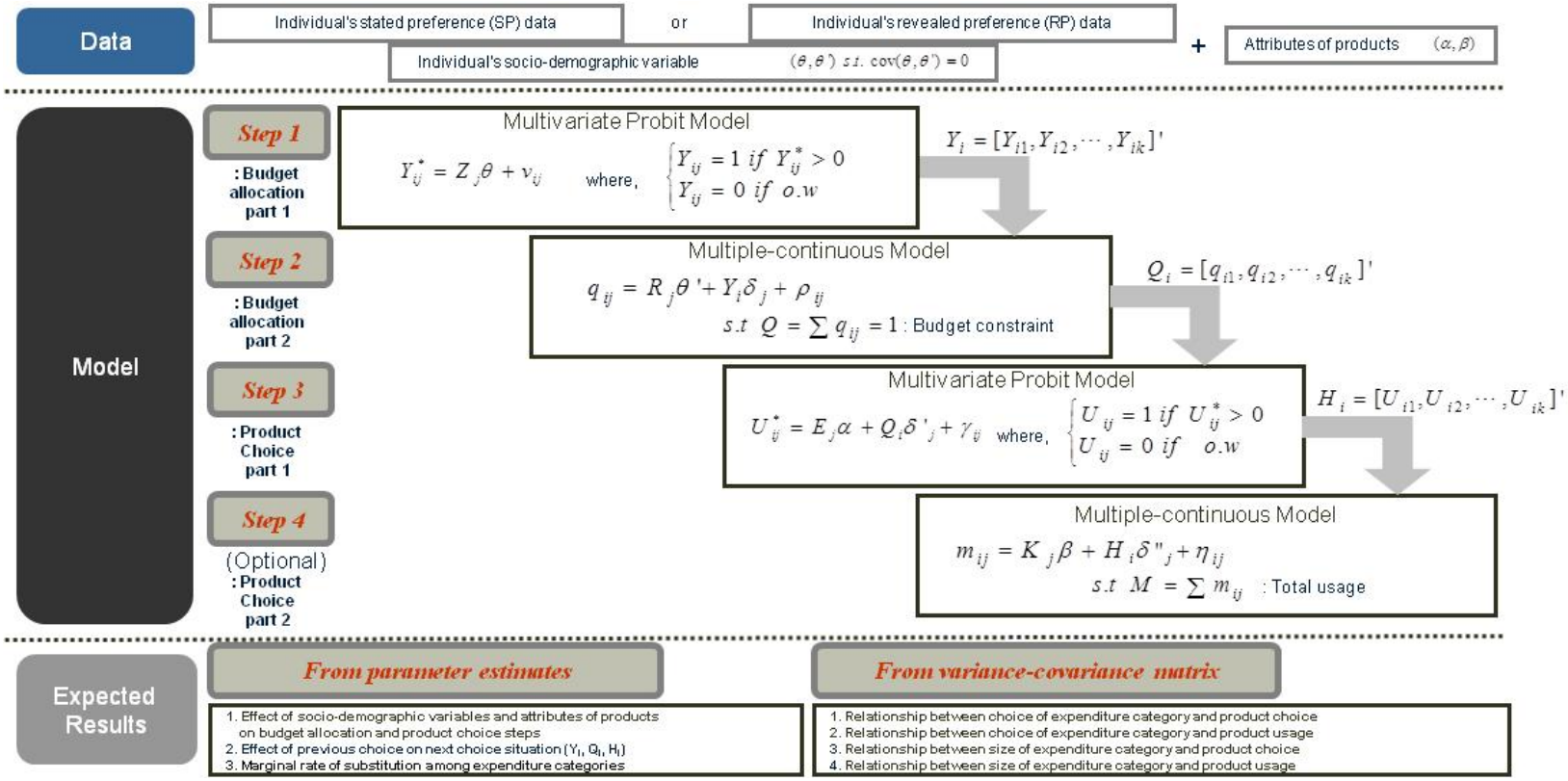

Figure 3. Schematic diagram of the proposed model in this dissertation

In case 1, the MVP–multiple continuous–MNP model considers the situation in which the single product choice is suggested as the multi-stage and multi-category discrete choice model; the MVP–multiple continuous–MVP model—the extended model—is then suggested to consider the situation of a multiple products choice. By using the proposed models of case 1, in chapter 4, the consumer preference of ICT products and the impact of the budget allocation stage on product choice are analyzed. Due to the introduction of smart devices and services, consumer consumption patterns have changed for those products; therefore, consumer preference vis-à-vis smart products and the impact of budget allocation on choice of smart devices should be analyzed to understand the impact of introducing smart devices and services. Thus, the impact of budget allocation on the choice of ICT products is analyzed by using the case 1 model. In addition, the case 1 model is also applied to an analysis of the impact of budget allocation on the choice of products in the household products industry; specifically, among them, eco-friendly products are analyzed, because these products have received considerable attention owing to the current interest in "being green."

In case 2, the MVP–multiple continuous–MNP–single continuous model considers the situation of single product choice and single product usage and is suggested as the multi-stage and multi-category discrete-continuous model; then, the MVP–multiple continuous–MVP–multiple continuous model, which is the extended model, is suggested to consider the situation of multiple-product choice and multiple-product use. The proposed models in case 2 are used to analyze consumer preference regarding the next

generation of automobiles, in terms of choice and usage; they are also used to analyze the impact of budget allocation on the choice and usage of the next generation of vehicles. In particular, the smart car is considered a next-generation automobile, and the impact of budget allocation on the choice and usage of smart cars is analyzed.

In summary, Table 1 outlines the relationship between the proposed models in chapter 3 and the empirical studies in chapter 4.

Table 1. Relationship between Proposed Model and Empirical Study

	Empirical Study 1	Empirical Study 2	Empirical Study 3
Related Industry	ICT Devices	Eco-friendly Products	Smart Cars
Proposed Model	Case 1 Model	Case 1 Model	Case 2 Model
Research Subject	Rigorous Analysis of New Products in the Market by Using Proposed Models that Consider the Budget Allocation Stage		
Estimation Method	Bayesian Estimation Method		

3.2 Multi-stage and Multi-category Discrete-continuous Choice Model with Outside goods

3.2.1 Previous Model: Multi-stage and Single-category Model

Before proposing a choice model that considers multi-stage and multi-category aspects, this section reviews previous models that consider multi-stage and single-category aspects—e.g., the recursive bivariate probit model and the SLDV model. As mentioned in

chapter 2, the recursive bivariate probit model and the SLDV model are included in simultaneous equation models that have both continuous and discrete endogenous variables. These models assume that the results from a preceding choice will affect the second choice, in the form of continuous and discrete endogenous variables (Wilde, 2000). The recursive bivariate probit model and the SLDV model are classified, based on the forms of the first and second equations.

If the structure of both the first and second equations follows the probit model, this type of model is called a bivariate probit model with an endogenous dummy, or a recursive bivariate probit model. With a recursive structure, the choice results from the first equation act as the preceding stage that affects the outcome of the second, subsequent equation, and the form of a potentially endogenous dummy variable is included in the second equation (Febbri et al., 2004). Therefore, the structure of a recursive bivariate probit model is as follows in Eq. (18).

$$(18)\quad \begin{aligned} y_{1i}^{*} &= \beta_1^{'} x_{1i} + u_{1i} \\ y_{2i}^{*} &= \beta_2^{'} x_{2i} + u_{2i} = \delta_1 y_{1i} + \delta_2^{'} z_{2i} + u_{2i} \end{aligned}$$

where y_{1i}^{*} and y_{2i}^{*} represent latent variables, and y_{1t} and y_{2t} describe choice data; these take a value of 0 or 1. x_{1i} and z_{2i} represent exogenous variables. In addition, choice data and the error term bear the following structure:

(19) $\begin{cases} y_{li} = 1, \ if \ \ y_{li}^{*} > 0 \\ y_{li} = 0, \ if \ \ y_{li}^{*} \le 0 \end{cases}, l = 1, 2$

$\begin{pmatrix} u_{1i} \\ u_{2i} \end{pmatrix} \sim IIDN\left(\begin{bmatrix} 0 \\ 0 \end{bmatrix}, \begin{bmatrix} 1 & \rho \\ \rho & 1 \end{bmatrix} \right)$: independently and identically distributed

where the distribution of (u_{1i}, u_{2i}) follows a bivariate normal distribution that has a mean zero and a variance–covariance matrix Ω. The correlation coefficient, ρ, in the variance–covariance matrix, Ω, represents the correlation between choice in the first equation and choice in the second equation. Therefore, if $\rho = 0$, there is no correlated relationship between choice in the first equation and choice in the second equation—that is, consumers make decisions independently. For this case, researchers need only analyze two equations independently. However, if $\rho \neq 0$, there is a correlated relationship between choice in the first equation and choice in the second equation, in which case researchers would need to estimate parameters by using the first and second equations simultaneously.

On the other hand, if the structure of the first equation follows the probit model and the structure of the second equation follows the tobit model, this type of overall model is called the SLDV model. Similar to results from a recursive bivariate probit model, the outcome of the second equation is affected by the results of the first equation, which take the form of endogenous dummy variables (Li, 1998). Therefore, the structure of the SLDV model is similar to that of the latent variable in Eq. (18), and the choice data and the error term have the following structure:

$$(20)\quad \begin{cases} y_{1i} = 1, \; if \;\; y_{1i}^{*} > 0 \\ y_{1i} = 0, \; if \;\; y_{1i}^{*} \le 0 \end{cases}, \begin{cases} y_{2i} = y_{2i}^{*}, \; if \;\; y_{2i}^{*} > 0 \\ y_{2i} = 0, \quad if \;\; y_{2i}^{*} \le 0 \end{cases}$$

$$\begin{pmatrix} u_{1i} \\ u_{2i} \end{pmatrix} \sim BVN\left(0_2, \Sigma\right), \; \Sigma = \begin{pmatrix} 1 & \sigma_{12} \\ \sigma_{21} & \sigma_{22} \end{pmatrix}$$

where y_{1i} and y_{2i} represent observed data. In Eq. (20), the structure of the first equation follows a type of probit model, and the structure of the second equation follows a type of tobit model. Similar to a recursive bivariate probit model, the correlation between the two equations is reflected in σ_{12}. Therefore, whenever $\sigma_{12} \neq 0$, researchers need to analyze the two equations simultaneously, if they wish to consider endogeneity.

Various estimation methods have been developed to estimate simultaneous equation models that contain both continuous and discrete endogenous variables. Blundell and Smith (1989) suggest a two-stage algorithm to estimate an SLDV model that considers the consumer's sequential choice situation. However, Li (1998) compares the FIML estimator to the two-stage algorithm and shows that while the FIML estimation method bears a computational burden, it is a more efficient estimation method than the two-stage algorithm. Additionally, Li (1998) suggests the Bayesian estimation method to overcome the computation difficulty inherent in the FIML estimation method.

Because they offer a variety of merits—such as improvements to estimation methods, the simultaneous consideration of consumer choices, and a consideration of endogeneity—simultaneous equation models with both continuous and discrete

endogenous variables have been applied in a variety of fields; however, among these models, only binary choice models are considered discrete choice models. Because consumers face multinomial or multivariate choices in a considerable proportion of cases, multinomial and multivariate choice types that include more than two alternatives should be considered in simultaneous equation models. Therefore, this dissertation suggests the use of choice models that consider the budget allocation and product choice stages, which are in turn based on the basic model structure of simultaneous equation models that contain both continuous and discrete endogenous variables. Moreover, extended choice models are proposed to consider multiple choice and usage.

3.2.2 Proposed Model

3.2.2.1 Multi-stage and Multi-category Discrete Choice Model and Extended Model

This section proposes choice models that consider multi-stage and multi-category situations. The structure of the proposed models is divided into two parts: budget allocation and product choice. In particular, in the product choice stage, a model that bears a multinomial choice structure that considers more than two alternatives as a choice option is suggested (case 1 base model). When a consumer decides to purchase products, he or she is faced with a multiple-choice situation, wherein he or she must choose more than two alternatives among the various alternatives. Therefore, this section also proposes

an extended model, by which one can consider multiple choices (case 1 extended model).

First, the case 1 base model follows an MVP–multiple continuous–MNP structure. In the case 1 base model, the structure of the budget allocation stage has an MVP–multiple continuous type; the structure of the product choice stage, meanwhile, has an MNP type that considers a single choice situation. In addition, the proposed model assumes that the ith consumer has a sequential decision-making process that consists of a budget allocation and product choice stages, and the choice results of the preceding stage will affect the choice made in each subsequent stage. In the budget allocation stage, consumers choose more than one expenditure category among the consumption-expenditure categories, in order to change the expenditure structure. Because consumers need money to buy new products, consumers gather a sufficient amount of money from changes made to the expenditure structure. The latent utility vector for the choice of expenditure categories is assumed to be $Y_i^* = [Y_{i1}^*, Y_{i2}^*, ..., Y_{ik}^*]'$. Additionally, the multiple continuous vector for changing the budget size for the expenditure categories is assumed to be $q_i = [q_{i1}, q_{i2}, ..., q_{ik}]'$; therefore, the choice of consumption-expenditure categories and decisions regarding changes to the sizes of consumption-expenditure categories in the budget allocation stage are shown as follows:

$$(21)\quad \begin{cases} Y_{i1}^{*} = Z_i\theta_1 + \nu_{i1} \\ Y_{i2}^{*} = Z_i\theta_2 + \nu_{i2} \\ \quad\vdots \\ Y_{ik}^{*} = Z_i\theta_k + \nu_{ik} \end{cases}$$

where, $\begin{cases} Y_{ij} = 1 \; if \;\; Y_{ij}^{*} > 0 \\ Y_{ij} = 0 \; if \;\; o.w \end{cases}$, $Y_i = [Y_{i1}, Y_{i2}, \cdots, Y_{ik}]'$

Z_i = [income, age, house type, number of family members, and other socio-demographic variables].

$$(22)\quad \begin{cases} q_{i1} = R_i\theta_1' + Y_i'\delta_1 + \rho_{i1} \\ q_{i2} = R_i\theta_2' + Y_i'\delta_2 + \rho_{i2} \\ \quad\vdots \\ q_{ik} = R_i\theta_k' + Y_i'\delta_k + \rho_{ik} \end{cases} \quad s.t.\;\; Q = \sum q_{ij} \text{ :budget constraint}$$

where R_i = [income, age, house type, number of family members, and other socio-demographic variables].

Eq. (21) represents the choice of consumption-expenditure categories that have changed and are of an MVP type. Eq. (22) represents the decision to change the budget size for each expenditure category, and it is of a multiple-continuous equation type. Z_i and R_i describe socio-demographic variables that affect the choice of consumption-expenditure categories and the decision to change the budget size for each category. When consumers decide to change the budget size for each expenditure category in Eq. (22), the choice results from Eq. (21)—which describes the choice of consumption-expenditure category—affect the decision vis-à-vis changes to budget size, and they are

included in Eq. (22) as endogenous dummy variables. Here, the endogenous dummy variables are described as $Y_i = [Y_{i1}, Y_{i2}, \cdots, Y_{ik}]'$; therefore, when consumers choose a specific category, $Y_{ij} = 1$, the effect of the preceding choice is considered in the subsequent choice model.

From the budget-allocation results, several things are assumed to describe the product choice stage : the ith consumer chooses the jth alternative among J alternatives, and the latent utility from the jth alternative is assumed to be U_{ij}^*. When the ith consumer does not choose any alternative among J alternatives, the latent utility from the no-choice option is assumed to be U_{i0}^*; this utility is interpreted as the utility from outside goods. Therefore, the consumers' utility function in the product choice stage is shown as follows:

(23) $$\begin{cases} U_{i1}^* = E\alpha_1 + Z_i^{'}\delta_1' + \gamma_{i1} \\ U_{i2}^* = E\alpha_2 + Z_i^{'}\delta_2' + \gamma_{i1} \\ \qquad \vdots \\ U_{iJ}^* = E\alpha_J + Z_i^{'}\delta_2' + \gamma_{i1} \end{cases}$$

where, $\begin{cases} U_{ij} = 1 \; if \; \max(U_i^*) = U_{ij}^* > 0 \\ U_{ij} = 0 \; if \; \max(U_i^*) = U_{ij}^* < 0 \end{cases}$

$Z_i^{'} = [(q_1, q_2, ..., q_K) \otimes I]$, $q_1, ..., q_K$ = budget share for each expenditure category

E = product attributes.

(24) $U_{i0}^* = E\alpha_0 + Z_i^{'}\delta_0' + \gamma_{i0}$

where, $U_{i0}^* > 0 \; if \; U_{ij}^* < 0 \quad for \;\; \forall j$

Eq. (23) represents the product choice stage, and it follows the MNP type. E represents the product's attributes as explanatory variables that affect product choice. As per Eq. (23), when consumers choose a product, the results of their budget allocation in Eq. (22) will affect the product choice and are included in Eq. (23) as endogenous continuous variables. The endogenous continuous variables are assumed to be $Z_i^{'} = [(q_1, q_2, ..., q_K) \otimes I]$; therefore, if the budget share for each expenditure category exceeds zero, $q_{ij} > 0$, the impact of the preceding choice on the subsequent choice will be analyzed. Eq. (24) represents the latent utility for outside goods when the ith consumer does not choose any alternative among J alternatives.

In considering the budget allocation and product choice stages concurrently, the model is as follows in Eq. (25).

(25) $W_i^* = X_i \beta + \varepsilon_i$

$$\begin{bmatrix} Y_i^* \\ q_i \\ U_i^* \end{bmatrix} = \begin{bmatrix} Z_i & & & & \\ & R_i & Y_i & & \\ & & & E_i & Z_i^{'} \end{bmatrix} \begin{bmatrix} \theta \\ \theta' \\ \delta \\ \alpha \\ \delta' \end{bmatrix} + \begin{bmatrix} \nu_i \\ \rho_i \\ \gamma_i \end{bmatrix}$$

where, $\varepsilon_i = \begin{bmatrix} \nu_i \\ \rho_i \\ \gamma_i \end{bmatrix} \sim N(0, \Sigma),\ \Sigma = \begin{bmatrix} \Sigma_{\nu\nu} & \Sigma_{\nu\rho} & \Sigma_{\nu\gamma} \\ & \Sigma_{\rho\rho} & \Sigma_{\rho\gamma} \\ & & \Sigma_{\gamma\gamma} \end{bmatrix}$

In Eq. (25), the endogeneity among choice alternatives could be captured through the variance–covariance matrix between the error terms. In other words, by

simultaneously analyzing the consumer sequential decision-making process, the proposed model considers the impact of the preceding choice on the subsequent choice and the endogeneity that results from the consumer sequential decision-making process. In addition, through the results of the variance–covariance matrix, the proposed model analyzes two kinds of relationships: (1) the relationship between the choice of expenditure categories and product choice, $\Sigma_{v\gamma}$, and (2) the relationship between the decision on budget size for each category and product choice, $\Sigma_{\rho\gamma}$; both are analyzed in the variance–covariance matrix. Moreover, in Eqs. (23) and (24)—which describe the product choice stage—the MRS between inside and outside goods is analyzed. In other words, to compensate for the loss of utility that derives from reducing the budget size for inside goods, the change in budget size for outside goods will be analyzed in MRS. MRS analysis is conducted via Eq. (26):

$$
\begin{aligned}
& U_{ij}^{*}(E, Z_i^{'}) = E\alpha_j + Z_i^{'}\delta'_j + \gamma_{ij}, where\, Z_i^{'} = [q_{i1}, q_{i2}, \cdots, q_{ik}]' \\
& \Rightarrow dU(\cdot) = 0 = \frac{\partial U(\cdot)}{\partial q_{i1}} dq_{i1} + \frac{\partial U(\cdot)}{\partial q_{i2}} dq_{i2} + \cdots \\
& \qquad\qquad = \delta_1 dq_{i1} + \delta_2 dq_{i2} + \cdots \\
& \Rightarrow \frac{dq_{i2}}{dq_{i1}} = -\frac{\delta_1}{\delta_2} (MRS\, between\, inside\, goods \text{ and } outside\, goods\, 2)
\end{aligned}
\tag{26}
$$

Until now, the budget allocation and product choice stages, which examine a single-choice situation, are considered in the case 1 base model. When consumers decide

to purchase products, they frequently face multiple-choice situations in which more than one alternative can be chosen. Therefore, to accommodate a multiple-choice situation, the case 1 extended model is proposed. The structure of the budget allocation stage in the case 1 extended model is similar to that in the case 1 base model; thus, the budget allocation stage in case 1 is the extended model that follows Eqs. (21) and (22). However, the latent utility from product choice is different, because the case 1 extended model considers multiple choices in the product choice stage. The latent utility of product choice in the case 1 extended model is shown as follows in Eq. (27).

(27) $$\begin{cases} U_{i1}^{*} = E\alpha_1 + Z_i^{'}\delta_1 ' + \gamma_{i1} \\ U_{i2}^{*} = E\alpha_2 + Z_i^{'}\delta_2 ' + \gamma_{i2} \\ \qquad \vdots \\ U_{ik}^{*} = E\alpha_k + Z_i^{'}\delta_k ' + \gamma_{ik} \end{cases}$$

where, $\begin{cases} U_{ij} = 1 \; if \;\; U_{ij}^{*} > 0 \\ U_{ij} = 0 \; if \;\; o.w \end{cases}$

$Z_i^{'} = [(q_1, q_2, ..., q_K) \otimes I]$, $q_1, ..., q_K =$ budget share for each expenditure category

$E =$ Product attributes

(28) $U_0^{*} = E\alpha_0 + Z_i^{'}\delta_0 ' + \gamma_0$

Eq. (27) represents the product choice stage and follows the MVP type. Thus, the structure of the case 1 extended model includes Eqs. (21) and (22) as the budget allocation stage, and Eqs. (27) and (28) as the product choice stage; if the two parts are

combined, the final structure of the case 1 extended model resembles Eq. (25). Although the final structure derives from the case 1 base model—which considers the single choice, as the MNP type is similar to the final structure derived from the case 1 extended model (which considers the multiple choices as the MVP type)—the identification for each model distinguishes them. Identification issues will be discussed in section 3.2.3.

3.2.2.2 Multi-stage and Multi-category Discrete-continuous Choice Model and Extended Model

In the previous section, the case 1 base model—which considers the budget allocation stage and the single-product choice—was proposed from multi-stage and multi-category perspectives. Moreover, the case 1 extended model is suggested, to consider a multiple-choice situation in product choice. The current section proposes a choice model that considers not only product choice but also product usage in the product choice stage. Consumers gain utility from product choice, and they gain additional utility from the use of a given product they have chosen (Bhat 2005, 2008). Therefore, the product choice stage is divided into two parts: the product-choice stage and the decision stage of product usage. Additionally, the single-choice or multiple-choice situation is handled in the product-choice stage, and the single-use or multiple-use situation is also addressed in the decision-making stage of the product usage stage.

In this section, the proposed model includes the budget allocation; the product choice, which considers the multinomial choice type; and the product usage stage, which

considers the single continuous outcome (case 2 base model). There are other cases in which consumers choose more than one alternative, and where the chosen alternatives affect total use. For instance, consumers can purchase more than one electric heating appliance, and total electric-power consumption is affected by using those appliances in their homes. Therefore, a multiple-choice situation with respect to product choice and a single continuous outcome with respect to product usage are considered in the proposed model (case 2-1 extended model). Finally, consumers choose more than one alternative, and each product usage is affected by each product. For example, consumers who purchase more than one automobile have a decision-making process whereby they determine the amount of car use for each automobile. Therefore, multiple choices with respect to the product choice stage and multiple continuous outcomes with respect to the product usage stage are considered in the proposed model (case 2-2 extended model)

First, the case 2 base model follows an MVP–multiple continuous–MNP–single continuous structure. The structure of the budget allocation stage in the case 2 base model resembles that of the case 1 base model, which follows an MVP–multiple continuous structure, and that of the product choice stage, which follows an MNP–single continuous structure that considers single choice and single use. In addition, the proposed model assumes that the preceding choice of the ith consumer affects the subsequent choice in each stage. For instance, the choice of expenditure categories in changing the consumption expenditure structure affects the decision to change the budget size for each category; the decision to change the budget size in each category then affects product

choice, and product choice, in turn, affects product use. As mentioned, the budget allocation stage has a structure similar to that of the case 1 base model, which is described by Eqs. (20) and (21).

Following the results of the budget allocation stage, this section focuses on and discusses the product choice stage. For product choice, the ith consumer chooses the jth alternative among J alternatives, and the latent utility from the jth alternative is assumed to be U_{ij}^{*}. If the ith consumer chooses no option—i.e., no alternatives among J alternatives are chosen—the latent utility from this case is assumed to be U_{i0}^{*}, and U_{i0}^{*} represents the utility of the outside goods. The single usage derived from the chosen alternative is assumed to be m_{ij}; therefore, the consumer's utility function at the product choice stage that considers product choice and use is shown as follows:

$$(29)\quad \begin{cases} U_{i1}^{*} = E\alpha_1 + Z_i^{'}\delta_1{}' + \gamma_{i1} \\ U_{i2}^{*} = E\alpha_2 + Z_i^{'}\delta_2{}' + \gamma_{i2} \\ \qquad \vdots \\ U_{iJ}^{*} = E\alpha_J + Z_i^{'}\delta_2{}' + \gamma_{iJ} \end{cases}, \quad U_{i0}^{*} = E\alpha_0 + Z_i^{'}\delta_0{}' + \gamma_{i0}$$

where, $\begin{cases} U_{ij} = 1 \; if \; \max(U_i^{*}) = U_{ij}^{*} > 0 \\ U_{ij} = 0 \; if \; \max(U_i^{*}) = U_{ij}^{*} < 0 \end{cases}$, $U_{i0}^{*} > 0 \; if \; U_{ij}^{*} < 0 \;\; for \;\; \forall j$

$Z_i^{'} = [(q_1, q_2, ..., q_K) \otimes I]$, $q_1, ..., q_K =$ budget share for expenditure categories

$E =$ Product attributes

$$(30)\quad m_{ij} = K_1\beta + H_i\delta_1{}'' + \eta_{i1}$$

where, $H_i = [(U_{i1}, U_{i2}, \cdots, U_{ik}) \otimes I]'$

Eq. (29) represents the product choice stage and follows an MNP type. E represents the product attributes as explanatory variables that affect consumer choice. When consumers decide to purchase the product in Eq. (29), the results of budget allocation in Eq. (22) affect the product choice and are included in Eq. (29) as endogenous continuous variables, which are assumed to be $Z_i^{'} = [(q_1, q_2, ..., q_K) \otimes I]$; therefore, decisions regarding the budget share for each expenditure category affects product choice, and this is a recursive structure. Eq. (30) represents the decision regarding the product usage stage and follows the single continuous structure. K describes product attributes as explanatory variables that affect product use. When consumers decide on product usage (Eq. (30)), the results of product choice (Eq. (29)) affect the product use decision and are included in Eq. (30) as endogenous dummy variables. These endogenous dummy variables are assumed to be $H_i = [(U_{i1}, U_{i2}, \cdots, U_{ik}) \otimes I]'$.

Considering together the budget allocation, product choice, and product usage stages, the model is shown as follows in Eq. (31):

(31) $W_i^* = X_i \beta + \varepsilon_i$

$$\begin{bmatrix} Y_i^* \\ q_i \\ U_i^* \\ m_i \end{bmatrix} = \begin{bmatrix} Z_i & & & & & \\ & R_i & Y_i & & & \\ & & & E_i & Z_i^{'} & \\ & & & & K_i & H_i \end{bmatrix} \begin{bmatrix} \theta \\ \theta' \\ \delta \\ \alpha \\ \delta' \\ \beta \\ \delta'' \end{bmatrix} + \begin{bmatrix} \nu_i \\ \rho_i \\ \gamma_i \\ \eta_i \end{bmatrix}$$

where, $\varepsilon_i = \begin{bmatrix} \nu_i \\ \rho_i \\ \gamma_i \\ \eta_i \end{bmatrix} \sim N(0, \Sigma),\ \Sigma = \begin{bmatrix} \Sigma_{\nu\nu} & \Sigma_{\nu\rho} & \Sigma_{\nu\gamma} & \Sigma_{\nu\eta} \\ & \Sigma_{\rho\rho} & \Sigma_{\rho\gamma} & \Sigma_{\rho\eta} \\ & & \Sigma_{\gamma\gamma} & \Sigma_{\gamma\eta} \\ & & & \Sigma_{\eta\eta} \end{bmatrix}$

In Eq. (31), the endogeneity that results from the impact of the preceding choice stage on subsequent decision stage could be captured through a variance–covariance matrix between error terms. As mentioned in the previous section, by simultaneously analyzing the consumer's sequential decision-making process, the endogeneity that results from the structure of the sequential decision-making process is considered in the proposed model. Moreover, on account of the results of the variance–covariance matrix, the proposed model analyzes four kinds of relationships: (1) the relationship between the choice of the expenditure category and the product choice, $\Sigma_{\nu\gamma}$, (2) the relationship between the decision of the budget size for each category and product choice, $\Sigma_{\rho\gamma}$, (3) the relationship between the choice of expenditure category and product use, $\Sigma_{\nu\eta}$, and (4) the relationship between the decision vis-à-vis budget size for each category and product use, $\Sigma_{\rho\eta}$; these relationships are analyzed within the results of the variance–covariance matrix. Additionally, MRS between consumption-expenditure categories is analyzed by using Eq. (26) in section 2.3.2.1.

Until now, the case 2 base model has considered budget allocation; product choice, which assumes a single-choice situation; and the product usage stage, which assumes a single-use situation. When consumers decide to buy products, they often face a

multiple-choice situation. Therefore, the case 2-1 extended model is proposed, as it considers both multiple choices and single use in the product-choice stage. The structure of the budget allocation stage in the case 2-1 extended model is similar to that of the budget allocation stage of the case 1 base model, which follows Eqs. (21) and (22). Moreover, the utility function of the multiple choices in the product choice stage follows Eq. (27), which is the utility function from the case 1 extended model. Therefore, the case 2-1 extended model includes Eqs. (21) and (22) as the budget allocation stage and Eqs. (27) and (30) as the product choice stage.

Consumers also face multiple-choice situations, as well as decisions regarding multiple usages for products. As mentioned, the consumer choice situation for automobiles is one example, and to consider this choice situation in the choice model, the case 2-2 extended model is suggested. The structure of the budget allocation stage in the case 2-2 extended model also resembles that of the budget allocation stage in the case 1 base model, which follows Eqs. (21) and (22). In addition, the utility function from the multiple choices in the product choice stage follows the utility function of Eq. (27) in the case 1 extended model, and the equations for multiple uses are shown as follows:

$$(32)\quad \begin{cases} m_{i1} = K_1\beta + H_i\delta_1'' + \eta_{i1} \\ m_{i2} = K_2\beta + H_i\delta_2'' + \eta_{i1} \\ \qquad\vdots \\ m_{ik} = K_k\beta + H_i\delta_2'' + \eta_{i1} \end{cases}$$

where, $H_i = [(U_{i1}, U_{i2}, \cdots, U_{ik}) \otimes I]'$

Thus, the case 2-2 extended model includes Eqs. (21) and (22) as the budget allocation stage and Eqs. (27) and (32) as the product choice stage. The final structure of the case 2-2 extended model is similar to that of Eq. (31). As mentioned in section 3.2.2.1, the identification issues relating to the proposed models will be discussed in section 3.2.3.

3.2.3 Identification Issue and Estimation Method

When a dependent variable presents a discrete rather than continuous outcome, an identification problem in estimating parameters occurs (Train, 2003). For discrete outcomes, a consumer's latent utilities are not observed, and only the binary choice outcomes that equal 0 or 1 are observed, from the consumer decision-making process. Therefore, if a consumer's utility function is increasing the scale *c* (see Eq. (34))—or if the level of the consumer's utility function is increased by *m* (see Eq. (35))—the choice outcome does not change. In other words, because the following three equations have the same choice outcomes, the problem is that the identified parameters are not estimated.

(33) $Z_i = X\beta + \varepsilon_i$

(34) $cZ_i = c(X\beta + \varepsilon_i) = X(c\beta) + c\varepsilon_i$

(35) $Z_i + m = X\beta + \varepsilon_i + m$

where Z_i represents the latent utility. The choice outcomes of Eqs. (34) and

(35) are the same, and this means that the increasing scale *c* on latent utility and the added level *m* on latent utility are not different. Therefore, to identify the utility function with the discrete choice outcome, identifications of both level and scale are considered (Train, 2003). According to Koop (2003), if researchers estimate a model that is not applied for the identification process, unidentifiable parameters are estimated and the value of the standard error for estimation results will be higher than the identified results. Thus, resolving the identification problem in the discrete choice model is essential to model validation.

This dissertation proposes models that are types of multiple equation probit models with endogenous dummies or continuous regressors. Previous studies suggest that for the identification of the 2nd equation, the 2nd equation must include at least more than one exogenous variable that is not included in the 1st equation (Maddala, 1983).[6] However, Wilde (2000) criticizes Maddala's (1983) results, and shows that with the identification process, it is not necessary to satisfy the exclusion restriction. Therefore, in the identification process, one need only include at least more than one exogenous variable in each equation—in other words, multiple equation probit models with endogenous dummies or continuous regressors are only needed to consider the identification issue for each equation, respectively. For instance, previous studies that used recursive bivariate probit models resolve the identification problem by normalizing one of the variances to 1 in each binary probit model, as a form of the 1st and 2nd

[6] Maddala (1983) mentions that the exclusion restriction between exogenous variable in the 1st equation and exogenous variable in the 2nd equation is satisfied for identification.

equations (Baslevent and El-hamidi, 2009; Marra and Radice, 2011).

Because the proposed models in this dissertation are of the MNP, MVP, and continuous model types, this section focuses on the MNP and MVP portions that have discrete outcomes and relate to identification issues. Thus, this section reviews the identification methods inherent in MNP and MVP models, and explains the identification process of the model proposed in this dissertation. In addition, based on the identification process, the estimation process for each proposed model is also explained.

First, let us review the identification process for an MNP model. According to Train (2003) and Greene (2008), the scale identification of logit and nested logit models is automatically achieved by the distribution assumption of the error term, and the identification issue from the utility level is resolved by considering the utility difference in the choice model. However, these studies suggest that when researchers use a probit model, they should conduct normalization to resolve both scale and level identification problems; therefore, an MNP model should consider identification issues related to level and scale. Identification for the utility level is resolved by using the utility difference in the choice model, and identification for the utility scale is resolved by normalizing one of the diagonal elements to 1 in the error term matrix (Train, 2003). McCulloch and Rossi (1994) suggest a scale identification method for an MNP model in the Bayesian estimation process; according to their suggestion, the covariance matrix (Σ) is estimated initially in the unidentified model; however, for identification, the estimated covariance matrix (Σ) in the unidentified model is multiplied by the matrix *D*, which is the

normalization matrix, to transpose one of the diagonal elements in the estimated covariance matrix (Σ) into 1. Thus, the identified covariance matrix ($\tilde{\Sigma}$) is analyzed by way of Eq. (36):

$$(36)\quad \tilde{\Sigma} = D^{-1}\Sigma D^{-1},\ where\ D = \begin{bmatrix} 1 & 0 & 0 & 0 \\ & \ddots & \vdots & \vdots \\ & & 1 & 0 \\ & & & \sigma_{nn} \end{bmatrix}$$

To examine the identification issues inherent in the MVP model—a model that considers multiple-choice situations—the MNP model is chosen as the most preferred alternative; the utility structure of the MVP model is assumed to choose more than one alternative if the utility for each alternative is greater than 0. Therefore, for identification, the MVP model is needed to normalize all diagonal elements in the variance–covariance matrix to 1 (Greene, 2008). Chib and Greenberg (1998) and Edwards and Allenby (2003) each suggest an identification method for the MVP model in the Bayesian estimation process, wherein the covariance matrix (Σ) is estimated first in the unidentified model; then, the estimated covariance matrix (Σ) is normalized, to transpose all diagonal elements in the estimated covariance matrix (Σ) to 1 by using matrix *C*, which is a transpose matrix. Thus, the identified covariance matrix ($\tilde{\Sigma}$) is analyzed by using Eq. (37):

$$(37)\quad \tilde{\Sigma} = C^{-1}\Sigma C^{-1},\ where\ C = \begin{bmatrix} \sigma_{11} & 0 & 0 & 0 \\ & \ddots & \vdots & \vdots \\ & & \sigma_{n-1n-1} & 0 \\ & & & \sigma_{nn} \end{bmatrix}$$

In the Bayesian estimation process, McCulloch et al. (2000) propose another identification method for MNP and MVP models. They directly use an identified parameter in the prior of the Bayesian estimation process, and put this prior into the posterior distribution. Finally, the parameters are estimated through the identified posterior distribution.[7] However, according to Nobile (2000), any estimation method that directly uses a prior of the covariance matrix, including identified parameters, will have a slow convergence speed and incur computational difficulties, so this method is not efficient. Therefore, Nobile (2000) suggests an identification process, that when the covariance matrix is drawn by a posterior distribution that is derived from an unidentified prior, the covariance matrix in the posterior distribution is drawn by Wishart distribution or an inverted wishart distribution with $\sigma_{11} = 1$. Webb and Forster (2008), Barnard et al. (2000), and others also propose various identification processes for MNP and MVP models.

The identification process for the model proposed in this dissertation is similar to that of Jeong (2008), which combines the identification methods of McCulloch and Rossi (1994) and Nobile (2000). For the estimation methods, this dissertation uses a

[7] A prior distribution with the identification constraint $\sigma_{11} = 1$ is suggested.

Bayesian estimation method to overcome the computational difficulties incurred by the FIML estimation method used in the SLDV model. The Bayesian estimation method has various advantages over classical estimation methods: it avoids direct computational complexity, initial point problem, and other issues (Edwards and Allenby, 2003; Allenby and Rossi, 1999; Train, 2003). Thus, by using the Bayesian estimation process, the current study examines the estimation process while considering the identification issues inherent in the proposed models.

Except for the continuous stage—which is unnecessary to the identification process—this dissertation initially categorizes the proposed models as MVP–MNP or MVP–MVP types, with respect to identification issues. Therefore, in using this typology, five proposed models in this dissertation are classified as follows in Table 2.

Table 2. Identification Types

Identification Type	Proposed Model
MVP–MNP Type model	- Case 1 Base Model: MVP–Multiple Continuous–MNP - Case 2 Base Model: Case 2 Base Model: MVP–Multiple Continuous–MNP–Single Continuous
MVP–MVP Type Model	- Case 1 Extended Model: MVP–Multiple Continuous–MVP - Case 2-1 Extended Model: MVP–Multiple Continuous–MVP–Single Continuous - Case 2-2 Extended Model: MVP–Multiple Continuous–MVP–Multiple Continuous

Identification and Estimation Process in MVP–MNP type models

In terms of MVP–MNP type models which are proposed in this dissertation, the case 1 base model (MVP–multiple continuous–MNP) and the case 2 base model (MVP–multiple continuous–MNP–single continuous) are included. The Bayesian estimation method is used in the proposed models, as are the identification processes of McCulloch and Rossi (1994) and Nobile (2000). As mentioned, the estimation process of MVP and MNP models have discrete outcomes, as dependent variables are needed to consider identification issues vis-à-vis scale shift and utility level.

To achieve identification for the utility level in the MNP model, one of the alternatives is set as a reference choice; then, the utility function of the remaining alternatives is defined as the difference in utility, relative to the reference choice. In addition, to achieve identification of the scale shift of utility in the MNP model, the identification process of Nobile (2000) is used. In other words, the 3rd stage of the decision-making process in this dissertation—which considers the budget allocation stage—follows the MNP model, and the scale identification of the MNP model in the 3rd stage is solved by following the thinking of Nobile (2000): when an inverted wishart distribution is drawn in the Bayesian estimation process, one of the diagonal elements of the covariance matrix in the MNP model is restricted to 1. Based on the restriction of the inverted wishart distribution, the variance–covariance matrix is drawn.

The identification of the MVP model in the 1st stage remains to be done. To

achieve scale identification in the MVP model, I appeal to the thinking of McCulloch and Rossi (1994): for scale identification, estimated parameters from the unidentified prior in the Bayesian estimation process are divided by each variance of the error terms, σ_{ii}, and the diagonal elements of the estimated variance–covariance matrix from the unidentified prior in the Bayesian estimation process are divided by each variance in the error terms, σ_{ii}, to make all diagonal elements equal to 1. Finally, in this dissertation, the identified parameters in the course of the identification process are retained in each iteration.

Thus, the prior distribution that considers the identification of the MVP–MNP-type model in the Bayesian estimation process is as follows:

$$(38)\quad \begin{cases} \beta \sim Normal\ distribution(\mu_\beta, V_\beta) \\ \Sigma^{-1} \sim Wishart\ distribution(v, \Omega^{-1}) I(\sigma_{k-1}^2 = 1) \end{cases}$$

where k represents the number of alternatives. The indicator function in the Wishart distribution means that variance for the $(k-1)^{th}$ alternative among the diagonal elements of the variance–covariance matrix, Σ, is restricted to 1 for the identification of the MNP model in the 3rd stage (Koop, 2007). Therefore, the joint posterior distribution is derived by using the prior distribution and the likelihood function. To estimate the parameter from the derived joint posterior distribution, Gibbs samplers that consist of a conditional posterior distribution for each parameter are used.

$$(39)\quad \begin{bmatrix} \beta \,|\, \Sigma^{-1}, W^*, y \\ \Sigma^{-1} \,|\, W^*, \beta, y \\ W_i^* \,|\, \beta, \Sigma^{-1}, y \end{bmatrix}$$

where $W*$ is a vector and consists of latent utility (Y_i^*, U_i^*) and continuous outcome (q_i, m_i), and y represents observed discrete choice data (Y_i, U_i). The conditional posterior distribution for each parameter is shown as follows in Eqs. (40)–(42).

$$(40)\quad \begin{aligned} &\beta \,|\, \Sigma^{-1}, W^*, y \sim \textit{Normal distribution}(D_\beta d_\beta, D_\beta) \\ &\textit{where, } D_\beta = \left(\sum_i X_i' \Sigma^{-1} X_i + V_\beta^{-1} \right)^{-1}, \quad d_\beta = \left(\sum_i X_i' \Sigma^{-1} W_i^* + V_\beta^{-1} \mu_\beta \right) \end{aligned}$$

$$(41)\quad \Sigma^{-1} \,|\, W^*, \beta, y \sim \textit{Wishart distribution}\left(N + v, \left(\Omega + \sum_i \varepsilon_i \varepsilon_i' \right)^{-1} \right) I(\sigma_{k-1}^2 = 1)$$

$$(42)\quad \begin{aligned} &Z_i^* \,|\, \beta, \Sigma^{-1}, y, L \overset{ind}{\sim} \textit{conditional Multivariate truncated normal distribution}_{R_i(y_i)} \left(\bar{\mu}, \bar{\Sigma} \right) \\ &\textit{where, } Z_i^* = [Y_i^*, U_i^*]' \textit{ and } L = [q_i, m_i]' \\ &\bar{\mu} = \mu_{Z_i^*} + \Sigma_{Z_i^* L} \Sigma_{LL}^{-1} \left[L - \mu_L \right], \quad \bar{\Sigma} = \Sigma_{Z_i^* Z_i^*} - \Sigma_{Z_i^* L} \Sigma_{LL}^{-1} \Sigma_{L Z_i^*} \end{aligned}$$

In Eq. (41), Σ^{-1} is drawn by using the identification process of Nobile (2000). In Eq. (42), because the L vector—which has a continuous outcome (q_i, m_i)—is not needed to draw latent utility, only the value of latent utility for the discrete outcome

(Y_i , U_i) is needed to be drawn from the multivariate truncated normal distribution of Z_i^* on L. In other words, when the discrete outcome has zero value—i.e., $y_i = 0$—the latent utility, Z_i^*, for this discrete outcome is drawn from the R_i region, which has a negative value. When the discrete outcome has one value—i.e., $y_i = 1$—the latent utility, Z_i^*, for this discrete outcome is drawn from the R_i region, which has a positive value. The process for deriving a conditional multivariate normal distribution from a multivariate normal distribution is discussed in Appendix A.

The estimated parameters in the above process achieve identification, but only for the MNP model in the 3rd stage and through the Gibbs sampler, which is constructed through by way of Nobile's (2000) idea. Thus, McCulloch and Rossi's (1994) idea is used to conduct identification for the MVP model in the 1st stage. In other words, after dividing the extracted β by each variance, σ_{ii}, β / σ_{ii} is retained, and the covariance matrix is identified by using Eq. (37).

Identification and Estimation Process in MVP–MVP type Models

In the MVP–MVP type models proposed in this dissertation, the case 1 extended model (MVP–multiple continuous–MVP), case 2-1 extended model (MVP–multiple continuous–MVP–single continuous), and case 2-2 extended model (MVP–multiple continuous–MVP–multiple continuous) are included. The Bayesian estimation method is used for estimation, while the ideas of McCulloch and Rossi (1994) and Nobile (2000)

are used for identification.

First, Nobile's (2000) idea is applied to achieve identification for the MVP model in the 3rd stage; this means that the 3rd stage of the decision-making process in the MVP–MVP type model—which considers the budget allocation stage—follows the MVP model, and that the identification process for the MVP model is conducted by following Nobile's (2000) idea. As such, when the inverted wishart distribution is drawn in the Bayesian estimation process, one of the diagonal elements of the covariance matrix in the MVP model is restricted to 1. Based on the restriction of the inverted wishart distribution, the variance–covariance matrix is drawn.

Scale identification is still essential to the remaining equation in the 3rd stage and the MVP model in the 1st stage. To achieve the scale identification of these parts, I appeal to McCulloch and Rossi's (1994) work: by dividing the estimated parameters from the unidentified prior in the Bayesian estimation process by each variance in the error terms, σ_{ii}, scale identification for each parameter is accomplished. By dividing the diagonal elements of the estimated variance–covariance matrix from the unidentified prior in the Bayesian estimation process by each variance in the error terms, σ_{ii}—thus making all diagonal elements into 1—the scale identification for the variance–covariance matrix is realized. Finally, the identified parameters derived via the identification process in this dissertation are retained in each iteration.

Thus, to estimate the parameters, the prior and posterior distribution in the Bayesian estimation process—which includes the identification process for the proposed

models of the MVP–MVP type model—are similar to the prior and posterior distributions of the MVP–MNP type model that follow Eq. (38) as a prior and Eqs. (39)–(42) as a posterior. However, the estimated parameters from Eqs. (38)–(42), which are constructed through the use of Nobile's (2000) idea achieves identification for only one equation of the MVP model, in the 3rd stage. Thus, McCulloch and Rossi's (1994) work is used to conduct identification for the remaining equation of the MVP model in the 3rd stage and for the MVP model in the 1st stage.

3.2.4 Validation of Proposed Models

In this section I report on the validation test for the proposed models. To undertake validation for the proposed models, a simulation study is performed by using an arbitrary dataset. The simulation study is divided into two parts: a multi-stage and multi-category discrete choice model (case 1) that includes both the budget allocation and product choice stages, and a multi-stage and multi-category discrete-continuous choice model (case 2) that includes the budget allocation, product choice, and product usage stages. An outline of the simulation study undertaken in this section is provided in Table 3.

Table 3. Outline of the Simulation Study

Category	Proposed Model	Structure of Model	Identification Type
Case 1	Case 1 Base Model	- MVP-Multiple Continuous-MNP	- MVP-MNP Type Model
	Case 1 Extended Model	- MVP-Multiple Continuous-MVP	- MVP-MVP Type Model
Case2	Case 2 Base Model	- MVP-Multiple Continuous-MNP-Single Continuous	- MVP-MNP Type Model
	Case 2-1 Extended Model	- MVP-Multiple Continuous-MVP-Single Continuous	- MVP-MVP Type Model
	Case 2-2 Extended Model	- MVP-Multiple Continuous-MVP-Multiple Continuous	- MVP-MVP Type Model

A simulation study for the case 1 base model among the case 1 proposed models, as well as a simulation study for the case 2-2 extended model among case 2 proposed models, is conducted, and the results thereof are reported in this section. Simulation studies for the remaining proposed models are discussed in Appendix B.

Simulation Study: Case 1 Base Model

For the simulation study of the case 1 base model (MVP–multiple continuous–MNP), the consumer decision-making process is assumed to take place as follows. In the first stage of the decision-making process, consumers choose the expenditure categories that are to be changed, from among two consumption-expenditure categories relating to the product to be purchased. In the second stage of the decision-making process, consumers decide upon the degree of change to the budget size among two consumption-expenditure

categories, given the choice made in the first stage. In the third stage of the decision-making process, consumers decide to buy an alternative among three alternatives, as a result of the budget allocation in the second stage. In the third stage, one of three alternatives is set as the reference alternative; therefore, the structure of the case 1 base model is as shown in Eq. (43):

$$
\begin{aligned}
&1st\ step: \begin{bmatrix} Y_{i1}^* = \theta_{10} + z_{i11}\theta_{11} + z_{i12}\theta_{12} + \nu_{i1} \\ Y_{i2}^* = \theta_{20} + z_{i21}\theta_{21} + z_{i22}\theta_{22} + \nu_{i2} \end{bmatrix} \\
&2nd\ step: \begin{bmatrix} q_{i1} = r_{i11}\alpha_{11} + r_{i12}\alpha_{12} + y_{i1}\delta_{11} + y_{i2}\delta_{12} + \rho_{i1} \\ q_{i2} = r_{i21}\alpha_{21} + r_{i22}\alpha_{22} + y_{i1}\delta_{21} + y_{i2}\delta_{22} + \rho_{i2} \end{bmatrix} \quad s.t.\ Total\ budget(Q) = q_{i1} + q_{i2} \\
&\qquad where, \begin{cases} y_{ij} = 1\ if\ \ Y_{ij}^* > 0 \\ y_{ij} = 0\ if\ \ o.w \end{cases}, j = 1,2 \\
&3rd\ step: \begin{bmatrix} U_{i1}^* = e_{i11}\beta_{11} + e_{i12}\beta_{12} + q_{i1}\delta_{11}^{'} + q_{i2}\delta_{12}^{'} + \gamma_{i1} \\ U_{i2}^* = e_{i21}\beta_{21} + e_{i22}\beta_{22} + q_{i1}\delta_{21}^{'} + q_{i2}\delta_{22}^{'} + \gamma_{i2} \end{bmatrix} \\
&\qquad where, (q_{i1}, q_{i2}) : budget\ share\ \ for\ each\ category
\end{aligned}
\tag{43}
$$

Under the model structure of Eq. (43), the 1,000 arbitrary explanatory variables (z_i, r_i, e_i) are drawn randomly from a normal distribution with mean zero and a variance of 1. In addition, the variance–covariance matrix (Σ) is assumed to have an equi-correlated covariance structure, and it is drawn from this assumption. The true value of the parameters and the variance–covariance matrix, the latter of which has an equi-correlated covariance structure, are defined as follows:

$$[v_{i1}, v_{i2}, \rho_{i1}, \rho_{i2}, \gamma_{i1}, \gamma_{i2}]' \sim N(0, \Sigma)$$

(44) $where, \Sigma = diag(\sigma_{ii})((1-\rho)I_6 + \rho J_6)diag(\sigma_{ii})\ with\, \sigma_{ii} = (1, 1, \sqrt{2}, \sqrt{1.5}, \sqrt{3}, 1),\ \rho = 0.5$

J_6 *is the* 6×6 *matrix of ones*

$$(\theta_{10}, \theta_{11}, \theta_{12}, \theta_{20}, \theta_{21}, \theta_{22}) = (1, -1.5, 1.5, -1, 1, -1)$$

(45) $(\alpha_{11}, \alpha_{12}, \alpha_{21}, \alpha_{22}) = (1, -1, 1.5, -1.5),\ (\delta_{11}, \delta_{12}, \delta_{21}, \delta_{22}) = (1, -1.5, -1.5, 0.5)$

$(\beta_{11}, \beta_{12}, \beta_{21}, \beta_{22}) = (0.5, -1, -0.9, 1.5),\ (\delta'_{11}, \delta'_{12}, \delta'_{21}, \delta'_{22}) = (0.5, -0.6, -0.9, 1)$

The endogenous regressors (i.e., y_{ij}, q_{ij}) are determined by the choice made in the 1st stage and the budget size for the 2nd stage, respectively. Finally, to perform the Bayesian estimation, the prior distribution is assumed to be diffused as follows: $\mu_\beta = 0_{6\times1}, V_\beta = 20^2 \cdot I_{6\times6}, v = 9, \Omega^{-1} = I_{6\times6}$. Therefore, through the Bayesian estimation process, the parameters are estimated by using 1,000 observations that are randomly generated. In other words, 12,000 draws are conducted from each chain by using Gibbs sampling; then, 2,000 draws among those 12,000 draws are discarded, to eliminate the initial point effect. Thus, based on the remaining 10,000 draws, the mean and variance of the parameters are estimated. The estimation results are presented in Tables 4 and 5.

The results show that the value of the estimated parameters converge to the true value of the parameters. Moreover, the root mean square deviation (RMSD)[8] between the estimated parameter from the proposed model and the true parameters is 0.0795, which is quite a low value.

[8] RMSD is also referred to as the root mean square error (RMSE), and this measure shows the difference between the estimator and the true value.

Table 4. Estimation Results of Parameters in the Simulation Study (Case 1 Base Model)

Model	Parameter		True value	Estimated Value	
				beta	sd
MVP (2)	Y1	θ_{10}/σ_{11}	1	1.0019	0.0745
		θ_{11}/σ_{11}	-1.5	-1.5123	0.0962
		θ_{12}/σ_{11}	1.5	1.5247	0.0934
	Y2	θ_{20}/σ_{22}	-1	-1.0263	0.0631
		θ_{21}/σ_{22}	1	1.0210	0.0632
		θ_{22}/σ_{22}	-1	-0.9802	0.0639
Multiple continuous (1)	Q1	α_{11}	1	0.9927	0.0363
		α_{12}	-1	-0.9935	0.0372
		δ_{11}	1	1.0374	0.0648
		δ_{12}	-1.5	-1.6163	0.1075
	Q2	α_{21}	1.5	1.5005	0.0320
		α_{22}	-1.5	-1.5202	0.0319
		δ_{21}	-1.5	-1.4168	0.0578
		δ_{22}	0.5	0.3422	0.0982
MNP (3)	U1	β_{11}	0.5	0.4475	0.0779
		β_{12}	-1	-0.9080	0.1148
		δ'_{11}	0.5	0.5291	0.1646
		δ'_{12}	-0.6	-0.4499	0.1855
	U2	β_{21}	-0.9	-0.8823	0.0810
		β_{22}	1.5	1.5254	0.1137
		δ'_{21}	-0.9	-1.1284	0.1961
		δ'_{22}	1	1.0419	0.1276
RMSD of MVP-Multiple continuous-MNP model = 0.0795					

Table 5. Estimation Results of Variance–Covariance Matrix in the Simulation Study (Case 1 Base Model)

Variance-Covariance matrix	True value	Estimated Value	
		beta	sd
σ_{33}^2	2	1.9301	0.0912
σ_{44}^2	1.5	1.5481	0.0740
σ_{55}^2	3	2.5603	0.5807
ρ	0.5	0.5033	0.0245

Simulation Study: Case 2-2 Extended Model

For the simulation study of the case 2-2 extended model (MVP–multiple continuous–MVP–multiple continuous), the consumer decision-making process is assumed to take place as follows. The decision-making process from the 1st stage to the 3rd stage is assumed to be similar to that in the simulation study of the case 1 base model; however, in the 3rd stage of the decision-making process, the case 2-2 extended model is assumed to be a multiple-choice situation, and the 4th stage in the structure of the case 2-2 extended model is added to the structure of the case 1 base model. Thus, in this 4th stage of the decision-making process, consumers decide the degree to how much they will use alternatives, given the choice made in the 3rd stage. The decision made in the 4th stage is assumed to be a multiple-use situation; therefore, the structure of the case 2-2 extended model as shown in Eq. (46):

$$
\begin{aligned}
&1st\ step: \begin{bmatrix} Y_{i1}^* = \theta_{10} + z_{i11}\theta_{11} + z_{i12}\theta_{12} + v_{i1} \\ Y_{i2}^* = \theta_{20} + z_{i21}\theta_{21} + z_{i22}\theta_{22} + v_{i2} \end{bmatrix} \\
&2nd\ step: \begin{bmatrix} q_{i1} = r_{i11}\alpha_{11} + r_{i12}\alpha_{12} + y_{i1}\delta_{11} + y_{i2}\delta_{12} + \rho_{i1} \\ q_{i2} = r_{i21}\alpha_{21} + r_{i22}\alpha_{22} + y_{i1}\delta_{21} + y_{i2}\delta_{22} + \rho_{i2} \end{bmatrix} \quad s.t.\ Total\ budget(Q) = q_{i1} + q_{i2} \\
&\qquad where, \begin{cases} y_{ij} = 1\ if\ Y_{ij}^* > 0 \\ y_{ij} = 0\ if\ o.w \end{cases}, j = 1,2 \\
&3rd\ step: \begin{bmatrix} U_{i1}^* = e_{i11}\beta_{11} + e_{i12}\beta_{12} + q_{i1}\delta_{11}^{'} + q_{i2}\delta_{12}^{'} + \gamma_{i1} \\ U_{i2}^* = e_{i21}\beta_{21} + e_{i22}\beta_{22} + q_{i1}\delta_{21}^{'} + q_{i2}\delta_{22}^{'} + \gamma_{i2} \end{bmatrix} \\
&\qquad where, (q_{i1}, q_{i2}): budget\ share\ for\ each\ category \\
&4th\ step: \begin{bmatrix} m_{i1} = k_{i11}\lambda_{11} + k_{i12}\lambda_{12} + U_{i1}\delta_{11}^{*} + U_{i2}\delta_{12}^{*} + \eta_{i1} \\ m_{i2} = k_{i21}\lambda_{21} + k_{i22}\lambda_{22} + U_{i1}\delta_{21}^{*} + U_{i2}\delta_{22}^{*} + \eta_{i1} \end{bmatrix} \\
&\qquad where, \begin{cases} U_{ij} = 1\ if\ \max(U_i^*) = U_{ij}^* > 0 \\ U_{ij} = 0\ if\ \max(U_i^*) = U_{ij}^* < 0 \end{cases}, j = 1,2
\end{aligned}
\tag{46}
$$

Under the model structure of Eq. (46), the 1,000 arbitrary explanatory variables (z_i, r_i, e_i, k_i) are drawn randomly from a normal distribution with mean zero and variance 1. In addition, the variance–covariance matrix (Σ) is assumed to have an equi-correlated covariance structure, and it is drawn from this assumption. The true value of the parameters and the variance–covariance matrix, the latter of which has an equi-correlated covariance structure, are defined as follows:

$$
\begin{aligned}
&[v_{i1}, v_{i2}, \rho_{i1}, \rho_{i2}, \gamma_{i1}, \gamma_{i2}, \eta_{i1}, \eta_{i2}]' \sim N(0, \Sigma) \\
&where, \Sigma = diag(\sigma_{ii})((1-\rho)I_8 + \rho J_8)diag(\sigma_{ii}) \\
&\qquad with\ \sigma_{ii} = (1, 1, \sqrt{2}, \sqrt{1.6}, \sqrt{1.3}, 1, \sqrt{1.6}, \sqrt{0.7}),\ \rho = 0.5 \\
&J_8\ is\ the\ 8 \times 8\ matrix\ of\ ones
\end{aligned}
\tag{47}
$$

$$
(48)\quad
\begin{aligned}
&(\theta_{10},\theta_{11},\theta_{12},\theta_{20},\theta_{21},\theta_{22}) = (1,-1.5,1.5,-1,1,-1)\\
&(\alpha_{11},\alpha_{12},\alpha_{21},\alpha_{22}) = (1,-1,1.5,-1.5),\ (\delta_{11},\delta_{12},\delta_{21},\delta_{22}) = (1,-1.5,-1.5,0.5)\\
&(\beta_{11},\beta_{12},\beta_{21},\beta_{22}) = (0.5,-1,-0.9,1.5),\ (\delta_{11}^{'},\delta_{12}^{'},\delta_{21}^{'},\delta_{22}^{'}) = (0.5,-0.9,-0.6,1)\\
&(\lambda_{11},\lambda_{12},\lambda_{21},\lambda_{22}) = (0.3,-0.5,0.7,0.2),\ (\delta_{11}^{''},\delta_{12}^{''},\delta_{21}^{''},\delta_{22}^{''}) = (-0.6,0.4,0.9,-0.7)
\end{aligned}
$$

The endogenous regressors (i.e., y_{ij}, q_{ij}, U_{ij}) are determined by the choice made in the 1st stage, the budget size for the 2nd stage, and the choice made in the 3rd stage, respectively. Finally, to perform the Bayesian estimation, the prior distribution is assumed to be diffused as follows: $\mu_\beta = 0_{8\times1},\ V_\beta = 20^2 \cdot I_{8\times8},\ v = 11,\ \Omega^{-1} = I_{8\times8}$. Therefore, through the Bayesian estimation process, the parameters are estimated by using 1,000 observations that are randomly generated: 12,000 draws are conducted from each chain by using Gibbs sampling and 2,000 draws among those 12,000 draws are discarded, to eliminate the initial point effect. Thus, based on the remaining 10,000 draws, the mean and variance of the parameters are estimated. The estimation results are presented in Tables 6 and 7.

The results of the case 2-2 extended model also show that the values of the estimated parameters converged with the true values of the parameters. Moreover, the RMSD has a value of 0.0619, which is quite low.

Table 6. Estimation Results of Parameters in the Simulation Study

(Case 2-2 Extended Model)

Model	Parameter		True value	Estimated Value	
				beta	sd
MVP (3)	Y1	θ_{10}/σ_{11}	1	0.9505	0.0477
		θ_{11}/σ_{11}	-1.5	-1.4487	0.0622
		θ_{12}/σ_{11}	1.5	1.4155	0.0608
	Y2	θ_{20}/σ_{22}	-1	-0.9508	0.0415
		θ_{21}/σ_{22}	1	0.9572	0.0443
		θ_{22}/σ_{22}	-1	-0.9531	0.0432
Multiple continuous (1)	Q1	α_{11}	1	1.0301	0.0257
		α_{12}	-1	-1.0078	0.0255
		δ_{11}	1	0.9973	0.0440
		δ_{12}	-1.5	-1.6841	0.0720
	Q2	α_{21}	1.5	1.4662	0.0213
		α_{22}	-1.5	-1.5309	0.0222
		δ_{21}	-1.5	-1.5372	0.0369
		δ_{22}	0.5	0.5033	0.0614
MVP (4)	U1	β_{11}/σ_{55}	$0.5/\sqrt{1.3}$ =0.4385	0.3964	0.0296
		β_{12}/σ_{55}	$-1/\sqrt{1.3}$ =-0.8770	-0.8823	0.0362
		δ'_{11}/σ_{55}	$0.5/\sqrt{1.3}$ =0.4385	0.4325	0.0675
		δ'_{12}/σ_{55}	$-0.9/\sqrt{1.3}$ =-0.7893	-0.8348	0.0621
	U2	β_{21}	-0.9	-0.8567	0.0437
		β_{22}	1.5	1.4379	0.0572
		δ'_{21}	-0.6	-0.5714	0.0792
		δ'_{22}	1	0.8565	0.0715
Multiple continuous (2)	M1	λ_{11}	0.3	0.3082	0.0210
		λ_{12}	-0.5	-0.5030	0.0203
		δ''_{11}	-0.6	-0.5051	0.0561
		δ''_{12}	0.4	0.2789	0.0426

M2	λ_{21}	0.7	0.7329	0.0149
	λ_{22}	0.2	0.1871	0.0149
	δ^{*}_{21}	0.9	0.8604	0.0391
	δ^{*}_{22}	-0.7	-0.7201	0.0300
RMSD of MVP-Multiple continuous-MVP-Multiple continuous model = 0.0619				

Table 7. Estimation Results of Variance–Covariance Matrix in the Simulation Study (Case 2-2 Extended Model)

Variance-Covariance matrix	True value	Estimated Value	
		beta	sd
σ^2_{33}	2	2.2368	0.0733
σ^2_{44}	1.5	1.5299	0.0507
σ^2_{77}	1.6	1.6029	0.0539
σ^2_{88}	0.7	0.7265	0.0246
ρ	0.5	0.4986	0.0172

Additionally, the findings derived through the process reported in this section verify that the proposed models are more reflective of consumer purchasing behavior than a single-stage model. To do this, cross-validation is conducted: Cross-validation is conducted based on SP data that had been collected by survey to facilitate the analysis of consumer purchasing behavior vis-à-vis smart pads (chapter 4). In other words, the SP dataset (N = 950) is divided into the calibration sample (N = 800) and the holdout sample (N = 150). Based on the estimation results from the calibration sample, predictions are made with regard to the holdout sample; then, the predicted values are compared to the true values of the holdout sample, in the proposed model and in the single-stage models

that include each decision stage independently.

To conduct the cross-validation, four single-stage models (with outside goods/without outside goods) are estimated independently, and the proposed model (case 1 base model) is estimated by using the calibration sample (N = 800) extracted from SP data concerning consumer purchasing behavior vis-à-vis smart pads. Based on the estimation results, the holdout sample (N = 150) is predicted. Results regarding the predictive performance of the proposed model and the single-stage model (with outside goods/without outside goods) are shown in Table 8.

Table 8. A Comparison of Performance between the Proposed Model and the Single-stage Model in the Holdout Sample

(Consumer Purchasing Behavior of Smart Pad)

Performance	Proposed Model (RMSD)	Single-stage Model with Outside goods (RMSD)	Single-stage Model without Outside goods (RMSD)
MVP (1st Stage)	0.2479	0.2689	–
Multiple Continuous (2nd Stage)	0.0400	0.0543[1]	–
MNP (3rd Stage)	0.2327	0.2562	0.2628
Overall	0.1964	0.2152	–
Holdout sample = 150			

Note: [1]If each single continuous equation among the multiple continuous equations (2nd stage) is estimated independently by using ordinary least squares, the RMSD will be 0.0751, which is larger than the RMSD of the multiple continuous equations (2nd stage).

The RMSD of the single-stage model (with outside goods/without outside

goods) in each stage is larger than the RMSD of the proposed model (case 1 base model). To show the significant difference between the proposed model and the single-stage model (with outside goods/without outside goods), various datasets are needed for a hypothetical test that resembles the process of Hu et al. (1999). However, given the low availability of various datasets, this dissertation cannot demonstrate such a significant difference. According to Dasgupta et al. (1994), however, in the absence of this hypothetical test, the results of cross-validation testing are still valid; therefore, the proposed model performs better in terms of the ability to reflect consumer purchasing behavior than the single-stage model (with outside goods/without outside goods).

3.2.5 Implications of Proposed Models

The proposed models in this dissertation have an advantage in analyzing consumer purchasing behavior comprehensively. To identify the implications of the proposed models, this section shows what could be analyzed through the proposed models and the difference between the results of simple statistics and those of the proposed models.

The estimation results vis-à-vis consumer purchasing behavior via the proposed models are divided into two cases: the results from the estimated parameters, and those from the variance–covariance matrix. Through the estimated parameters, the proposed models could analyze five kinds of relationships: (1) the impact of household (i.e., household income, number of family members, etc.) and agent characteristics (i.e., gender,

age, education level, etc.) on the consumption-expenditure structure, (2) the impact of the choice of consumption-expenditure categories on the decision relating to expenditure size, (3) the identification that changes, in which the consumption-expenditure categories significantly affect the product choice, (4) the consumer preference regarding product attributes, and (5) the MRS among consumption-expenditure categories. In addition, through the variance–covariance matrix, the relationship among consumer decision-making stages is analyzed.

Theoretically, the average consumer's consumption patterns can be analyzed based on simple statistics, but the marginal effects on consumption patterns are not analyzed when socio-demographics change. In addition, regression analysis in which the equation consists of more than one explanatory variable could analyze the marginal effect of each explanatory variable on consumption patterns. According to Anderson et al. (2009), based on the estimation results derived from regression analysis, the change in sales can be predicted as a function of changes to explanatory variables. In other words, compared to simple statistics, regression analysis has an advantage: it can predict the value of a dependent variable, as a function of the future value of explanatory variables.

In addition to the general difference between simple statistics and the results of regression analysis, the estimation results derived from the proposed models show additional information about consumer purchasing behavior: for instance, it is possible to analyze the impact of a structural change in consumption-expenditure on product choice, consumer preference vis-à-vis products, and the expected consumption-expenditure share

as a function of socio-demographic variables. Therefore, the estimation results derived from the proposed models provide more implications in terms of consumer purchasing behavior and consumption patterns in the market, and they can be used to establish efficient marketing strategies and policies.

Chapter 4. Empirical Study

This chapter conducts empirical studies based on the models proposed in chapter 3. Because the impact of budget allocation for each expenditure category on product choice is expected to differ among expenditure categories, this chapter analyzes three products from different categories: ICT, household products, and automobiles. Thus, comparative analysis is conducted about how the effect of budget allocation on product choice differs among products in each industry. Section 4.1 analyzes the effect on choice of smart devices within the ICT industry by taking into account the budget allocation stage of the choice model. Smart phones, smart televisions, and smart pads (or tablet PCs) are all part of the smart device industry; among these products, I focus on the smart pad, in order to analyze the effect of budget allocation on to the fastest-growing segment of the smart pad industry.

Section 4.2 analyzes the effect on choice of green products within the household products industry, by considering the budget allocation stage of the choice model. The eco-friendly household products industry is a hot issue, as it creates and diffuses environmental products used to reduce greenhouse gas emissions and thus help preserve the environment. Therefore, among green products that are made with while considering greenhouse gases and environmental conservation, I focus on eco-friendly laundry detergent to analyze the effect of budget allocation, as eco-friendly laundry detergent is

frequently used and purchased in real life.

Finally, section 4.3 analyzes the effect on the choice of smart car in the automobile industry by considering the budget allocation stage of the choice model. Smart cars are just now being introduced to the automobile market by converging IT technologies with electric-car technologies. Therefore, the effect of budget allocation on the consumer's choice of smart cars—which include electric cars—is analyzed in section 4.3.

4.1 Multi-stage and Multi-category Discrete Choice Model with Outside goods in ICT

4.1.1 Introduction

The term "convergence" is a ubiquitous buzzword in the ICT industry; it refers to the affiliation of disparate fields, such as the internet service sector and the manufactured products sector. According to Eastwood (2006), producers are increasingly releasing and selling new types of convergence devices. From this perspective, a quantitative analysis of consumer preferences for ICT devices and services is essential. Among ICT devices, the recent emergence of smart phones, tablet PCs, and smart televisions has piqued people's interest; especially following the release of Apple's iPad—which has distinctive characteristics such as a user interface (UI) and a substantial amount of content—the tablet PC market has grown explosively.

Prior to the launch of Apple's iPad, many people had raised questions about whether it could survive in the actual market, as it shares so many features with the iPhone, a mobile product bearing the same content as the iPad, and the netbook. However, unlike previous tablet PCs, the iPad has a touch-screen and improves user convenience through the ongoing development of consumer UIs. In addition, unlike the iPhone, the iPad has an expanded screen size and satisfies the consumer need for visual multimedia content. Therefore, games, movies, books, and other media content can be easily utilized on the iPad.

In fact, according to market research companies such as Strategy Analytics (SA),[9] Caris & Company (2010), Gartner, and JP Morgan, the iPad is the leading tablet PC product, and Apple sold 10 million units in the eight months following its April 2010 launch through Apple's new positioning strategy. In addition, the tablet PC market is expected to grow more quickly than the netbook market and is predicted to experience explosive growth. In particular, the Apple iPad has the highest sales among other tablet PCs, and sales of iPads have been in the tens of millions.

In 2011, due to the success of Apple's iPad, various competitors entered the tablet PC market by using the Android operating system; representative products include the Samsung Galaxy Tab series, the LG G-Slate, and the HTC Flyer. Currently, major market research companies (i.e., IDC[10] and Barclays Capital) expect the Apple iPad to continue to dominate the tablet PC market; however, the tablet PC market in South Korea

[9] See http://www.strategyanalytics.com.
[10] See http://www.idc.com.

has a market situation slightly different from those of other countries, because a representative Android-based tablet PC, the Samsung Galaxy Tab, is released 15 days earlier than the iPad; as such, the Apple iPad does not have the first-mover advantage in that country.

Considering this slightly different market situation in South Korea, this section estimates consumer utility for tablet PCs with considering the budget allocation stage; it does so through the use of the proposed model and SP data collected via a conjoint survey. The proposed methodology enables researchers to pinpoint the directions of various firms' strategies vis-à-vis optimal screen size, device performance, and price factor, all of which are hot issues in the tablet PC market. In addition, I analyze the effect of budget allocation on each category, with regard to a consumer's choice of tablet PC. The results of this section can be used by companies to support and establish a business strategy by which they can enter the tablet PC market, and to establish policy related to the tablet PC market.

4.1.2 Data and Empirical Model

This section reports on SP data collected via a conjoint survey. The survey has been conducted from March 2012 to May 2012, and it sampled 1,000 consumers aged 20–59 years from among six metropolitan cities in South Korea (i.e., Seoul, Busan, Daegu, Inchon, Gwangju, and Daejeon). In addition, to improve the reliability of the results, this

survey is conducted through the execution of one-to-one individual interviews. The sample of this survey is extracted on the basis of gender and age through a purposive quota sampling method, to reflect with reasonable accuracy the characteristics of the actual population (i.e., population proportions in terms of gender and age). Thus, the empirical study discussed in this section make use of a sample comprising 950 survey respondents, except those for which there are missing or unreliable data.[11] The demographic properties of the sample extracted from the survey results are shown in Table 9.

[11] To improve the reliability of the responses to the questions, I remove from the dataset records that featured inconsistent answers. For instance, when some respondents provide answers regarding changes to budget allocation based on the original structure of their budget allocation, some of them change their budget size for particular categories so that they exceeded the original budget size; I extract these records from the dataset, to facilitate more consistent analysis. In other cases, I remove from the dataset used in the analysis extreme outliers.

Table 9. Demographic Properties of Respondents in Empirical Study 1

Category		Respondents	Percentage (%)	Average	Standard Deviation
Total		950	100	-	-
Gender	Male(1)	477	50.2	0.50	0.50
	Female(0)	473	49.8		
Age	20s	248	26.1	38 year	10.86
	30s	248	26.1		
	40s	276	29.0		
	50s	178	18.7		
Education Level	Under high school graduate	368	38.7	-	-
	In college or college graduate	551	58.0		
	Over university graduate	31	3.3		
Average monthly income per household (10,000 won)	Under 199	23	2.4	4.19 million won	162.4
	200 ~ 299	105	11.1		
	300 ~ 399	305	32.1		
	400 ~ 499	234	24.6		
	500 ~ 599	175	18.4		
	Over 600	108	11.4		

To describe the various tablet PC devices available on the market, several attributes are selected from a pilot test conducted by the author's group members; operating system (OS), screen size, weight, delay time, and price are chosen as the core attributes, to analyze consumer preferences vis-à-vis tablet PCs. The attribute details and their relative levels, as addressed in this study, are outlined in Table 10.

Table 10. Attributes and Attribute Levels of Smart Pad Devices

Attribute	Level	Explanation
OS	iOS/ Android/ Windows Mobile	iOS: OS by Apple Android: OS by Google Windows Mobile: OS by Microsoft
Screen Size (inch)	7/9/11	7 inch: 3 times larger than usual smart phone (4 inch) 9 inch: 5 times larger than usual smart phone (4 inch) 11 inch: 7.5 times larger than usual smart phone (4 inch)
Weight (gram)	400/700/ 1,000	The weights of representative smart pads is shown below: iPad 2 (610 g), Galaxy Tab 7.7 (370 g), Galaxy Tab 10.1 (575 g)
Delay Time (second)	Fast(1)/ Normal(5)/ Slow(10)	The amount of time it will take to open the web page such as a navigation homepage
Player Price (10,000 KRW)	50/75/100	Price of tablet PCs

Note: OS, operating system; KRW, South Korean won

Based on the five aforementioned attributes and the relative attribute levels, the total number of possible alternative combinations is 3 × 3 × 3 × 3 × 3 = 243. However, this total number of possible alternatives is too large for respondents to choose from, for the purposes of this survey; therefore, this empirical study features a fractional factorial design and extracts optimal alternatives through the use of SPSS. Through the use of a fractional factorial design, 18 alternative choice-cards are extracted, and these alternative choice-cards are divided into six choice sets that consist of four alternative cards that include a no-purchase option. Additionally, I divide the respondents into two groups and

show three different choice sets; thus, the respondents are to choose the one alternative that represents the highest utility in each choice set. Each respondent answers three times.

Before choosing the conjoint cards, the respondents answer questions regarding the structure of their household consumption expenditures, which equal their household monthly income. This dissertation makes use of 10 consumption-expenditure categories, based on the classification of individual consumption by purpose (COICOP)[12]: food and nonalcoholic beverages, alcoholic beverages and tobacco, clothing and fashion accessories, household commodities, transport, communication, recreation and culture, education and health, saving and insurance, and housing/electricity/other. After answering these questions about their current structure of household consumption expenditures, they answer by how much they would change consumption expenditure in each category in order to facilitate the purchase of a smart pad; for this purpose, the price of a smart pad is assumed to be 0.7 million South Korean won (KRW).[13]

Based on the survey data, this section analyzes consumer purchasing behavior with regard to smart pads, through the use of the proposed model (case 1 base model), which considers budget allocation stage. The empirical model is shown in Eq. (49):

[12] The COICOP is first suggested by the United Nations (UN) statistics division in 1999. Statistics Korea (wwww.kostat.go.kr) uses COICOP from 2009 to survey the structure of household consumption expenditure. Consumption expenditure categories in COICOP include 12 categories: food and nonalcoholic beverages, alcoholic beverages/tobacco/narcotics, clothing and footwear, housing/water/electricity/gas/other fuels, furnishings/household equipment/routine household maintenance, health, transport, communication, recreation and culture, education, restaurants and hotels, and miscellaneous goods and services.

[13] As of January 2012, the average exchange rate per U.S. dollar is 1,159.60 Korean won.

$$1st\ step:\ Y_{ij}^{*} = [(One, Income, Sex, Age, Edu) \otimes I]\theta + v_{ij},\quad j = 1,...,10$$

$$2nd\ step: q_{ij} = [(Income, Sex, Age, Edu) \otimes I]\alpha + \sum_{k=1}^{10} y_{ik}\delta_k + \rho_{ij}\ s.t.$$

$$s.t.\ \sum_j q_{ij} = \sum the\ change\ of\ each\ expenditure\ category$$

(49)

$$where, \begin{cases} y_{ij} = 1\ if\ Y_{ij}^{*} > 0 \\ y_{ij} = 0\ if\ o.w \end{cases},\ j = 1,...,10$$

$$3rd\ step:\ U_{il}^{*} = [Delay, (Screen/Weight, \mathrm{Price}) \otimes I]\beta + \sum_{k=1}^{10} z_{ik}\delta_k^{'} + \gamma_{ij}, l = 1,2,3$$

$$where, z_{ik}: the\ share\ of\ changed\ expenditure\ for\ k^{th}\ category$$

where, *Income* describes the household monthly income, which is one of the variables for household characteristics. *Sex*, *Age*, and *Edu* represent agent characteristics: *Sex* takes a value of 1 if a respondent is male and 0 if female, *Age* describes an agent's age level, and *Edu* describes an agent's education level. *Delay*, *Screen/Weight*, and *Price* are attributes of smart pad devices.

In the product-choice segment, four alternative choices—i.e., smart pads with iOS, Android OS, Windows mobile OS, and a no-purchase option—are offered to respondents. For identification, the no-purchase option is used as the base alternative.

4.1.3 Results and Discussion

Based on the case 1 base model and the identification process used in this dissertation, I analyze consumer purchasing behavior with consumer budget allocations, for smart pad

devices. To perform the Bayesian estimation process, the prior distribution is assumed to be diffused; therefore, parameters are estimated through 10,000 draws, which are drawn from each Markov chain. To exclude the initial point effect, the first 1,000 draws among the 10,000 draws generated from the Bayesian estimation process are discarded; in other words, these records are considered part of a burn-in period. Based on the remaining 9,000 draws, the mean and variance of parameters are estimated, and the estimation results are provided below in Tables 11 and 12.

Table 11. Estimation Results of Empirical Study 1

1st step	Variables	Food (Y1)		Alcoholic beverage/ Tobacco (Y2)		Clothing/Fashion Accessories (Y3)		Household commodities (Y4)		Transport (Y5)		Communication (Y6)		Recreation/ Culture (Y7)		Education/Health (Y8)		Saving/Insurance (Y9)		Housing/electricity/ Others (Y10)	
		beta	s.d.	beta	s.d.	beta	s.d.	beta	s.d.	beta	s.d.	beta	s.d.	beta	s.d.	beta	s.d.	beta	s.d.	beta	s.d.
MVP	Constant	**1.8154***	0.2223	**0.7047***	0.2200	**0.5524***	0.2161	**-0.5002***	0.2309	**1.1897***	0.2151	**0.6139***	0.2242	0.0181	0.2455	-0.1150	0.2683	**0.7941***	0.2692	0.2198	0.2175
	Income	**-0.8044***	0.1601	-0.1456	0.1582	**-0.3794***	0.1567	**0.3330***	0.1691	**-0.3326***	0.1576	**-0.3057***	0.1695	**-0.4066***	0.1836	**-0.4175***	0.2018	**-0.3310***	0.1733	**-0.5324***	0.1585
	Sex	**-0.1348***	0.0502	**0.2297***	0.0483	**-0.1067***	0.0474	0.0011	0.0527	0.0334	0.0492	0.0302	0.0514	-0.0744	0.0533	**-0.1120***	0.0576	-0.0591	0.0569	**0.0908***	0.0478
	Age	**-0.6737***	0.2531	**-0.7948***	0.2520	**-0.5823***	0.2457	**0.5808***	0.2738	**-1.0483***	0.2456	**-1.6107***	0.2577	**-0.7707***	0.2703	**-1.1120***	0.2935	-0.0687	0.3011	**0.5317***	0.2476
	Edu	**-1.6099***	0.4762	**-1.8201***	0.4647	-0.1846	0.4587	**-1.6892***	0.4867	**-2.4471***	0.4543	**-1.4522***	0.4774	-0.5991	0.5146	-0.5137	0.5634	**1.0468***	0.5591	-0.7419	0.4648

2nd step	Variables	Size (Q1)		Size (Q2)		Size (Q3)		Size (Q4)		Size (Q5)		Size (Q6)		Size (Q7)		Size (Q8)		Size (Q9)		Size (Q10)	
		beta	s.d.	beta	s.d.	beta	s.d.	beta	s.d.	beta	s.d.	beta	s.d.	beta	s.d.	beta	s.d.	beta	s.d.	beta	s.d.
Multiple continuous	Income	-0.0092	0.0138	**0.0202***	0.0121	0.0176	0.0122	0.0064	0.0120	0.0037	0.0120	0.0049	0.0119	-0.0020	0.0122	0.0004	0.0121	0.0237	0.0205	0.0013	0.0129
	Sex	0.0015	0.0045	0.0004	0.0039	-0.0039	0.0039	-0.0003	0.0039	0.0015	0.0039	0.0018	0.0038	-0.0051	0.0040	-0.0038	0.0040	**0.0112***	0.0067	-0.0044	0.0042
	Age	**0.0353***	0.0181	0.0127	0.0155	0.0230	0.0156	0.0077	0.0156	0.0221	0.0154	0.0066	0.0148	0.0190	0.0158	0.0052	0.0159	**0.1999***	0.0278	**0.0609***	0.0171
	Edu	**0.1499***	0.0292	**0.0629***	0.0249	**0.0422***	0.0253	**0.0425***	0.0245	**0.0469***	0.0247	0.0274	0.0243	0.0357	0.0246	0.0346	0.0254	**0.3764***	0.0432	**0.0689***	0.0270
	Y1	**0.1552***	0.0075	-0.0029	0.0063	-0.0034	0.0062	-0.0011	0.0061	0.0002	0.0060	0.0006	0.0060	0.0077	0.0061	**0.0105***	0.0062	**-0.0828***	0.0112	-0.0034	0.0068
	Y2	-0.0053	0.0076	**0.1014***	0.0062	-0.0070	0.0064	-0.0022	0.0063	-0.0011	0.0062	-0.0047	0.0059	-0.0026	0.0064	-0.0028	0.0064	**-0.0445***	0.0119	-0.0011	0.0070
	Y3	-0.0117	0.0078	-0.0027	0.0065	**0.1051***	0.0064	-0.0016	0.0062	-0.0004	0.0062	-0.0010	0.0060	-0.0054	0.0063	-0.0052	0.0063	**-0.0557***	0.0116	-0.0068	0.0069
	Y4	**-0.0171***	0.0082	-0.0101	0.0066	-0.0073	0.0067	**0.0886***	0.0064	**-0.0108***	0.0065	-0.0062	0.0064	-0.0092	0.0065	-0.0075	0.0067	-0.0132	0.0120	**-0.0125***	0.0073
	Y5	-0.0054	0.0075	-0.0021	0.0063	-0.0059	0.0064	-0.0057	0.0062	**0.0983***	0.0060	-0.0075	0.0061	-0.0074	0.0062	-0.0067	0.0061	-0.0173	0.0112	-0.0024	0.0070
	Y6	-0.0067	0.0079	-0.0090	0.0065	-0.0066	0.0065	-0.0082	0.0065	**-0.0108***	0.0062	**0.0896***	0.0060	-0.0065	0.0063	-0.0086	0.0064	-0.0009	0.0117	-0.0018	0.0071
	Y7	-0.0025	0.0107	**-0.0158***	0.0095	-0.0097	0.0094	**-0.0169***	0.0092	0.0009	0.0090	0.0031	0.0089	**0.0813***	0.0091	0.0036	0.0092	**-0.0274***	0.0150	0.0005	0.0101
	Y8	-0.0157	0.0111	0.0045	0.0100	0.0039	0.0099	0.0090	0.0097	-0.0057	0.0094	-0.0037	0.0093	**0.0609***	0.0096	**0.1282***	0.0098	-0.0113	0.0151	-0.0096	0.0108
	Y9	**-0.0586***	0.0081	**-0.0278***	0.0068	**-0.0193***	0.0069	**-0.0139***	0.0066	**-0.0252***	0.0066	**-0.0122***	0.0064	**-0.0134***	0.0067	**-0.0123***	0.0067	**0.2762***	0.0131	**-0.0419***	0.0074
	Y10	**-0.0183***	0.0077	-0.0053	0.0065	-0.0090	0.0064	-0.0016	0.0062	-0.0035	0.0062	0.0009	0.0062	-0.0027	0.0063	-0.0008	0.0065	**-0.0469***	0.0113	**0.1222***	0.0072

3rd step	Variables		Smart pad with iOS (U1)		Smart pad with Android OS (U2)		Smart pad with window mobile OS (U3)	
			beta	s.d.	beta	s.d.	beta	s.d.
MNP	Screen size/weight		**-0.5489***	0.1109	**0.2303***	0.0564	**-0.1762***	0.0473
	Price		**-1.0569***	0.2124	**-1.3039***	0.2248	**-0.4051***	0.1545
	Q1		**1.4993***	0.4953	0.4467	0.4555	-0.3774	0.3589
	Q2		1.2672	0.8406	1.1419	0.7533	0.2079	0.5610
	Q3		**4.9713***	0.9191	**3.1732***	0.8491	**2.3260***	0.6400
	Q4		-0.5384	1.1031	-0.1050	0.9143	0.1384	0.7052
	Q5		1.5373	1.0498	0.9990	0.9389	-0.0778	0.7281
	Q6		**5.7379***	1.3379	**3.9217***	1.1988	**3.1413***	0.8632
	Q7		**2.8585***	1.5547	**2.6688***	1.3583	**1.9475***	1.0629
	Q8		0.9589	1.6124	-1.3855	1.4104	**-2.6966***	1.0893
	Q9		**1.3391***	0.2881	**0.7527***	0.2665	**0.6834***	0.1890
	Q10		-0.5764	0.6193	0.1336	0.5200	0.3066	0.3732
	Delay	beta	**-0.2883***					
		s.d.	0.0562					

Note: * significant at 10% level

Table 12. Variance–Covariance Matrix of Empirical Study 1

	Y1	Y2	Y3	Y4	Y5	Y6	Y7	Y8	Y9	Y10	Q1	Q2	Q3	Q4	Q5	Q6	Q7	Q8	Q9	Q10	U1	U2	U3
Y1	1.0000	0.3194*	0.4815*	0.3323*	0.4721*	0.3333*	0.4043*	0.4112*	-0.2271*	0.4798*	-0.0006	-0.0015	0.0003	-0.0012	-0.0013	-0.0017	-0.0028	-0.0037	-0.0483*	-0.0009	-0.0477	-0.0043	0.1816*
Y2		1.0000	0.4651*	0.4494*	0.2502*	0.1295*	0.1459*	0.0838*	-0.3233*	0.2622*	0.0016	0.0005	0.0014	0.0002	0.0002	-0.0001	-0.0006	-0.0013	-0.0451*	0.0008	-0.0284	0.0558	0.0857
Y3			1.0000	0.6027*	0.3974*	0.3407*	0.2086*	0.1581*	-0.2638*	0.3608*	0.0010	0.0001	0.0013	-0.0001	-0.0001	-0.0005	-0.0010	-0.0018	-0.0541*	0.0002	-0.1722	-0.1062	-0.0252
Y4				1.0000	0.3234*	0.4345*	0.2274*	0.1718*	-0.2324*	0.4249*	0.0028	0.0016	0.0023	0.0014	0.0009	0.0004	0.0003	-0.0005	-0.0592*	0.0014	0.0214	0.0187	-0.0207
Y5					1.0000	0.4892*	0.3818*	0.4233*	-0.3719*	0.4932*	0.0020	0.0008	0.0021	0.0003	0.0001	-0.0005	-0.0006	-0.0018	-0.0609*	0.0007	-0.0724	-0.0502	0.1189
Y6						1.0000	0.3663*	0.3916*	-0.2882*	0.4407*	0.0033	0.0017	0.0024	0.0014	0.0009	0.0002	0.0008	0.0000	-0.0593*	0.0011	-0.3614*	-0.2916*	-0.1484*
Y7							1.0000	0.9116*	-0.3003*	0.3966*	0.0029	0.0010	0.0022	0.0009	0.0004	-0.0001	0.0001	-0.0008	-0.0492*	0.0008	-0.1170	-0.0544	0.1316
Y8								1.0000	-0.2981*	0.4034*	0.0028	0.0009	0.0022	0.0008	0.0003	-0.0002	0.0001	-0.0009	-0.0478*	0.0007	-0.1021	-0.0467	0.1508*
Y9									1.0000	-0.3821*	-0.0045	-0.0015	-0.0026	-0.0013	-0.0007	-0.0003	-0.0008	-0.0001	0.0362*	-0.0013	0.0499	0.0281	-0.0903
Y10										1.0000	0.0018	0.0005	0.0019	0.0002	-0.0001	-0.0006	-0.0010	-0.0022	-0.0543*	0.0004	0.2512*	0.1355	0.1649*
Q1											0.0134*	0.0001	0.0002	0.0000	0.0000	0.0000	0.0000	0.0000	-0.0036*	-0.0005*	-0.0019	-0.0009	-0.0006
Q2												0.0104*	0.0001	0.0001	0.0001	0.0000	0.0000	0.0000	-0.0013*	-0.0002	-0.0012	-0.0010	-0.0007
Q3													0.0105*	0.0001	0.0000	0.0000	0.0000	0.0000	-0.0016*	-0.0001	-0.0004	-0.0007	-0.0001
Q4														0.0100*	0.0001	0.0000	0.0001	0.0001	-0.0008*	-0.0001	-0.0005	-0.0007	-0.0007
Q5															0.0100*	0.0001	0.0001	0.0001	-0.0009*	-0.0001	-0.0004	-0.0007	-0.0008
Q6																0.0097*	0.0000	0.0000	-0.0006	0.0000	-0.0007	-0.0007	-0.0007
Q7																	0.0102*	0.0009*	-0.0008*	0.0000	-0.0009	-0.0014	-0.0015
Q8																		0.0103*	-0.0005	-0.0001	-0.0016	-0.0018	-0.0017
Q9																			0.0290*	-0.0015*	0.0151	0.0107	-0.0002
Q10																				0.0120*	0.0007	0.0004	-0.0005
U1																					2.9540*	1.5365*	0.6643*
U2																						2.2916*	0.7983*
U3																							1.0000

Note: * significant at 10% level

Several factors can be analyzed from the estimation results. First, differences in purchasing behavior, as a function of household and agent characteristics, can be analyzed. In other words, product choice and the budget sizes of the consumption-expenditure categories for purchasing a smart pad are identified, as a function of household and agent characteristics. For instance, if a consumer has a higher income level, he or she is more likely to choose a household commodity category for changing the budget size. With respect to the choice of consumption-expenditure categories, consumers with higher income levels have a greater tendency to reduce their budget size in the alcoholic beverage/tobacco category to facilitate the purchase of a smart pad; this means that because the alcoholic beverage/tobacco category is considered to comprise nonessential products that satisfy only personal desires, people with higher incomes are more likely to reduce their demand for alcoholic beverages/tobacco and transfer that demand to the purchase of a smart pad device.

Second, the relationships among consumption-expenditure categories are identified from the estimation results in the 2nd stage and via the variance–covariance matrix. To purchase a smart pad, the choice in the saving/insurance category has a negative effect on the budget size for the other categories: if the budget size for the saving/insurance category increases, the budget size for each of the other categories is reduced. Therefore, there is a substitution relationship between the choice in the saving/insurance category and the budget size for the other categories. However, the choice regarding the education/health category has a positive effect on the budget size for

the recreation/culture category, and the budget sizes for these categories are positively correlated. Thus, there is complementary relationship between the education/health and recreation/culture categories.

Third, the estimation results identify the consumption-expenditure categories wherein changes will significantly affect consumer purchasing behavior vis-à-vis smart pads. In other words, the estimation results show that a change in the ratio of consumption-expenditure categories will affect the choice probability of smart pads. For example, if consumers reduce a higher ratio of food category expenditure to purchase a smart pad, they are more likely to purchase a smart pad that features iOS; in other words, consumers with a higher tendency to buy smart pads that feature iOS are more willing to reduce their expenditure for an essential category such as food. Similarly, if consumers are willing to reduce their education/health category expenditure ratio to purchase a smart pad, they are more likely to purchase smart pads that feature Windows mobile OS.

Finally, the 3rd stage estimation results also show the preference of attributes vis-à-vis smart pads. For the screen size, if the screen size is bigger, the choice probability of smart pads with Android OS increases, but the choice probability of smart pads with iOS and Windows mobile OS decreases. Thus, based on the preference of screen size, smart pads with Android OS should be diffused by providing products with diversified screen size, but the proliferation of smart pads with iOS or Windows mobile OS should be facilitated through the provision of products with a compact screen size. According to Park et al. (2011), the smart device market should be examined synthetically to analyze

consumer purchasing behavior, while considering the overall value chain, which consists of the related content, platform, network, and device. In particular, a device's OS platform has the greatest effect on consumer purchasing behavior among smart devices and content in the smart device market. Therefore, when the consumer purchasing behavior with regard to smart pads is examined in future research, the effect of the penetration rates of smart phones and smart televisions based on an OS platform should be considered, in order to analyze the effect of the OS platform. Indeed, the interrelationships of smart devices should be considered in future research.

In the following section, several scenario analyses are conducted to analyze trends in expenditure categories and how they affect the purchase of smart pad devices.

Scenario Analysis

To examine expenditure category trends with regard to the purchase of smart pads, the expected expenditure share of the representative consumer is analyzed, based on the estimation results. According to socio-demographic information on the representative consumer and estimation results, changes to the consumption-expenditure structure in each category, and how it is brought to bear on the purchase of smart pads, are shown in Figure 4.

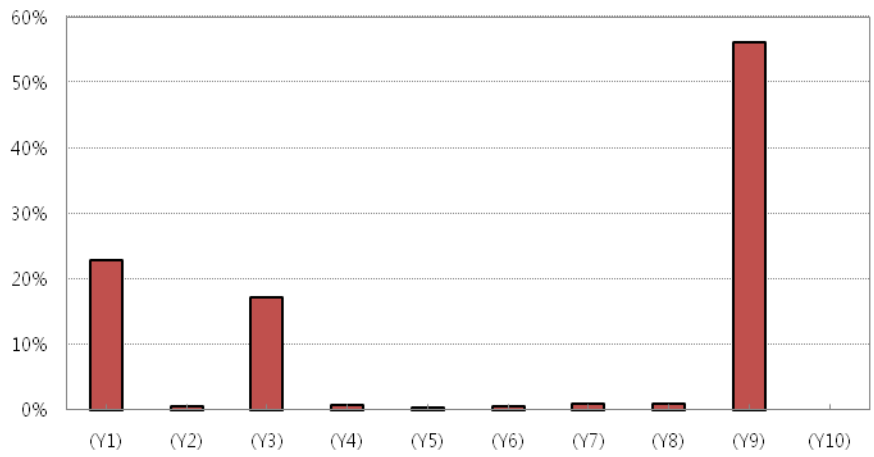

Figure 4. Changes to the Consumption-Expenditure Structure for the Representative Consumer, in Empirical Study 1

From Figure 4, the representative consumer mainly reduces expenditures in his or her food, clothing/fashion accessories, and saving/insurance categories to facilitate the purchase of a smart pad; reductions in the budget sizes for these categories accounts for about 97% of the total overall expenditure change. To compare the impact of socio-demographic level on changes to the consumption-expenditure structure, four scenario analyses are conducted by using four socio-demographic variables: education level, age level, income level, and gender. The expected expenditure ranking and expected expenditure share of the representative consumer, as a function of education level, are analyzed in scenario 1, the results of which are shown in Figure 5.

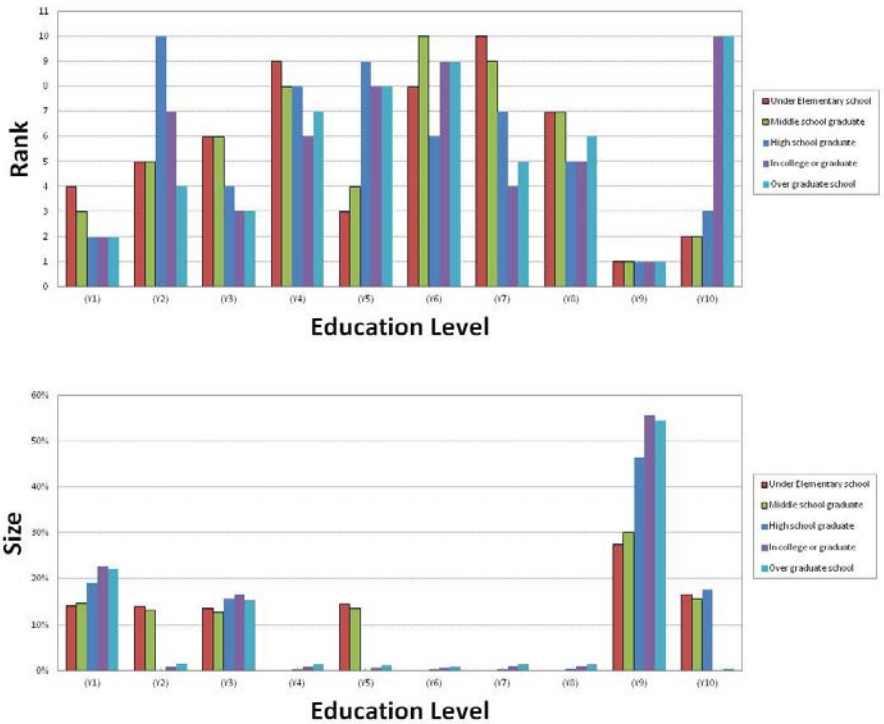

Figure 5. Changes to the Consumption-Expenditure Structure for the Representative Consumer, as a Function of Education Level, in Empirical Study 1

Figure 5 shows that the higher the consumer education level is, the more likely the ratios of the food, clothing/fashion accessories, recreation/culture, and saving/insurance expenditure categories will increase in order to facilitate the purchase of a smart pad. Based on the 3rd stage results in Table 10, one can see that among consumers with higher education levels, there is a greater willingness to buy smart pads that feature iOS than others.

The expected expenditure ranking and expected expenditure share of the representative consumer, as a function of age, are analyzed in scenario 2, the results of

which are shown in Figure 6.

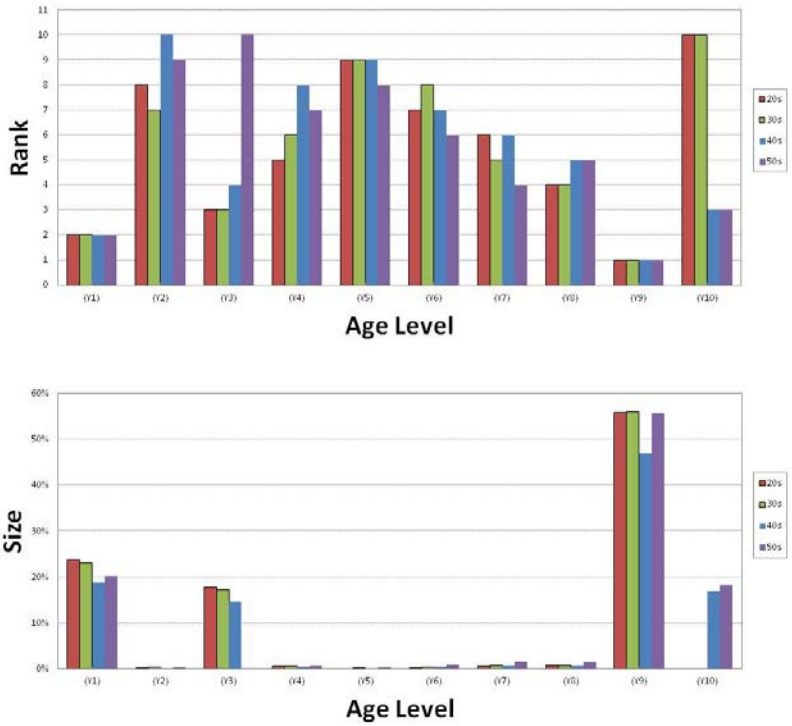

Figure 6. Changes to the Consumption-Expenditure Structure for the Representative Consumer, as a Function of Age Level, in Empirical Study 1

Figure 6 shows that the higher the consumer age level is, the more likely the ratios of the food and clothing/fashion expenditure categories will be reduced in order to facilitate the purchase of a smart pad. Based on the 3rd stage results in Table 10, one can see that among older consumers, there is a lower willingness to buy smart pads that feature iOS, but they exhibit a relatively higher willingness to buy smart pads that feature an Android or Windows mobile OS.

The expected expenditure ranking and expected expenditure share of the representative consumer, as a function of income level, are analyzed in scenario 3, the results of which are shown in Figure 7.

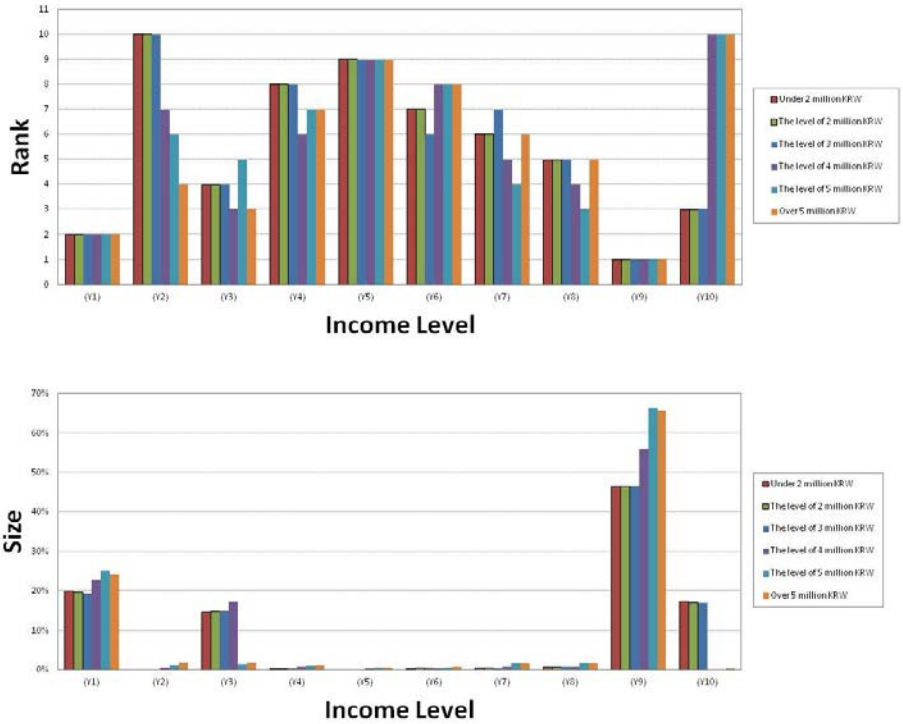

Figure 7. Changes to the Consumption-Expenditure Structure for the Representative Consumer, as a Function of Income Level, in Empirical Study 1

Figure 7 shows that the higher the consumer income level is, the more likely the ratios of the food, alcoholic beverage/tobacco, and saving/insurance categories will be increased in order to facilitate the purchase of a smart pad. Based on the 3rd stage results

in Table 10, one can see that among consumers with higher incomes, there is a greater willingness to buy smart pads that feature iOS, compared to other types of OS.

The expected expenditure ranking and expected expenditure share of the representative consumer, as a function of gender, are analyzed in scenario 4, the results of which are shown in Figure 8.

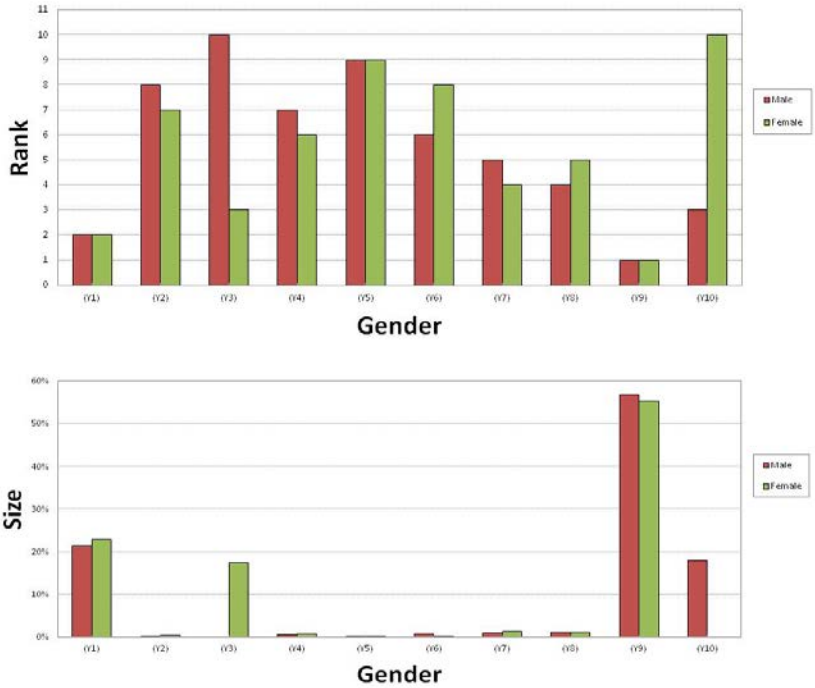

Figure 8. Changes to the Consumption-Expenditure Structure for the Representative Consumer, as a Function of Gender, in Empirical Study 1

Figure 8 shows that among female consumers, the ratio of the clothing/fashion accessories category is more likely to increase in order to facilitate the purchase of a

smart pad, compared to male consumers. Based on the 3rd stage results in Table 10, one can see that among female consumers, there is a greater willingness than among males to buy smart pads that feature iOS, after controlling other factors.

According to Frank et al. (1972) and Fine (1980), analysis of the consumption-expenditure pattern provides a proper standard by which one can understand the tendencies of a target group, and that such analysis is useful for policy-makers as they establish welfare programs or efficient policy. Therefore, the estimation results in this dissertation could be used to establish efficient policies that encourage the adoption of smart devices. In addition, because consumers' consumption-expenditure patterns could be predicted through estimation results and socio-demographic information, the results in this dissertation could be used to develop and undertake market segmentation strategies. In other words, if a company's product is part of the ICT category—the category to which smart pad devices belong—estimation results vis-à-vis consumption-expenditure patterns as a function of household or agent characteristics could inform efficient marketing strategies that increase product sales.

4.2 Multi-stage and Multi-category Discrete Choice Model with Outside goods in Household Products

4.2.1 Introduction

Large and growing volumes of greenhouse gas emissions have prompted an increase in the earth's average temperature, resulting in climate change and environmental problems that have been the subject of perpetual interest. Consumer recognition of global climate change and environmental problems is increasing, and most countries have shown growing concern over these problems. At the global level, attempts at international cooperation have been made to increase collaboration that will work to reduce greenhouse gas emissions; one such attempt is the Kyoto Protocol, which was generated in February 2005. In particular, the Kyoto Protocol identifies the need to control six greenhouse gases—i.e., CO_2, CH_4, N_2O, HFCs, PFCs, and SF_6—as a means of mitigating environmental problems. The bulk of the attention is paid to CO_2 emissions, which many countries are trying to reduce. At the national level, countries are making the effort to encourage consumers to increase green consumption[14] and firms to make green products.[15]

[14] The literature offers a variety of definitions of "green consumption." This section follows that of the Ministry of the Environment in South Korea, according to whom "green consumption" is shown when consumers purchase environmentally friendly products. Therefore, green consumption can help to preserve the environment.

[15] Various definitions of green products have been used in previous researches. For instance, Peattie (1995) defines "green products" as products that exhibit higher environmental and societal performance during production, usage, and disposal, whereas Dangelico and Pontrandolfo (2010) define them as products that have a higher environmental performance, but not necessarily societal performance. Green products are classified according to their characteristics (Rombouts, 1998), level of environmental effects (Hanssen, 1999;

One effective method of resolving environmental issues is to disseminate green products among consumers at the national level, and to encourage their purchase and use. In other words, the revitalization of green consumption, as it is brought to bear on consumer purchasing behavior, should be conducted through government subsidy policies or firm strategies. According to Kilbourne and Beckmann (1998) and Grunert–Beckmann et al. (1997), consumers' environmental concerns constitute an important factor in product purchase decision-making. Therefore, a segmentation strategy based on consumer preference is an effective way of targeting consumers who have environmental concerns (Prendergast and Thompson, 1998). However, Gupta and Ogden (2009) mention that environmental consumerism often bears attitude–behavior inconsistencies—e.g., consumers with environmental concerns are unwilling to pay more money for eco-friendly products (Ottman, 1992; Schlossberg, 1991)—but some studies have derived contrary results (Arbuthnot, 1977; Kellgren and Wood, 1986). To determine the reason for these mixed results vis-à-vis buying behavior, consumer purchase behavior with regard to green products should be analyzed, in detail, on the basis of consumer preference.

In addition, the Korea Consumer Agency (2009)[16] conducted a consumer survey to determine the key factor by which to revitalize green consumption. The results of this survey show that both the South Korean government and consumers have major roles in the diffusion of green products and increases in their consumption. Consequentially,

---

Dangelico and Pontrandolfo, 2010), and types of environmental protection strategy (Park et al., 1999; Rose et al., 1999), among others. In line with Dangelico and Pontrandolfo (2010)—who consider the level of environmental effects in classifying green products—this dissertation includes eco-friendly detergents under the rubric of "green products" that are composed of natural materials.

[16] See http://www.kca.go.kr.

because government policy should be established based on consumer needs, studies of consumer preference for green products is essential to the efficient diffusion of green products.

In consideration of global environmental interests, this section estimates consumer utility with regard to green products while considering the budget allocation stage, via the proposed model; it uses SP data that is collected via a conjoint survey. In particular, this section considers eco-friendly laundry detergent, because most households purchase and use laundry detergent frequently. The results stemming from the proposed methodology highlight implications that can inform various government policies and firm strategies with respect to eco-friendly household products. In addition, I analyze the effect of the budget allocation structure for each category on consumer choices vis-à-vis eco-friendly laundry detergent.

4.2.2 Data and Empirical Model

To collect the SP data, this section uses a conjoint survey method. I have been collected survey data from 1,000 respondents of various ages (20–59 years), between March and May 2012; this survey is executed by a specialized survey company. The interviews are one-to-one—a survey type that generates more reliable data; random sampling is used through a purposive quota sampling method. The empirical study discussed in this section makes use of a sample comprising 957 survey respondents, except those for which there

were missing or unreliable data. The demographic properties of the sample are shown in Table 13.

Table 13. Demographic Properties of Respondents in Empirical Study 2

Category		Respondents	Percentage (%)	Average	Standard Deviation
Total		957	100	-	-
Gender	Male(1)	476	49.7	0.50	0.50
	Female(0)	481	50.3		
Age	20s	248	25.9	38.54 year	10.87
	30s	251	26.2		
	40s	277	28.9		
	50s	181	18.9		
Education Level	Under High school graduate	370	38.7	-	-
	In college or college graduate	556	58.1		
	Over university graduate	31	3.2		
Average monthly income per household (10,000 won)	Under 199	23	2.4	4.19 million won	162.6
	200 ~ 299	107	11.2		
	300 ~ 399	305	31.9		
	400 ~ 499	237	24.8		
	500 ~ 599	176	18.4		
	Over 600	109	11.4		

To analyze consumer preference vis-à-vis eco-friendly laundry detergent, the following six core attributes are selected: production type, brand, possibility of skin irritation, bio-degradable, detergent type, and price. Detailed explanations of each

attribute and its relative level are denoted in Table 14.

Table 14. Attributes and Attribute Levels of Eco-friendly Detergents

Attribute	Level	Explanation
Production Type	Chemical Synthesis Type/ Natural Type	Chemical synthesis type: In the manufacturing process, chemical wastes and pollutants are generated Natural type: In the manufacturing process, a small amount of pollutant is generated
Brand	No Experience with Brand (0)/ Experience with Brand (1)	Consumer experience with the brand
Possibility of Skin Irritation	Yes (0)/ No (1)	Possibility of skin irritation after wearing clothes washed in this detergent
Bio-degradable	Impossible (0)/ Possible (1)	Whether or not the detergent is bio-degradable
Detergent Type	Liquid Form (0)/ Powder Form (1)	The type of detergent
Price (10,000 KRW)	1/3/5	Price of the detergent

Each attribute is set to have different levels; from these attributes and levels, the number of possible alternatives is found to be 96 (2 × 2 × 2 × 2 × 2 × 3 = 96). Since this number of alternatives may burden the respondents, the actual conjoint survey contained only 16 cards in eight equally divided choice sets, which had been constructed via a fractional factorial design. The respondents choose the most preferred alternative from

the three alternatives of each choice set (i.e., among two choices and a no-purchase option). In addition, I divide the respondents into two groups and show them four choice sets. Thus, each respondent provides four answers.

Before choosing the conjoint cards, respondents provide information on the structure of their household consumption expenditures, which equal their household monthly income. This process is similar to that detailed in section 4.1.2. After providing said information, they answer by how much they would change the consumption expenditure for each category in order to facilitate the purchase of eco-friendly detergent; for this purpose, the price of eco-friendly detergent is assumed to be 0.1 million KRW.

Based on the survey data, this section analyzes consumer purchasing behavior with regard to eco-friendly detergent; it does so through the use of the proposed model (case 1 base model), which considers the budget allocation stage. The empirical model is shown as follows, in Eq. (50):

$$
\begin{aligned}
&1st\ step:\ Y_{ij}^{*} = [(One, Income, Sex, Age, Edu) \otimes I]\theta + v_{ij},\ \ j = 1,...,10 \\
&2nd\ step: q_{ij} = [(Income, Sex, Age, Edu) \otimes I]\alpha + \sum_{k=1}^{10} y_{ik}\delta_k + \rho_{ij} \\
&\qquad s.t.\ \sum_j q_{ij} = \sum the\ change\ of\ each\ expenditure\ category \\
&\qquad where, \begin{cases} y_{ij} = 1\ if\ \ Y_{ij}^{*} > 0 \\ y_{ij} = 0\ if\ \ o.w \end{cases}, j = 1,...,10 \\
&3rd\ step:\ U_{il}^{*} = [(P\_type, Brand, P\_skin, Bio, D\_type, price) \otimes I]\beta + \sum_{k=1}^{10} z_{ik}\delta_k^{'} + \gamma_{ij} \\
&\qquad where, z_{ik}: the\ share\ of\ changed\ expenditure\ for\ k^{th}\ category,\ l = 1,2,3
\end{aligned}
\tag{50}
$$

where *Income* describes the household monthly income, which is one of the variables for household characteristics. *Sex*, *Age*, and *Edu* represent agent characteristics: *Sex* takes a value of 1 if a respondent is male and 0 if female, *Age* describes an agent's age level, and *Edu* describes an agent's education level. *P_type*, *Brand*, *P_skin*, *Bio*, *D_type*, and *Price* are attributes of eco-friendly detergent, and they describe the production type, brand, the possibility of skin irritation, whether or not it is bio-degradable, the detergent type, and product price, respectively.

In the product choice part of the survey, three alternatives (i.e., detergent of a chemical synthesis type, detergent of a natural type, and a no-purchase option) are provided to respondents. For identification, the no-purchase option is used as the base alternative.

4.2.3 Results and Discussion

Based on the case 1 base model and the identification process used in this dissertation, I analyze consumer purchasing behavior in terms of consumer budget allocations, for eco-friendly detergent. To perform the Bayesian estimation process, the prior distribution is assumed to be diffused; therefore, parameters are estimated through 10,000 draws that had been drawn from each Markov chain. To exclude the initial point effect, the first 1,000 draws among the 10,000 draws generated from the Bayesian estimation process are discarded; in other words, these records are considered part of a burn-in period. Based on

the remaining 9,000 draws, the mean and variance of the parameters are estimated, and the estimation results are provided below in Tables 15 and 16.

Table 15. Estimation Results in Empirical Study 2

1st step	Variables	Food (Y1)		Alcoholic beverage/ Tobacco (Y2)		Clothing/Fashion Accessories (Y3)		Household commodities (Y4)		Transport (Y5)		Communication (Y6)		Recreation/ Culture (Y7)		Education/Health (Y8)		Saving/Insurance (Y9)		Housing/electricity/ Others (Y10)	
		beta	s.d.	beta	s.d.	beta	s.d.	beta	s.d.	beta	s.d.	beta	s.d.	beta	s.d.	beta	s.d.	beta	s.d.	beta	s.d.
MVP	Constant	**0.7257***	0.2063	**0.4309***	0.2161	-0.1438	0.2143	**-0.8491***	0.3073	0.1839	0.2471	0.0232	0.2490	0.0295	0.2737	-0.3388	0.2952	**-0.7125***	0.2249	**-0.6856***	0.2273
	Income	-0.0373	0.1268	-0.1162	0.1404	**-0.7563***	0.1499	**-0.4553***	0.2235	**-0.5909***	0.1794	**-0.9243***	0.1882	**-0.8154***	0.2144	-0.1848	0.2003	**0.2710***	0.1316	**-0.7214***	0.1645
	Sex	**-0.1154***	0.0413	0.0651	0.0431	0.0055	0.0439	**-0.1813***	0.0637	-0.0442	0.0522	-0.0333	0.0542	**0.1115***	0.0624	0.0008	0.0642	0.0443	0.0439	-0.0418	0.0478
	Age	-0.1153	0.2223	**-0.4548***	0.2299	-0.0834	0.2298	**-0.7987***	0.3263	**-0.6670***	0.2641	**-0.6997***	0.2646	**-0.7437***	0.3182	**-0.5572***	0.3193	-0.3019	0.2349	0.2446	0.2383
	Edu	**-0.8529***	0.4264	**-2.2998***	0.4535	-0.4619	0.4426	-0.3812	0.6549	**-2.3150***	0.5175	**-1.7980***	0.5199	**-2.6870***	0.5740	**-2.5946***	0.6208	0.0902	0.4609	0.0157	0.4838

2nd step	Variables	Size (Q1)		Size (Q2)		Size (Q3)		Size (Q4)		Size (Q5)		Size (Q6)		Size (Q7)		Size (Q8)		Size (Q9)		Size (Q10)	
		beta	s.d.	beta	s.d.	beta	s.d.	beta	s.d.	beta	s.d.	beta	s.d.	beta	s.d.	beta	s.d.	beta	s.d.	beta	s.d.
Multiple continuous	Income	**0.0980***	0.0188	0.0123	0.0148	**-0.0373***	0.0132	-0.0158	0.0102	0.0047	0.0107	0.0077	0.0101	-0.0088	0.0100	-0.0110	0.0100	**0.0418***	0.0156	**-0.0236***	0.0124
	Sex	**0.0191***	0.0060	**-0.0081***	0.0046	**-0.0154***	0.0041	0.0012	0.0032	0.0013	0.0035	-0.0006	0.0032	-0.0015	0.0032	-0.0031	0.0032	-0.0001	0.0050	**0.0125***	0.0039
	Age	**0.1308***	0.0231	**0.0861***	0.0178	**0.0876***	0.0159	**0.0359***	0.0121	**0.0349***	0.0133	0.0196	0.0122	**0.0422***	0.0119	0.0185	0.0120	**0.1299***	0.0192	**0.0698***	0.0152
	Edu	**0.2459***	0.0333	**0.2519***	0.0254	**0.2825***	0.0229	**0.0729***	0.0173	**0.1188***	0.0183	**0.0873***	0.0174	**0.0533***	0.0167	**0.0710***	0.0171	**0.2464***	0.0274	**0.2212***	0.0214
	Y1	**0.6069***	0.0153	**-0.0910***	0.0097	**-0.0765***	0.0080	**-0.0212***	0.0054	**-0.0326***	0.0061	**-0.0270***	0.0055	**-0.0194***	0.0053	**-0.0148***	0.0054	**-0.1685***	0.0107	**-0.0701***	0.0076
	Y2	**-0.1588***	0.0128	**0.4703***	0.0089	**-0.0789***	0.0081	**-0.0097***	0.0056	**-0.0313***	0.0061	**-0.0156***	0.0054	**-0.0134***	0.0053	-0.0072	0.0055	**-0.0360***	0.0114	**-0.0497***	0.0079
	Y3	**-0.1275***	0.0124	**-0.0978***	0.0097	**0.4173***	0.0075	**-0.0140***	0.0055	**-0.0306***	0.0060	**-0.0116***	0.0056	**-0.0130***	0.0055	**-0.0130***	0.0056	**-0.0294***	0.0110	**-0.0396***	0.0079
	Y4	**-0.0511***	0.0159	**0.0295***	0.0134	**-0.0678***	0.0111	**0.3282***	0.0080	**-0.0382***	0.0086	**-0.0324***	0.0079	**-0.0074***	0.0079	**-0.0416***	0.0080	**-0.0337***	0.0135	**-0.0668***	0.0105
	Y5	**-0.0854***	0.0139	**-0.0340***	0.0113	**-0.0530***	0.0092	**-0.0143***	0.0063	**0.3432***	0.0069	**-0.0285***	0.0063	-0.0056	0.0064	**-0.0113***	0.0064	-0.0003	0.0114	**-0.0654***	0.0086
	Y6	0.0200	0.0149	**-0.0427***	0.0110	**-0.0444***	0.0093	**-0.0210***	0.0067	**-0.0348***	0.0072	**0.3127***	0.0066	**-0.0180***	0.0067	**-0.0208***	0.0065	**-0.0733***	0.0114	**-0.0454***	0.0089
	Y7	-0.0162	0.0155	**-0.0541***	0.0122	**-0.0422***	0.0106	-0.0065	0.0076	**-0.0139***	0.0084	**-0.0334***	0.0075	**0.3115***	0.0073	-0.0105	0.0076	**-0.0588***	0.0128	**-0.0398***	0.0099
	Y8	-0.0258	0.0165	0.0023	0.0134	**-0.0556***	0.0111	**-0.0433***	0.0079	**-0.0350***	0.0088	**-0.0304***	0.0080	**-0.0174***	0.0079	**0.3208***	0.0077	**-0.0464***	0.0135	**-0.0425***	0.0106
	Y9	**-0.2617***	0.0166	**-0.1153***	0.0103	**-0.0964***	0.0087	**-0.0234***	0.0057	**-0.0395***	0.0063	**-0.0346***	0.0058	**-0.0268***	0.0055	**-0.0189***	0.0054	**0.7274***	0.0110	**-0.0867***	0.0078
	Y10	**-0.1220***	0.0137	**-0.0551***	0.0114	**-0.0428***	0.0092	**-0.0191***	0.0060	**-0.0363***	0.0064	**-0.0164***	0.0058	**-0.0190***	0.0058	**-0.0141***	0.0059	**-0.0910***	0.0105	**0.4366***	0.0073

3rd stepl	Variables	Detergent of natural type (U1)		Detergent of chemical synthesis type (U2)	
		beta	s.d.	beta	s.d.
MNP	Brand	0.0703	0.0675	**0.1390***	0.0506
	P_skin (“1”=no)	**0.2645***	0.0714	**0.3205***	0.0535
	Bio	**0.3301***	0.0667	-0.0721	0.0528
	D_type(“1”=powder)	0.0646	0.0578	0.0652	0.0475
	Price	**-3.6265***	0.3781	**-2.1683***	0.1719
	Q1	**1.5799***	0.2301	**0.9936***	0.1459
	Q2	**0.8416***	0.3579	**0.4324***	0.2487
	Q3	**0.6621***	0.3967	0.2488	0.2780
	Q4	0.6675	0.6058	0.0714	0.4597
	Q5	0.8849	0.5764	0.3726	0.4184
	Q6	-0.2747	0.6341	-0.7764	0.4759
	Q7	0.7264	0.6609	0.1958	0.5194
	Q8	-0.3849	0.6244	-0.0707	0.4839
	Q9	**0.7408***	0.1966	**0.3647***	0.1560
	Q10	**1.6356***	0.4456	**1.0303***	0.3296

Note: * significant at 10% level

Table 16. Variance–Covariance Matrix in Empirical Study 2

	Y1	Y2	Y3	Y4	Y5	Y6	Y7	Y8	Y9	Y10	Q1	Q2	Q3	Q4	Q5	Q6	Q7	Q8	Q9	Q10	U1	U2
Y1	1	**0.0701***	**0.1465***	**0.1485***	**0.1664***	**-0.1307***	0.0067	0.0240	**-0.5351***	**0.0803***	**-0.0234***	**-0.0139***	-0.0055	-0.0001	0.0000	0.0001	-0.0013	-0.0016	**0.0112***	-0.0023	**-0.3418***	**-0.2268***
Y2		1	**0.3546***	0.0033	**0.2459***	**0.3127***	**0.3405***	**0.1943***	**-0.3682***	**0.1817***	**-0.0357***	0.0007	0.0023	0.0014	**0.0053***	0.0020	0.0006	0.0011	**-0.0178***	0.0022	0.1279	0.1063
Y3			1	**0.3855***	**0.3827***	**0.4056***	**0.4548***	**0.3644***	**-0.2464***	**0.2194***	**-0.0447***	0.0007	**0.0062***	0.0040	**0.0081***	0.0035	0.0014	0.0028	**-0.0244***	0.0033	0.1232	**0.1860***
Y4				1	**0.4146***	**0.3461***	**0.3069***	**0.4995***	**-0.0893***	**0.3880***	**-0.0374***	**-0.0128***	0.0032	**0.0065***	**0.0083***	**0.0046***	0.0018	**0.0046***	**-0.0119***	**0.0075***	0.0383	0.0959
Y5					1	**0.5358***	**0.3177***	**0.3634***	**-0.1838***	**0.4219***	**-0.0483***	-0.0060	0.0030	**0.0052***	**0.0113***	**0.0053***	0.0015	0.0034	**-0.0203***	**0.0080***	-0.0070	0.1101
Y6						1	**0.4074***	**0.5079***	0.0341	**0.4326***	**-0.0570***	0.0020	0.0063	**0.0054***	**0.0112***	**0.0065***	0.0030	**0.0047***	**-0.0204***	**0.0092***	**0.3261***	**0.3758***
Y7							1	**0.4066***	-0.0379	**0.3895***	**-0.0485***	0.0016	0.0057	**0.0040***	**0.0076***	**0.0044***	**0.0036***	0.0035	**-0.0194***	**0.0062***	0.2027	0.1437
Y8								1	-0.0548	**0.3798***	**-0.0470***	-0.0063	0.0048	**0.0060***	**0.0087***	**0.0054***	0.0027	**0.0056***	**-0.0150***	**0.0076***	**0.3713***	**0.2013***
Y9									1	-0.0339	**0.0455***	**0.0116***	0.0005	-0.0010	-0.0014	-0.0019	-0.0011	0.0005	**-0.0268***	-0.0031	**0.1825***	0.0959
Y10										1	**-0.0412***	-0.0058	0.0006	**0.0045***	**0.0085***	**0.0047***	0.0024	0.0035	**-0.0143***	**0.0091***	-0.0048	0.0309
Q1											**0.0315***	**-0.0028***	**-0.0038***	**-0.0015***	**-0.0023***	**-0.0014***	**-0.0011***	**-0.0010***	**-0.0038***	**-0.0032***	-0.0087	-0.0114
Q2												**0.0197***	0.0001	**-0.0008***	**-0.0007***	**-0.0007***	-0.0003	**-0.0007***	**-0.0036***	**-0.0013***	0.0039	0.0030
Q3													**0.0156***	-0.0001	-0.0003	-0.0003	-0.0002	-0.0002	**-0.0017***	**-0.0012***	0.0042	0.0015
Q4														**0.0094***	0.0001	0.0001	0.0000	0.0002	**-0.0006***	-0.0001	0.0019	0.0015
Q5															**0.0109***	0.0002	-0.0001	0.0000	**-0.0014***	0.0000	0.0026	0.0031
Q6																**0.0094***	0.0001	0.0000	-0.0001	-0.0002	0.0014	0.0008
Q7																	**0.0090***	0.0001	0.0000	-0.0001	0.0010	0.0009
Q8																		**0.0093***	**-0.0007***	-0.0001	0.0023	0.0018
Q9																			**0.0225***	**-0.0007***	**-0.0115***	**-0.0086***
Q10																				**0.0144***	0.0026	0.0022
U1																					**1.8079***	**0.6315***
U2																						1

Note: * significant at 10% level

Several implications can be drawn from the estimation results. First, the difference in consumer purchasing behavior as a function of household and agent characteristics is analyzed: in other words, choice and the budget size for each consumption-expenditure category for purchasing eco-friendly detergent are identified as a function of household and agent characteristics. For instance, among consumers with higher income levels, there is a greater tendency to change the budget size for the saving/insurance category, compared to the other categories. Under the choice of the consumption-expenditure categories, consumers with higher income levels are more likely to reduce the budget size for the food and saving/insurance categories in order to facilitate the purchase of eco-friendly detergent, but consumers with lower income levels are more likely to reduce the budget size for the clothing/fashion accessories and housing/electricity/other categories to facilitate its purchase.

Second, the relationships among the consumption-expenditure categories are identified from the 2nd stage estimation results and the variance–covariance matrix. To purchase eco-friendly detergent, the choice of one consumption-expenditure category budget will have a negative effect on the budget sizes of the other categories. Because the price of eco-friendly detergent is much lower than smart devices or smart cars, there are substitution relationships among the consumption-expenditure categories.

Third, the estimation results suggest certain changes to consumption-expenditure categories that significantly affect the purchase behavior vis-à-vis eco-friendly detergent. In other words, those results indicate that a change in the ratios of certain consumption-

expenditure categories affects the choice probability of eco-friendly detergent being purchased: if consumers reduce a high ratio for the clothing/fashion accessories category in their consumption-expenditure structure in order to purchase an eco-friendly product, they will be more likely to purchase a natural-type detergent. In other words, consumers with a higher tendency to buy natural-type detergent are more willing to reduce their budgets related to the clothing/fashion accessories category, and instead use that money to preserve the clothing (i.e., through the use of that detergent).

Finally, the 3rd stage estimation results also show the consumer preference for certain attributes of eco-friendly detergent. With regard to brand, if consumers have a use experience with a brand, the choice probability of it being a chemical-type detergent is increased, but the choice probability of it being a natural-type detergent remains unchanged. As for being bio-degradable, if the detergent is of a bio-degradable type, the choice probability of it being a natural-type detergent is increased, but the choice probability of it being a chemical-type detergent remains unchanged. In the following section, several scenario analyses are conducted to examine trends vis-à-vis consumption-expenditure categories for purchasing eco-friendly detergent.

Scenario Analysis

To examine the trends among consumption-expenditure categories with regard to the purchase of eco-friendly detergent, the expected expenditure share of the representative

consumer is analyzed, based on the estimation results. According to socio-demographic information on the representative consumer and estimation results, changes to the consumption–expenditure structure in each category with respect to the purchase of eco-friendly detergent are as shown in Figure 9.

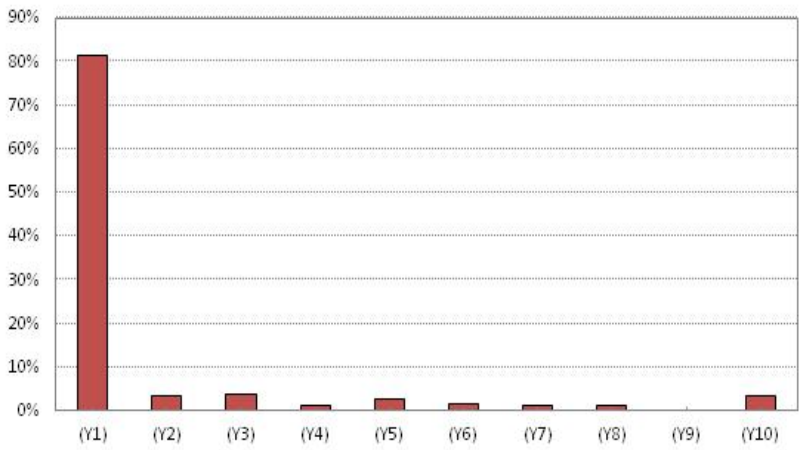

Figure 9. Changes to the Consumption-Expenditure Structure for the Representative Consumer, in Empirical Study 2

From Figure 9, one can see that the representative consumer generally reduces the food category budget in order to purchase eco-friendly detergent; this budget reduction accounts for about 80% of all expenditure change. To examine the impact of socio-demographic level on changes to the consumption-expenditure structure, four scenario analyses are conducted through the use of four socio-demographic variables: education level, age level, income level, and gender. The expected expenditure ranking

and expected expenditure share of the representative consumer, as a function of education level, are analyzed in scenario 1, the results of which are shown in Figure 10.

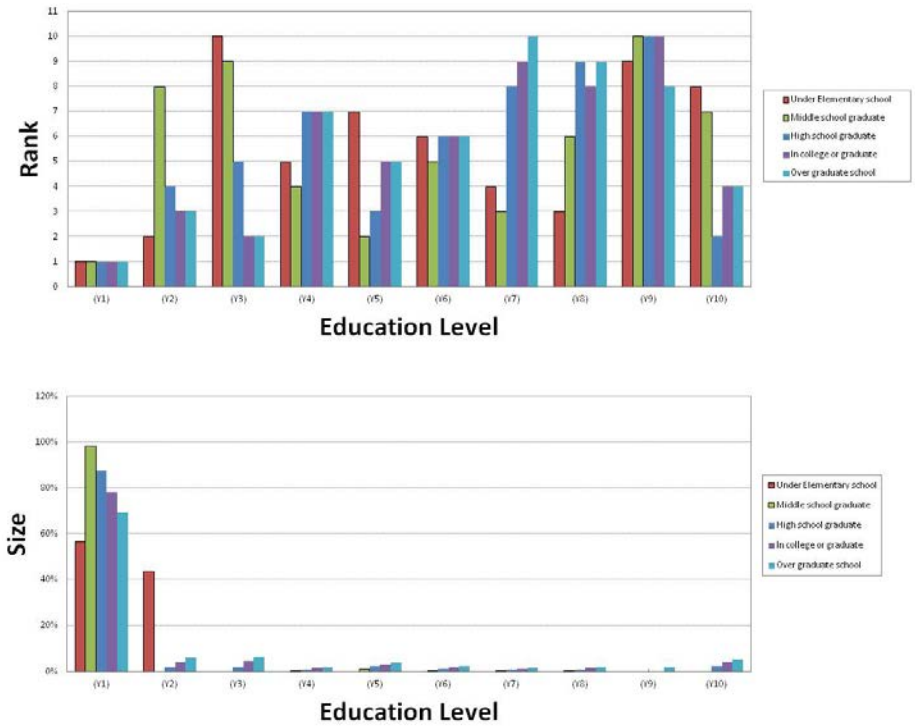

Figure 10. Changes to the Consumption-Expenditure Structure for the Representative Consumer, as a Function of Education Level, in Empirical Study 2

Figure 10 shows that among consumers with a higher education level, the ratio of expenditure for the clothing/fashion accessories category is increased to facilitate the purchase of eco-friendly detergent. Based on the 3rd stage results in Table 14, when consumers have a higher education level, they tend to be more willing to buy detergent of a natural type rather than of a chemical type.

The expected expenditure ranking and expected expenditure share of the representative consumer as a function of age level are analyzed in scenario 2, the results of which are shown in Figure 11.

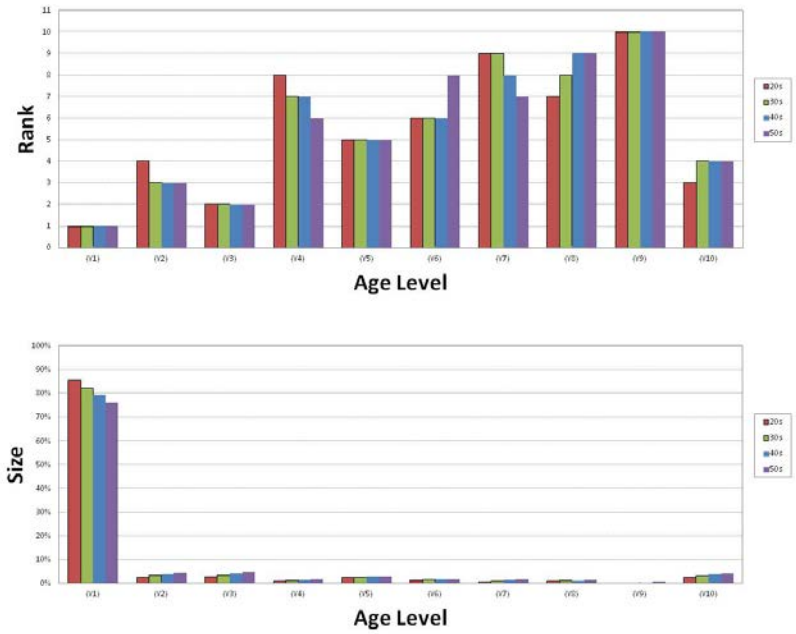

Figure 11. Changes to the Consumption-Expenditure Structure for the Representative Consumer, as a Function of Age Level, in Empirical Study 2

Figure 11 shows that the higher a consumer's age is, the more the ratio of the food category is likely to be reduced in order to purchase eco-friendly detergent. Based on the 3rd stage results in Table 14, the higher a consumer's age is, the less willing he or she will be to buy detergent of a natural type, but he or she will have a relatively higher willingness to buy detergent of a chemical type.

The expected expenditure ranking and expected expenditure share of the representative consumer as a function of income level are analyzed in scenario 3, the results of which are shown in Figure 12.

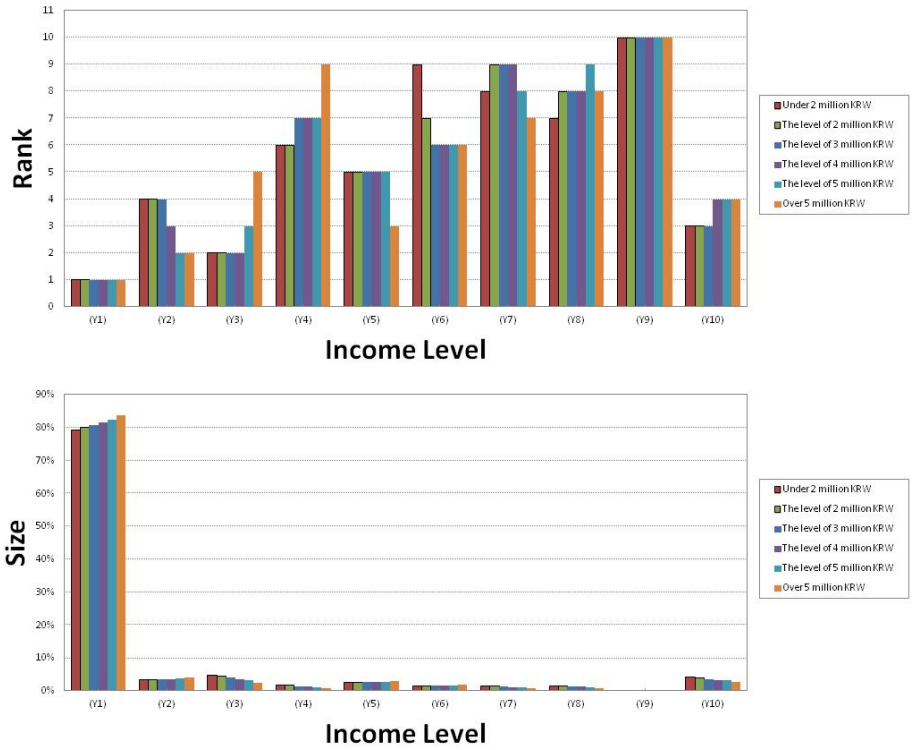

Figure 12. Changes to the Consumption-Expenditure Structure for the Representative Consumer, as a Function of Income Level, in Empirical Study 2

Figure 12 shows that the higher a consumer's income level is, the more the ratios of the clothing/fashion accessories and housing/electricity/other categories are likely to be increased, but the ratio of the food category will increase relatively in order to

facilitate the purchase of eco-friendly detergent. Based on the 3rd stage results in Table 14, among higher-income consumers, there is a greater willingness to buy detergent of a natural type rather than of a chemical type.

The expected expenditure ranking and expected expenditure share of the representative consumer as a function of gender are analyzed in scenario 4, the results of which are shown in Figure 13.

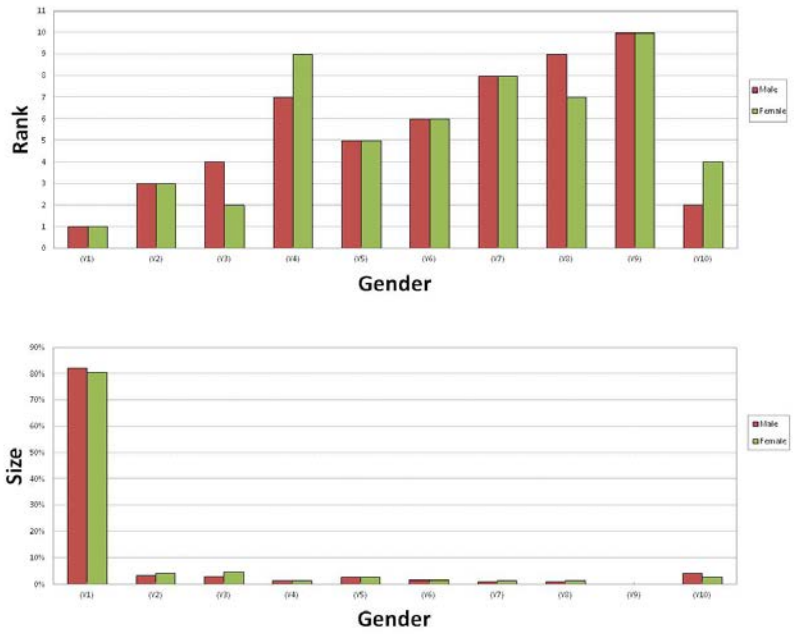

Figure 13. Changes to the Consumption-Expenditure Structure for the Representative Consumer, as a Function of Gender, in Empirical Study 2

Figure 13 shows that changes to the consumption-expenditure structure as a function of gender are almost similar to those expected for other categories. Therefore, no

different purchasing behavior is expected with regard to new detergent.

Because analysis of the consumption-expenditure pattern provides a proper standard by which one can understand the tendencies of a target group as policy-makers establish welfare programs or efficient policy, the estimation results outlined in this dissertation could be utilized to establish efficient policies that encourage the adoption of green products. In addition, because consumers' consumption-expenditure patterns can be predicted through estimation results and socio-demographic information, the results in this dissertation could be used to develop and undertake market segmentation strategies. In other words, if a company's product is part of the household commodity category—the category to which eco-friendly detergent belongs—estimation results vis-à-vis the consumption-expenditure pattern as a function of household or agent characteristics could inform efficient marketing strategies that increase product sales.

4.3 Multi-stage and Multi-category Discrete-continuous Choice Model with Outside goods in Automobiles

4.3.1 Introduction

Over the past 10 decades, humans have continuously pursued rapid economic growth through industrial restructuring. In the growing automotive industry, the technology gap between developing and developed countries has significantly diminished, making global competition more fierce than ever. In addition, nations have continued to enhance fuel-

efficiency capabilities and safety and environmental regulations, all of which are becoming more homogeneous in nature. Such activities have resulted from the formation of a worldwide social consensus about global warming and energy consumption. As a result, electric automobiles as a future mode of transportation have become a possible means of coping with high oil prices and CO_2 levels. For instance, according to the California Energy Commission (2007), the U.S. government and companies in the United States have accelerated the production of plug-in hybrid electric cars and stepped up efforts to make fully electric comprise 75% of all automobiles by 2050.

Due to industrial development in developed and developing countries alike, fossil fuel depletion has become a serious problem, causing increases in both oil prices and industrial development costs. To resolve the problem of high oil prices, many researchers around the world have developed renewable energy and alternative technologies. In addition, industrial growth in both developed and developing nations has generated environmental pollution, in the form of greenhouse gas emissions, *inter alia*. Under these circumstances, electric automobiles have emerged as viable alternatives that can mitigate high oil prices and address environmental concerns.

According to Suehiro et al. (2010), energy usage in developing countries is expected to increase 1.6-fold as a result of a sharp increase in automotive sector needs. Because of this, the level of CO_2 emissions in 2050 is expected to be 1.8-fold higher than the current level. Resolutions to problems associated with high energy consumption, CO_2 emissions, and oil prices demand international cooperation. This section discusses how

electric automobiles are a viable means of addressing high oil prices and environmental concerns; it also considers smart cars, a recent trend in the automobile industry.

To enable the successful vitalization of electric and smart cars, policies should be based on consumers' perspectives on technology development, rather than the providers' technology-oriented policies or development analyses. In other words, consumer demand analysis should first be conducted to analyze the direction of technology development regarding the electric car's battery charge time, capacity, charging method, and smart options, as well as other technology-related issues. Therefore, consumer demand and market analyses are essential to the enhancement of activities among related industries and will thus improve the likelihood of success of a new product in the automotive market.

Due to relatively short cycles of innovation and rapidly changing consumer preferences, market forecasting vis-à-vis new products determines not only the legitimacy of investment for producers but directs the industry's policy-makers to eliminate uncertainty. Therefore, the purpose of this section is to analyze consumers' car-usage behavior after the introduction of the electric car and the smart car. This section uses consumer-SP data and the multi-stage and multi-category discrete-continuous choice model. Because the electric and smart automobiles considered in this section have not yet arrived on the market, an RP approach would not extract sufficient consumer preference information about electric and smart cars. Accordingly, the consumer-stated preference (i.e., SP) approach—especially, a conjoint analysis—is more relevant to meeting the

objective of this section. In addition, consumer usage patterns should be considered. Because the case 2 base model can be used to consider product choice and usage, it proves adequate for the analysis of consumer choice patterns and the usage of passenger cars. In addition, the effect of budget allocation for each category on the choice of electric and smart car is analyzed.

4.3.2 Data and Empirical Model

This section uses survey data to analyze the impact of introducing electric and smart cars to the existing automobile market, and especially consumer usage behavior. Conjoint analysis is suitable for this research, because only a few market datasets relating to electric and smart cars exist. In other words, the electric and smart automobile markets are still nascent.

The survey has been conducted from March 2012 to May 2012 among 675 households whose surveyed member is aged between 20 and 59 years and who lived in one of six metropolitan cities in South Korea—i.e., Seoul, Busan, Daegu, Inchon, Gwangju, or Daejeon. The survey is executed by an experienced domestic survey company; a structured questionnaire is used for individual interviews, to help ensure the accuracy of the survey results. The sample from this study is extracted by using a purposive quota sampling method, especially with respect to gender and age. Thus, the empirical study addressed in this section uses a dataset comprising 616 participants,

except for those records for whom data are missing or unreliable. The demographic characteristics of the sample that participated in the survey are presented in Table 17.

Table 17. Demographic Characteristics of the Sample in Empirical Study 3

Category		Respondents	Percentage (%)	Average	Standard Deviation
Total		616	100	-	-
Gender	Male(1)	294	47.7	0.48	0.50
	Female(0)	322	52.3		
Age	20s	164	26.6	38.44 year	10.89
	30s	164	26.6		
	40s	170	27.6		
	50s	118	19.2		
Education Level	Under High school graduate	249	40.4	-	-
	In college or college graduate	351	57.0		
	Over university graduate	16	2.6		
Average monthly income per household (10,000 won)	Under 199	11	1.8	4.14 million won	149.6
	200 ~ 299	76	12.3		
	300 ~ 399	204	33.1		
	400 ~ 499	140	22.7		
	500 ~ 599	119	18.3		
	Over 600	66	10.7		

This empirical study establishes the attributes and attribute levels of automobiles, to undertake conjoint analysis. The attributes of the automobiles include the fuel type,

vehicle type, fuel cost, purchase price, accessibility of fueling stations, and smart car option. The attributes and attribute levels of automobiles that are considered in this study in order to analyze consumer preferences vis-à-vis automobile choice are listed in Table 18.

Table 18. Attributes and Attribute Levels of Smart Cars

Attribute	Attribute Level	Explanation
Fuel Type	Gasoline/Diesel/Hybrid (Gasoline + Battery Type)/Electric (Battery)	Compared to the existing fossil fuel cars, electric automobiles need a battery-charging time of about 4 h, or a battery-replacement time of about 2 min
Vehicle Type	SUV (RV)/ General vehicle	Vehicle type
Fuel Cost (KRW/km)	200/100/50	The fuel cost is defined as the cost associated with 1 km of driving
Purchase Price (1 Million KRW)	25/30/35/40	Purchase price
Accessibility of Fueling Station (%)	100/80/50	When the current level of car's gasoline station is defined as 100, accessibility of fueling station is defined by the number of fueling stations for its specific fuel type
Smart Car Option	Yes/No	Whether or not a smart car option is provided

The number of possible alternatives is 4 × 2 × 3 × 4 × 3 × 2 = 576, based on the attributes and attribute levels of automobiles as established for this empirical study.

Because respondents cannot express their preferences with regard to this total number of possible cases, I reduce the number of possible cases to 16, through the use of a fractional factorial design—a method used to maintain orthogonality among attributes. Based on these 16 alternatives, this study features four sets that consist of four alternatives; a survey is then executed by using these design cards. In addition, I divide the respondents into three groups and show them three choice sets. Thus, the respondents choose from each choice set the alternative that provides the highest utility; thus, each respondent ultimately provides three answers. In the next section, consumer preferences are analyzed based on the survey results, through the use of the proposed model.

Before choosing from the conjoint cards, the respondents are asked about the structure of their household consumption expenditures, which equal their household monthly income. This process is similar to that detailed in section 4.1.2. After providing said information, they answer by how much they would change the consumption expenditure for each category in order to facilitate the purchase of a smart car; for this purpose, I assume the price of a smart car to be 1 million KRW, on a 36-month installment-payment plan.

Based on the survey data, this section analyzes consumer purchasing behavior with regard to smart cars, via the proposed model (case 2 base model), which considers the budget allocation stage. The empirical model is shown as follows, in Eq. (51):

$$1st\ step:\ Y_{ij}^{*} = [(One, Income, Age, N\_Family, Drive) \otimes I]\theta + \nu_{ij},\ \ j = 1,...,10$$

$$2nd\ step: q_{ij} = [(Income, Sex, Age, Edu) \otimes I]\alpha + \sum_{k=1}^{10} y_{ik}\delta_k + \rho_{ij}$$

$$s.t.\ \sum_j q_{ij} = \sum the\ change\ of\ each\ expenditure\ category$$

$$where, \begin{cases} y_{ij} = 1\ if\ \ Y_{ij}^{*} > 0 \\ y_{ij} = 0\ if\ \ o.w \end{cases}, j = 1,...,10$$

$$3rd\ step:\ U_{il}^{*} = [(F\_\cos t, \Pr ice, F\_station, V\_smart) \otimes I]\beta + \sum_{k=1}^{10} z_{ik}\delta_k^{'} + \gamma_{ij}$$

$$where, z_{ik}: the\ share\ of\ changed\ expenditure\ for\ k^{th}\ category,\ l = 1,2,3$$

$$4th\ step:\ m_i = [(Income, Sex, Age, Edu, N\_Family) \otimes I]\gamma + \sum_{k=1}^{4} U_{ik}\delta_k^{'} + \eta_{ij}$$

$$where, \begin{cases} U_{ik} = 1\ if\ \max(U_i^{*}) = U_{ik}^{*} > 0 \\ U_{ik} = 0\ if\ o.w \end{cases} \tag{51}$$

where *Income* and *N_Family* represent household characteristics—namely, the household monthly income and the number of family members, respectively. *Sex*, *Age*, *Edu*, and *Drive* represent agent characteristics: *Sex* takes a value of 1 if a respondent is male and 0 if female, *Age* describes an agent's age level, and *Edu* describes an agent's education level. *Drive* describes whether or not the agent is a driver. *F_cost*, *Price*, *F_station*, and *V_smart* are attributes of a smart car, and they describe the fuel type, purchase price, accessibility of fueling station, and a dummy variable that represents the interaction between vehicle type and smart car option, respectively.

In the product choice part, three alternatives pertaining to automobile type—i.e., of a diesel, hybrid, or electric type—are provided to respondents. For identification, the

gasoline-type automobile is used as the base alternative.

4.3.3 Results and Discussion

Based on the case 2 base model and the identification process used in this dissertation, I analyze consumer purchasing behavior in terms of consumer budget allocations, for smart cars. To perform the Bayesian estimation process, the prior distribution is assumed to be diffused; therefore, the parameters are estimated through 10,000 draws that had been drawn from each Markov chain. To exclude the initial point effect, the first 1,000 draws among the 10,000 draws generated from the Bayesian estimation process are discarded; in other words, these records are considered part of a burn-in period. Based on the remaining 9,000 draws, the mean and variance of the parameters are estimated, and the estimation results are provided below in Tables 19 and 20.

Table 19. Estimation Results in Empirical Study 3

1st step	Variables	Food (Y1)		Alcoholic beverage/ Tobacco (Y2)		Clothing/Fashion Accessories (Y3)		Household commodities (Y4)		Transport (Y5)		Communication (Y6)		Recreation/ Culture (Y7)		Education/Health (Y8)		Saving/Insurance (Y9)		Housing/electricity/ Others (Y10)	
		beta	s.d.	beta	s.d.	beta	s.d.	beta	s.d.	beta	s.d.	beta	s.d.	beta	s.d.	beta	s.d.	beta	s.d.	beta	s.d.
MVP	Constant	**0.3890***	0.1317	-0.1110	0.1315	**0.5053***	0.1263	**-0.9416***	0.1436	**-0.3972***	0.1306	**-0.8691***	0.1381	-0.2076	0.1374	**-1.4491***	0.1556	**2.2283***	0.2102	0.0702	0.1258
	Income	**-1.1282***	0.2157	**-0.7550***	0.2139	**-1.1880***	0.2132	**-0.4730***	0.2356	**-0.7378***	0.2149	**-0.4427***	0.2254	**-0.8600***	0.2321	**-0.8341***	0.2402	-0.1344	0.2889	-0.3383	0.2107
	Age	**0.1297***	0.0659	**0.1827***	0.0633	**-0.1196***	0.0638	-0.0594	0.0707	**0.1144***	0.0639	**0.2124***	0.0661	-0.0513	0.0663	**0.1521***	0.0690	-0.0437	0.0903	0.0488	0.0626
	N_Family	**1.6763***	0.3320	0.4782	0.3239	-0.0580	0.3073	**0.7639***	0.3518	**0.6055***	0.3189	**0.7910***	0.3365	-0.4337	0.3332	**2.8681***	0.3750	**-2.0750***	0.5113	-0.1572	0.3126
	Drive	**-0.2092***	0.0643	-0.0753	0.0621	0.0754	0.0607	**0.2112***	0.0670	0.0610	0.0628	**-0.1125***	0.0653	**0.2092***	0.0652	**-0.1661***	0.0675	-0.0659	0.0918	-0.0817	0.0609

2nd step	Variables	Size (Q1)		Size (Q2)		Size (Q3)		Size (Q4)		Size (Q5)		Size (Q6)		Size (Q7)		Size (Q8)		Size (Q9)		Size (Q10)	
		beta	s.d.	beta	s.d.	beta	s.d.	beta	s.d.	beta	s.d.	beta	s.d.	beta	s.d.	beta	s.d.	beta	s.d.	beta	s.d.
Multiple continuous	Income	-0.0137	0.0254	0.0087	0.0219	0.0198	0.0217	0.0290	0.0210	0.0292	0.0220	0.0019	0.0205	0.0186	0.0206	0.0161	0.0223	0.0150	0.0383	0.0146	0.0242
	Sex	-0.0072	0.0073	0.0035	0.0060	0.0006	0.0062	0.0008	0.0059	0.0071	0.0062	0.0000	0.0059	0.0030	0.0058	-0.0047	0.0065	0.0010	0.0110	-0.0017	0.0068
	Age	**0.0702***	0.0300	0.0255	0.0246	0.0284	0.0256	-0.0066	0.0239	0.0077	0.0256	0.0181	0.0239	0.0239	0.0241	-0.0075	0.0265	**0.2468***	0.0426	**0.1169***	0.0281
	Edu	**0.2396***	0.0500	**0.0848***	0.0415	0.0592	0.0421	0.0329	0.0396	**0.0747***	0.0422	0.0372	0.0405	-0.0008	0.0396	**0.0973***	0.0448	**0.4403***	0.0728	**0.1677***	0.0465
	Y1	**0.1949***	0.0123	0.0041	0.0100	-0.0107	0.0103	-0.0019	0.0095	-0.0151	0.0102	0.0051	0.0095	0.0050	0.0096	0.0039	0.0107	**-0.0949***	0.0198	-0.0100	0.0115
	Y2	0.0129	0.0130	**0.1205***	0.0098	-0.0035	0.0101	-0.0025	0.0096	-0.0023	0.0101	0.0002	0.0093	0.0015	0.0096	-0.0009	0.0108	**-0.0800***	0.0200	0.0006	0.0117
	Y3	-0.0032	0.0123	0.0053	0.0100	**0.1184***	0.0102	0.0032	0.0095	-0.0060	0.0102	-0.0061	0.0094	0.0079	0.0097	-0.0013	0.0108	**-0.0409***	0.0189	-0.0153	0.0117
	Y4	-0.0216	0.0133	-0.0136	0.0109	0.0002	0.0109	**0.0981***	0.0102	-0.0075	0.0110	-0.0030	0.0101	**-0.0195***	0.0102	-0.0107	0.0115	0.0037	0.0201	-0.0179	0.0123
	Y5	-0.0045	0.0130	0.0044	0.0100	-0.0009	0.0103	-0.0005	0.0096	**0.1340***	0.0101	-0.0040	0.0093	-0.0040	0.0096	-0.0132	0.0110	**-0.0689***	0.0195	-0.0080	0.0119
	Y6	-0.0166	0.0135	-0.0081	0.0107	-0.0134	0.0109	-0.0067	0.0102	**-0.0277***	0.0108	**0.0861***	0.0102	-0.0111	0.0102	-0.0064	0.0113	0.0236	0.0195	-0.0038	0.0126
	Y7	-0.0145	0.0133	0.0000	0.0103	0.0117	0.0105	-0.0060	0.0099	-0.0023	0.0104	-0.0027	0.0099	**0.1045***	0.0099	-0.0160	0.0110	-0.0265	0.0199	**-0.0207***	0.0120
	Y8	-0.0184	0.0131	**-0.0250***	0.0104	0.0038	0.0101	-0.0104	0.0100	-0.0058	0.0106	0.0055	0.0099	-0.0102	0.0099	**0.1710***	0.0109	**-0.0620***	0.0198	-0.0126	0.0118
	Y9	**-0.1023***	0.0151	**-0.0460***	0.0119	**-0.0311***	0.0123	-0.0158	0.0113	**-0.0245***	0.0120	**-0.0233***	0.0116	**-0.0187***	0.0114	**-0.0298***	0.0128	**0.5223***	0.0228	**-0.0919***	0.0137
	Y10	**-0.0232***	0.0125	-0.0066	0.0102	-0.0150	0.0102	-0.0049	0.0095	-0.0062	0.0101	0.0011	0.0096	-0.0039	0.0093	-0.0021	0.0106	**-0.0618***	0.0194	**0.1710***	0.0118

3rd step	Variables		Vehicle with diesel type (U1)		Vehicle with hybrid type (U2)		Vehicle with electric type (U3)	
			beta	s.d.	beta	s.d.	beta	s.d.
MNP	Q1		-0.2847	0.6946	0.0340	0.6354	0.6463	0.4676
	Q2		1.6568	1.3459	**2.2120***	1.2969	1.0572	0.8976
	Q3		-0.1803	1.3419	0.7015	1.2668	0.6271	0.8105
	Q4		-2.0843	2.1404	-0.6372	2.1408	-0.9347	1.3353
	Q5		-0.0388	1.3112	1.0812	1.2732	**1.8914***	0.8244
	Q6		-2.1455	2.4110	-0.2214	2.0799	0.5902	1.3750
	Q7		**3.1800***	1.9313	**4.7023***	1.9231	**2.2946***	1.2666
	Q8		-0.6313	1.1077	**3.0757***	0.9907	**1.1610***	0.7081
	Q9		-0.1975	0.3781	-0.1695	0.2917	0.2555	0.2665
	Q10		-0.8131	0.7512	0.0426	0.6959	0.1648	0.4985
	F_cost	beta	**-0.4619***					
		s.d.	0.0503					
	Price	beta	**-0.2411***					
		s.d.	0.0544					
	F_station	beta	**0.7970***					
		s.d.	0.1697					
	V_smart	beta	**0.3724***					
		s.d.	0.0877					

4th step	Variables	Vehicle Usage	
		beta	s.d.
Single continuous	Income	**7.7060***	1.2447
	Sex	0.1731	0.3463
	Age	**-3.2739***	1.6814
	Edu	1.0473	3.2575
	N.Family	-0.1285	1.8753
	U1	**7.5261***	1.9509
	U2	**17.9369***	2.2383
	U3	**8.9020***	1.9747
	U4	**12.5253***	1.9540

Note: * significant at 10% level

Table 20. Variance–Covariance Matrix in Empirical Study 3

	Y1	Y2	Y3	Y4	Y5	Y6	Y7	Y8	Y9	Y10	Q1	Q2	Q3	Q4	Q5	Q6	Q7	Q8	Q9	Q10	U1	U2	U3	M
Y1	1.0000	**0.5598***	**0.6571***	**0.5203***	**0.4445***	**0.6131***	**0.4719***	**0.5002***	**-0.2979***	**0.5974***	-0.0010	-0.0028	0.0023	0.0002	0.0035	-0.0022	-0.0029	-0.0012	**-0.0806***	0.0034	0.0981	-0.0938	-0.0887	0.3531
Y2		1.0000	**0.6125***	**0.5219***	**0.3167***	**0.3723***	**0.5748***	**0.3142***	**-0.2429***	**0.4249***	-0.0009	-0.0016	0.0027	0.0011	0.0035	-0.0013	-0.0018	-0.0004	**-0.0759***	0.0053	-0.1654	-0.1858	-0.0872	**0.7111***
Y3			1.0000	**0.6835***	**0.3850***	**0.5010***	**0.7160***	**0.3944***	**-0.1926***	**0.5077***	-0.0012	-0.0022	0.0031	0.0009	0.0045	-0.0018	-0.0023	-0.0001	**-0.0889***	0.0061	0.0372	-0.2239	-0.1008	**0.4198***
Y4				1.0000	**0.5235***	**0.5810***	**0.6382***	**0.4614***	**-0.1592***	**0.6322***	-0.0005	-0.0009	0.0027	0.0021	0.0067	-0.0012	-0.0008	0.0009	**-0.0935***	0.0085	0.1261	-0.1415	-0.0385	-0.3685
Y5					1.0000	**0.6101***	**0.3930***	**0.4375***	**-0.2241***	**0.4895***	0.0007	-0.0006	0.0019	0.0014	0.0057	-0.0013	-0.0008	0.0016	**-0.0741***	0.0055	0.0574	**-0.3836***	**-0.1395***	**-0.8661***
Y6						1.0000	**0.5307***	**0.6999***	**-0.3479***	**0.6389***	0.0020	-0.0003	0.0026	0.0020	0.0072	-0.0013	-0.0006	0.0027	**-0.0922***	0.0065	0.1519	-0.2442	-0.1273	**-0.6928***
Y7							1.0000	**0.4787***	**-0.2284***	**0.5249***	-0.0006	-0.0009	0.0030	0.0019	0.0055	-0.0010	-0.0009	0.0013	**-0.0892***	0.0085	-0.0985	**-0.4310***	**-0.2339***	**1.0735***
Y8								1.0000	**-0.4020***	**0.5389***	0.0017	0.0002	0.0023	0.0020	0.0058	-0.0009	-0.0004	0.0025	**-0.0784***	0.0068	0.1345	-0.2319	**-0.1860***	0.1098
Y9									1.0000	**-0.4377***	-0.0009	-0.0002	-0.0018	-0.0004	-0.0027	0.0002	0.0000	-0.0026	**0.0244***	-0.0023	-0.0097	0.2372	0.0878	0.3458
Y10										1.0000	0.0008	-0.0012	0.0025	0.0013	0.0057	-0.0016	-0.0012	0.0009	**-0.0845***	0.0063	0.1142	-0.1630	-0.0265	-0.1514
Q1											**0.0230***	0.0004	0.0003	-0.0001	0.0003	0.0001	0.0000	0.0001	**-0.0077***	-0.0008	0.0069	0.0069	-0.0032	0.0018
Q2												**0.0164***	0.0002	0.0001	0.0000	0.0000	0.0002	0.0000	**-0.0017***	-0.0002	0.0012	0.0012	-0.0010	0.0000
Q3													**0.0168***	0.0000	0.0001	0.0000	0.0002	0.0000	**-0.0025***	-0.0001	-0.0013	-0.0013	-0.0015	-0.0010
Q4														**0.0153***	0.0001	0.0000	0.0001	0.0000	-0.0010	0.0001	-0.0009	-0.0009	-0.0008	-0.0005
Q5															**0.0169***	0.0001	0.0001	0.0001	**-0.0030***	-0.0002	0.0040	0.0040	-0.0032	-0.0003
Q6																**0.0153***	0.0000	0.0002	-0.0002	-0.0001	-0.0013	-0.0013	0.0005	-0.0001
Q7																	**0.0155***	0.0000	-0.0007	0.0000	0.0004	0.0004	0.0004	0.0005
Q8																		**0.0183***	**-0.0027***	-0.0004	-0.0015	-0.0015	-0.0020	-0.0007
Q9																			**0.0538***	**-0.0046***	-0.0095	-0.0095	0.0369	0.0151
Q10																				**0.0209***	-0.0065	-0.0065	-0.0025	-0.0035
U1																					**3.7106***	0.3462	**0.7297***	**-10.6053***
U2																						**2.8041***	**0.4438***	0.7875
U3																							1.0000	**-2.7965***
M																								**69.9613***

Note: * significant at 10% level

Several implications can be drawn from the estimation results. First, the difference in consumer purchasing behavior as a function of household and agent characteristics is analyzed; in other words, choice and the budget size for each consumption-expenditure category for purchasing a smart car are identified as a function of household and agent characteristics. For instance, among consumers from large families, there is a greater tendency to change the budget size for the food, household commodity, transport, communication, and education/health categories, compared to the other categories; those from small families, on the other hand, are more likely to choose to change the saving/insurance category. This means that people from large families are more likely to reduce their current consumption expenditures to purchase smart cars than to reduce the budget associated with the saving/insurance category.

Second, the relationships among the consumption-expenditure categories are identified from the 2nd stage estimation results and the variance–covariance matrix. To purchase a smart car, choosing to change the budget associated with the saving/insurance category has a negative effect on the budget sizes of the other categories; if the budget size for the saving/insurance category is increased, the budget size for other categories will be reduced. Therefore, there is a substitution relationship between the saving/insurance category budget size and those of the other categories.

Third, the estimation results suggest certain changes to consumption-expenditure categories that significantly affect the purchase behavior vis-à-vis smart cars. In other words, those results indicate that a change in the ratios of certain consumption-

expenditure categories affects the choice probability of smart cars being purchased; if consumers reduce a higher ratio for the alcoholic beverage/tobacco category in their consumption-expenditure structure in order to purchase smart cars, they will be more likely to purchase a hybrid-type automobile. If consumers reduce the higher ratio within their consumption-expenditure structures that is associated with the transport category—which excludes the purchase of automobiles—in order to purchase smart cars, they will be more likely to purchase electric-type automobiles.

Finally, the 3rd and 4th stage estimation results also show consumer preferences with regard to certain smart car attributes. With regard to SUV-type automobiles that feature a smart option, consumers have a positive preference for that alternative. As for car usage behavior in the 4th stage, if consumers have higher income levels and are young, their car usage is likely to be higher than among consumers with different income or age levels: because people with higher incomes are less sensitive to fuel costs, car usage is relatively higher. In addition, consumers who choose in the 3rd stage a diesel-type automobile show the highest level of car usage. In the following section, several scenario analyses are conducted to examine trends vis-à-vis consumption-expenditure categories for purchasing smart cars.

Scenario Analysis

To examine the trends among consumption-expenditure categories with regard to the

purchase of smart cars, the expected expenditure share of the representative consumer is analyzed, based on the estimation results. According to socio-demographic information on the representative consumer and estimation results, changes to the consumption-expenditures structure in each category with respect to the purchase of smart cars are as shown in Figure 14.

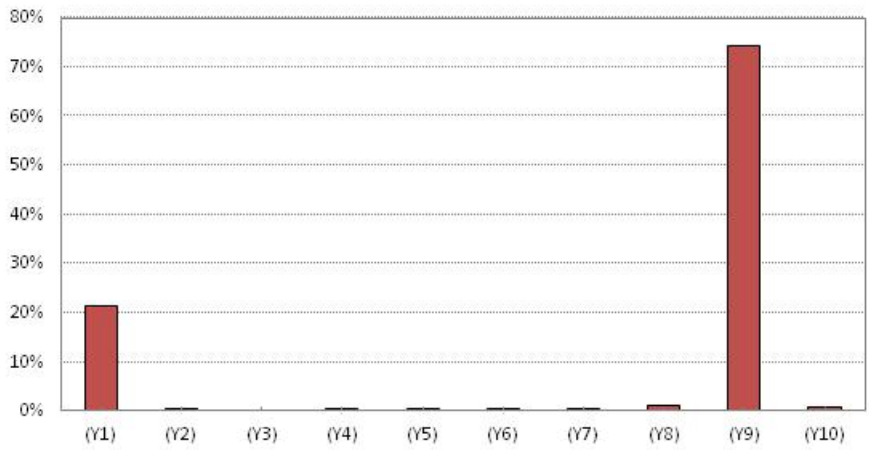

Figure 14. Changes to the Consumption-Expenditure Structure for the Representative Consumer, in Empirical Study 3

From Figure 14, one can see that the representative consumer generally reduces food and saving/insurance category budgets in order to purchase a smart car; this budget reduction accounts for about 95% of all expenditure change. To examine the impact of socio-demographic level on changes to the consumption-expenditure structure, four scenario analyses are conducted by using four socio-demographic variables: education

level, age level, income level, and gender. The expected expenditure ranking and expected expenditure share of the representative consumer, as a function of education level, are analyzed in scenario 1, the results of which are shown in Figure 15.

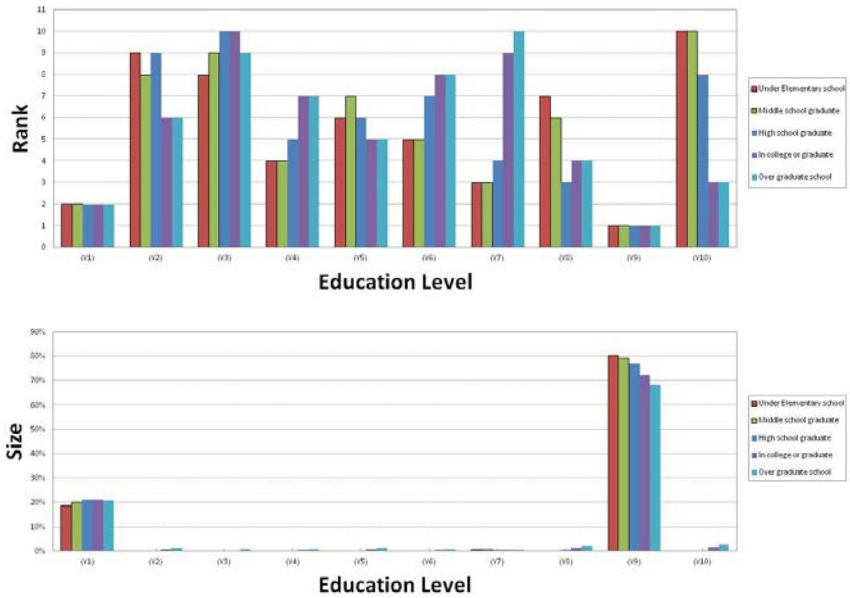

Figure 15. Changes to the Consumption-Expenditure Structure for the Representative Consumer, as a Function of Education Level, in Empirical Study 3

Figure 15 shows that among consumers with higher education levels, the ratio of expenditure for the saving/insurance category is reduced, and those of the alcoholic beverage/tobacco, transport, and education/health categories are increased, in order to facilitate the purchase of a smart car. Based on the 3rd stage results in Table 18, when consumers have a higher education level, they tend to be more willing to buy automobiles

of a hybrid or electric type than of a diesel type.

The expected expenditure ranking and expected expenditure share of the representative consumer as a function of age level are analyzed in scenario 2, the results of which are shown in Figure 16.

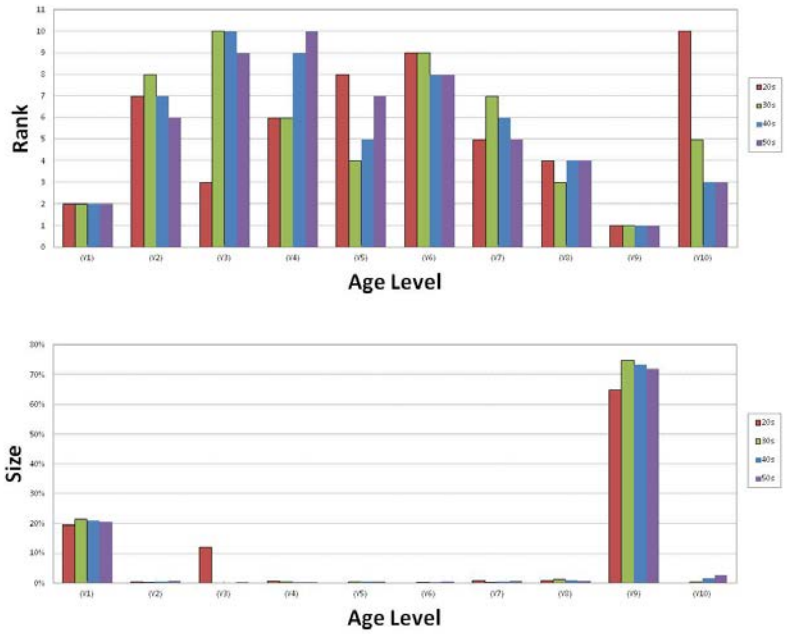

Figure 16. Changes to the Consumption-Expenditure Structure for the Representative Consumer, as a Function of Age Level, in Empirical Study 3

Figure 16 shows that the higher a consumer's age is, the more the ratio of the housing/electricity/others category is likely to be increased in order to purchase a smart car. Based on the 3rd stage results in Table 18, age level makes no difference on the choice of automobile type that is purchased.

The expected expenditure ranking and expected expenditure share of the representative consumer as a function of income level are analyzed in scenario 3, the results of which are shown in Figure 17.

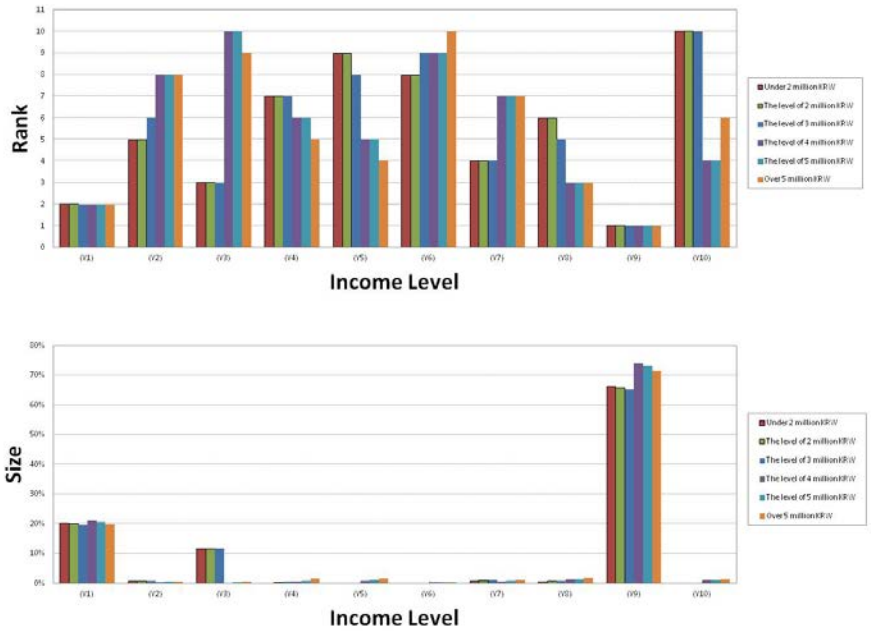

Figure 17. Changes to the Consumption-Expenditure Structure for the Representative Consumer, as a Function of Income Level, in Empirical Study 3

Figure 17 shows that the higher a consumer's income level is, the more the ratios of the alcoholic beverage/tobacco and recreation/culture categories are likely to be decreased, but the ratios of the transport category—which excludes the car purchases—and the education/health category will relatively increase in order to facilitate the purchase of a new smart car. Based on the 3rd stage results in Table 18, among higher-

income consumers, there is a greater willingness to buy an automobile of a hybrid or electric type than one of a diesel type.

The expected expenditure ranking and expected expenditure share of the representative consumer as a function of gender are analyzed in scenario 4, the results of which are shown in Figure 18.

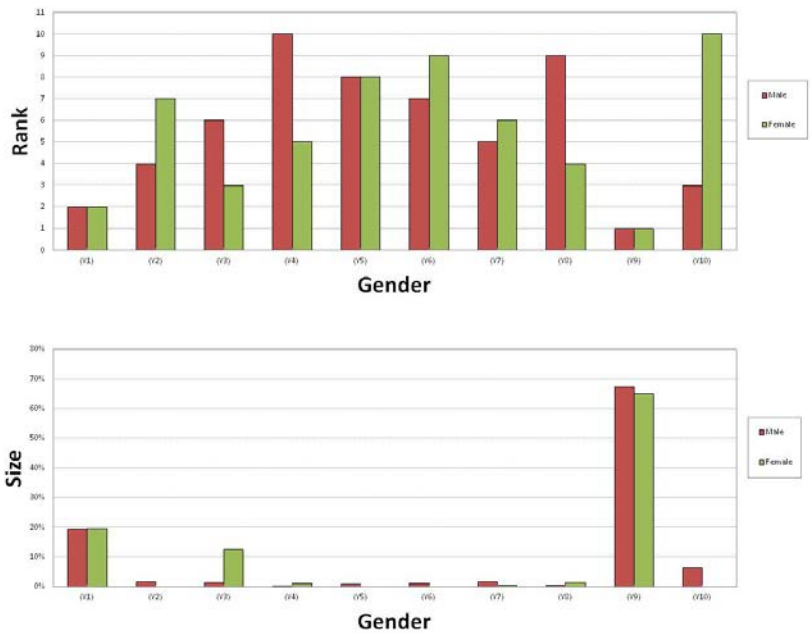

Figure 18. Changes to the Consumption-Expenditure Structure for the Representative Consumer, as a Function of Gender, in Empirical Study 3

Figure 18 shows that among male consumers, the ratios of the alcoholic beverage/tobacco, recreation/culture, and housing/electricity/other categories are higher than those among female consumers. However, among female consumers, the ratios of

the clothing/fashion accessories and education/health categories are higher than those among male consumers. Based on the 3rd stage results in Table 18, male consumers have a greater willingness than female consumers to buy automobiles of a hybrid type.

Because analysis of the consumption-expenditure pattern provides a proper standard by which one can understand the tendencies of a target group as policy-makers establish welfare programs or efficient policy, the estimation results outlined in this dissertation could be used to establish efficient policies that encourage the adoption of smart cars and electric automobiles. In addition, because consumers' consumption-expenditure patterns can be predicted through estimation results and socio-demographic information, the results in this dissertation could be used to develop and undertake market segmentation strategies. In other words, if a company's product is part of the transport category—including new car purchases—estimation results vis-à-vis consumption-expenditure pattern as a function of household or agent characteristics could inform efficient marketing strategies that increase product sales.

Chapter 5. Summary and Conclusion

According to Du and Kamakura (2008), marketers and public-policy makers need to understand how consumers spend their budgets in particular categories, how consumer spending is affected by other categories, and which categories have substitution or complementary relationships among cross-categories, because consumer expenditures—which involve seemingly unrelated categories—are clearly correlated in terms of budget constraints. With respect to the analysis of such matters, Du and Kamakura (2008) mention that it is important to suggest a demand model that considers consumer budget allocations and multiple categories. However, in the marketing literature, most consumer demand models focus on within-category purchase decisions, which means they consider only product choices within a single product category. Recently, a few studies have worked to develop demand models in order to analyze multiple categories; nonetheless, none consider the budget allocation stage in their model.

Meanwhile, several empirical studies within the marketing literature analyze how consumers allocate and consume budgets. In the economic literature, some studies consider budget allocation with respect to several consumption categories, but they relate only to broad commodity groups (Deaton and Muellbauer, 1980; Pollak and Wales, 1978) or analyze only a particular category of goods or services. Economically, outside goods (i.e., those from other categories) should be included in the choice model to analyze the feasibility of price policies (Allenby et al., 2004b), consumer preference, and the

relationships among cross-categories. Moreover, because the consumer decision-making process has more than one stage—except in situations involving simple choices—choice models that consider multiple stages of the consumer decision-making process need to be developed.

Therefore, this dissertation suggests new choice models that consider the budget-allocation stage, to include outside goods and the consumer decision-making process from multi-stage and multi-category perspectives. The choice models proposed in this dissertation are classified into two parts: a multi-stage and multi-category discrete choice model that consists of budget allocation and product choice stages, and a multi-stage and multi-category discrete-continuous choice model that additionally considers the product usage stage. To examine the validity of the proposed models, a simulation study is conducted based on an identification process that consists of preliminary and post-control methods. That simulation study shows that the models proposed in this dissertation are statistically valid.

Based on the two kinds of proposed models, this dissertation conducts empirical studies of representative products from the ICT (see section 4.1), household product (see section 4.2), and automobile industries (see section 4.3); three different industries are chosen, because the impact of budget allocations for one consumption-expenditure category on product choice is expected to differ from that for another category. From the estimation results derived from the proposed models, four matters can be analyzed. First, the difference in consumption-expenditure structure for purchasing products, as a

function of household and agent characteristics, is analyzed. Second, the relationships among consumption-expenditure categories can be identified from the estimation results in the 2nd stage and via the variance–covariance matrix. Third, the 3rd stage estimation results identify those consumption-expenditure categories in which changes can significantly affect purchase behavior. Finally, the 3rd stage estimation results also show consumer preference vis-à-vis product attributes. Through the three empirical studies outlined here, the utilization of proposed models and the implications thereof are examined.

Meanwhile, to examine trends with regard to each consumption-expenditure category and the purchase of products, the expected expenditure share of the representative consumer is analyzed, based on the estimation results. In addition, to compare the impact of socio-demographic level on consumption-expenditure structure changes, four scenario analyses are examined by using four socio-demographic variables: education level, age level, income level, and gender. Through analysis of the expected expenditure share of the representative consumer, consumers who are willing to buy smart pads are found to be more likely to reduce their unrelated consumption-expenditure category budgets. In other words, their budgets related to the communication, recreation/culture, education/health, and transport categories are changed somewhat, but those for unrelated expenditure categories—such as food, clothing/fashion accessories, and saving/insurance—are reduced greatly and together account for 97% of the total consumption-expenditure change. For eco-friendly detergent, consumers are more likely

to reduce their food category budgets, and this reduction accounts for about 80% of the total consumption-expenditure change. Given these results, I expect that eco-friendly detergent has a substitution relationship with the food category. For smart cars, consumers are more likely to focus on reducing saving/insurance category budgets, simply because automobiles are expensive.

According to Frank et al. (1972) and Fine (1980), the analysis of consumption-expenditure patterns provides a proper standard by which one can better understand the tendencies of a target group, and this is especially useful when policy-makers wish to establish welfare programs or efficient policy. In addition, analysis of budget allocation patterns points to the economic rationale behind consumer consumption expenditures.

On the analysis of consumer consumption-expenditure patterns, the literature shows that differences in consumer consumption-expenditure patterns can be analyzed in terms of consumer socio-demographic information (Fan and Lewis, 1999; Paulin, 2008; Lee, 2001; Maitra and Ranja, 2006; Du and Kamakura, 2008; Yusof and Duasa, 2010); it also mentions that, based on the estimation results vis-à-vis consumer consumption-expenditure patterns, substitution/complementary relationships among different categories can be useful in establishing efficient government policies and corporate marketing strategies (Allen and Rigby, 2005; Du and Kamakura, 2008; Yusof and Duasa, 2010).

Thus, based on the models proposed in this dissertation, it is possible to analyze consumer purchasing behavior more accurately by comprehensively analyzing the impact

of changes to allocated budgets between inside and outside goods on inside good product choice and usage. Additionally, the estimation results obtained in this dissertation could be utilized to establish efficient policies that encourage the adoption of particular products and to establish marketing strategies. From a methodological perspective, the proposed models are also expected to bear methodological implications with regard to analysis of the consumer decision-making structure.

Substantively, several directions can be taken to improve in future research the models proposed here. First, the proposed models could be extended to include the intertemporal allocation process and the feedback process, both from the equilibrium perspective. Generally, due to the high-dimensionality problem that comes with the existence of various products in each category, it is difficult for demand models to consider the structure of the equilibrium model; in other words, not all products in each category are considered together within a single model. Therefore, it is necessary to construct demand models from an equilibrium perspective, in order to reflect a more realistic situation pertaining to consumer purchasing behavior.

From this viewpoint, the proposed models should consider the intertemporal allocation process. Currently, the proposed models consider only static decision-making that bears a hierarchical decision structure. In a more realistic situation, consumers decide whether to consume more of a good in the current month, or save and consume next month, and the like. Thus, future study should consider dynamic decision-making that features an intertemporal allocation process.

In addition to the intertemporal allocation process, the proposed models should also consider including a feedback process. Because the decision-making results could affect consumer budget allocations as part of a feedback process, a hierarchical decision structure should be extended in future study. To reflect this feedback process, structural equation modeling could be used; given the advantage of a structural equation model—i.e., it considers both factor analysis and regression analysis (Breckler, 1990)—it can be and is utilized in a variety of fields. However, according to Kupek (2006), the structural equation model incurs some difficulties when used with categorical outcomes. Thus, whenever the framework of a structural equation model is used to reflect the feedback process, efforts should be made in future research to mitigate the difficulties that come with categorical outcomes.

Second, this dissertation uses SP data in its empirical examinations. Although the estimation results of consumer purchasing behavior based on SP data are useful in establishing policy and strategy-making directions, some gaps exist between results from hypothetical situations (SP data) and those from real-life situations (RP data). According to Brownstone et al. (2000) and Train (2003), the results of marginal willingness to pay from SP data could be more useful than those derived from RP data. However, few previous studies have analyzed consumer purchasing behavior while considering the budget-allocation stage; thus, if both SP and RP data are available, a combination of the two should be used, and in future research, these datasets should be applied to the proposed models in order to reflect more realistic situations.

Bibliography

Ackerberg, D.A. (2003). Advertising, learning, and consumer choice in experience good markets: An empirical examination. *International Economic Review*, 44 (3), 1007-1040.

Ahn, J., Jeong, G. & Kim, Y. (2008). A forecast of household ownership and use of alternative fuel vehicles: A multiple discrete-continuous choice approach. *Energy Economics*, 30, 2091-2104.

Allenby, G.M. & Rossi, P.E. (1999). Marketing models of consumer heterogeneity. *Journal of Econometrics*, 89, 57–78.

Allenby, G.M., Bakken, D.G. & Rossi, P.E. (2004a). The HB revolution. *Marketing Research*, 16, 20-25.

Allenby, G.M., Shively, T.S., Yang, S. & Garratt, M.J. (2004b). A Choice model for packaged goods: Dealing with discrete quantities and quantity discounts. *Marketing Science*, 23 (1), 95-108.

Allenby, G., Fennell, G., Huber, J., Eagle, T., Gilbride, T., Horsky, D., Kim, J., Lenk, P., Johnson, R., Ofek, E., Orme, B., Otter, T. & Walker, J. (2005). Adjusting choice models to better predict market behavior. *Marketing Letters*, 16, 197-208.

Allen, J. & Rigby, D. (2005). The consumer of 2020. reprinted from Gobal Agenda, published for the World Economic Forum Annual Meetings in Davos, Swizerland, 26-30.

Amemiya, T. (1978). The estimation of a simultaneous equation generalized probit model. *Econometrica*, 46, 1193-1205.

Anders, S. & Moser, A. (2010). Consumer choice and health: The importance of health attributes for retail meat demand in Canada. *Canadian Journal of Agricultural Economics*, 58 (2), 249-271.

Anderson, D.R., Sweeney, D.J. & Williams, T.A. (2009). *Essentials of statistics for business and economics*. Mason, OH: Thomson Higher Education.

Arbuthnot, J. (1977). The roles of attitudinal and personality variables in the prediction of environmental behavior and knowledge. *Environment and Behavior*, 9 (2), 217-232.

Baltas, G. (2004). A model for multiple brand choice. *European Journal of Operational Research*, 154 (1), 144-149.

Banfi, S., Farsi, M., Filippini, M. & Jakob, M. (2008). Willingness to pay for energy-saving measures in residential buildings. *Energy Economics*, 30, 503-516.

Barnard, J., McCulloch, R. & Meng, X. (2000). Modeling covariance matrices in terms of standard deviations and correlations, with application to shrink-age. *Statistica Sinica*, 10, 1281-1311.

Baslevent, C. & El-hamidi, F. (2009). Preference for early retirement among older government employees in Egypt. *Economics Bulletin*, 29 (2), 554-565.

Benitez-Silva, H., Buchinsky, M., Chan, H., Cheidvasser, S. & Rust, J. (2004). How large is the bias in self-reported disability?. *Journal of Applied Econometrics*, 19,

649-670.

Besanko, D., Gupta, S. & Jain, D. (1998). Logit demand estimation under competitive pricing behavior: An equilibrium framework. *Management Science*, 44 (11), 1533-1547.

Berry, S. (1994). Estimating discrete-choice models of product differentiation. *The RAND Journal of Economics*, 25 (2), 242-262.

Berry, S., Levinsohn, J. & Pakes, A. (1995). Automobile prices in market equilibrium. *Econometrica*, 63 (4), 841-890

Berry, S. & Haile, P. (2009). Identification of discrete choice demand from market-level data. *Working Paper*, Yale University

Bettman, J.R. (1979). *An Information Processing Theory of Consumer Choice.* Reading, MA: Addison-Wesley Publishing Co.

Bettman, J.R. & Park, C.W. (1980). Effect of prior knowledge and experience and phase of the choice process on consumer decision processes: a protocol analysis. *Journal of Consumer Research*, 7, 234-248.

Bhat, C.R. & Castelar, S. (2002). A unified mixed logit framework for modeling revealed and stated preferences: formulation and application to congestion pricing analysis in the San Francisco Bay area. *Transportation Research: Part B*, 36, 593–616.

Bhat, C.R. (2005). A multiple discrete-continuous extreme value model: formulation and application to discretionary time use decisions. *Transportation Research: Part B*, 39, 679–707.

Bhat, C.R. & Sen, S. (2006). Household vehicle type holdings and usage: an application of the multiple discrete-continuous extreme value (MDCEV) model. *Transportation Research: Part B*, 40, 35–53.

Bhat, C.R (2008). The multiple discrete-continuous extreme value (MDCEV) model: Role of utility function parameters, identification considerations, and model extensions. *Transportation Research: Part B*, 42, 274-303.

Blundell, R.W. & Smith, R.J. (1989). Estimation in a class of simultaneous equation limited dependent variable models. *The Review of Economic Studies*, 56 (1), 37-57.

Blundell, R., Pashardes, P. and Weber, G. (1993). What do we learn about consumer demand patterns from micro data. *The American Economic Review*, 83 (3), 570-597.

Boztug, Y. & Hildebrandt, L. (2006). *A market basket analysis conducted with a multivariate logit model*. Heidelberg: Springer Berlin.

Breckler, S.J. (1990). Applications of covariance structure modeling in psychology: Cause of concern?. *Psychological Bulletin*, 107 (2), 260-273.

Breiman, L. Freidman, J.H., Olshen, R.A. & Stone, C.W. (1984). *Classification and Regression Trees*. Belmont, CA.: Wadsworth International.

Brownstone, D. & Train, K. (1999). Forecasting new product penetration with flexible substitution patterns. *Journal of Econometrics*, 89, 109-129.

Brownstone, D., Bunch, D.S. & Train, K. (2000). Joint mixed logit models of stated and

revealed preferences for alternative fuel vehicles. *Transportation Research: Part B*, 34, 315–338.

Buchmueller, T.C., Grumbach, K., Kronick, R. & Kahn, J.G. (2005). Book review: The effect of health insurance on medical care utilization and implications for insurance expansion: A review of the literature. *Medical Care Research and Review*, 62, 3-30.

Calfee, J., Winston, C. & Stempski, R. (2001). Econometric issues in estimating consumer preference from stated preference data: a case study of the value of automobile travel time. *Review of Economics and Statistics*, 83 (4), 699-707.

California Energy Commission, (2007). *State Alternative Fuels Plan*. Arnold Schwarzenegger, Governor.

Caris and Company, (2010). iPads and tablets: are they cars or motorcycles?. [online]. Available from: http://www.cariscompany.com

Carpentier, A. & Guyomard, H. (2001). Unconditional elasticities in two-stage demand systems: an approximate solution. *American Journal of Agricultural Economics*, 83 (1), 222-229.

Chang, H.S., Griffiths, G. & Bettington, N. (2002). The demand for wine in Australia using a systems approach: industry implications. *Agribusiness Review*, 10, 1-12.

Chib, S. & Greenberg, E. (1998). Analysis of multivariate probit models. *Biometrika*, 85 (2), 347-361.

Chintagunta, P.K. (1992). Estimating a multinomial probit model of brand choice using

the method of simulated moments. *Marketing Science*, 11, 386-407.

Chintagunta, P.K. & Nair, H.S. (2010). Marketing models of consumer demand. *Research Papers 2072*, Standford University, Graduate School of Business.

Chib, S. & Greenberg, E. (1998). Analysis of multivariate probit models. *Biometrika*, 85 (2), 347–361.

Choi, J.Y., Shin, J. & Lee, J. (forthcoming). Strategic management of new products: ex-ante simulation and market segmentation. *International Journal of Market Research*.

Cotterill, R.W. & Putsis, W.P. (2000). Market share and price setting behavior for private labels and national brands. *Review of Industrial Organization*, 17 (1), 17-39.

Currim, I.S. (1982). Predictive testing of consumer choice models not subject to independence of irrelevant alternatives. *Journal of Marketing Research*, 19, 208-222.

Currim, I.S., Meyer, R.J. & Le, N.T. (1988). Disaggregate tree-structured modeling of consumer choice data. *Journal of Marketing Research*, 25 (3), 253-265.

Dangelico, R.M. & Pontrandolfo, P. (2010). From green product definitions and classifications to the green option matrix. *Journal of Cleaner Production*, 18, 1608-1628.

Dasgupta, C.G., Dispensa, G.S. & Ghose, S. (1994). Comparing the predictive performance of a neural network model with some traditional market response model. *International Journal of Forecasting*, 10, 235-244.

Deadman, D. & MacDonald, Z. (2004). Offenders as victims of crime? An investigation into the relationship between criminal behavior and victimization. *Journal of the Royal Statistical Society Series A*, 167, 53-67.

Deaton, A. & Muellbauer, J. (1980). An almost ideal demand system. *The American Economic Review*, 70 (3), 312-326.

Dubin, J.A. (1986). A nested logit model of space and water heat system choice. *Marketing Science*, 5, 112-124.

Duffy, M. (2003). Advertising and food, drink and tobacco consumption in the United Kingdom: a dynamic demand system. *Agricultural Economics*, 28 (1), 51-70.

Du, R.Y. & Kamakura, W.A. (2008). Where did all that money go? Understanding how consumers allocate their consumption budget. *Journal of Marketing*, 72, 109-131.

Eastwood, G. (2006). *The future of convergence – new devices services and growth opportunities*. Business insights.

Eaton, M.L. (1983). *Multivariate statistics: A vector space approach*. Wiley: New York

Edgerton, D.L. (1997). Weak Separability and the estimation of elasticities in multistage demand systems. *American Journal of Agricultural Economics*, 79 (1), 62-79.

Edwards, Y.D. & Allenby, G.M. (2003). Multivariate analysis of multiple response data. *Journal of Marketing Research*, 40 (3), 321–334.

Engel, J.F., Kollat D.T. & Blackwell R.D. (1968). *Consumer behaviour*. New York: Holt, Rinehart and Winston.

Engel, J.F., Blackwell R.D. & Miniard P.W. (1995). *Consumer behavior*. 8th ed. Fort

Worth: Dryden Press.

Ewing, G. & Sarigollu, E. (2000). Assessing Consumer Preferences for Clean-Fuel Vehicles: A Discrete Choice Experiment. *Journal of Public Policy & Marketing*, 19 (1), 106-118.

Fan, J.X. & Lewis, J.K. (1999). Budget allocation patterns of African Americans. *The Journal of Consumer Affairs*, 33 (1), 134-164.

Febbri, D., Monfaridi, C. & Radice, R. (2004). Testing exogeneity in the bivariate probit model: Monte carlo evidence and an application to health economics. *Working paper 514*, Department of Economics, University of Bologna, Italy.

Fine, S.H. (1980). Towards a theory of segmentation by objectives in social marketing. *Journal of Consumer Research*, 7 (1), 1-13.

Frank, R.E., Massy, W.F. & Wind, Y. (1972). *Market Segmentation*, Englewood Cliffs: Prentice-Hall

Garrido, R.A. & Mahmassani, H.S. (2000). Forecasting freight transportation demand with the space-time multinomial probit model. *Transportation Research Part B: Methodological*, 34, 403-418.

Gensch. D.H. (1987). A Two-stage disaggregate attribute choice model. *Marketing Science*, 6 (3), 223-239.

Goldberger, A.S. & Gamaletsos, T. (1970). A cross-country comparison of consumer expenditure patterns. *European Economic Review*, 1, 357-400.

Goldman, D.P., Bhattacharya, J., McCaffrey, D.F., Duan, N., Leibowitz, A.A., Joyce, G.F.

& Morton, S.C. (2001). Effect of insurance on mortality in and HIV-positive population in care. *Journal of the American Statistical Association*, 96, 883-894.

Goolsbee, A. & Petrin, A. (2004). The consumer gains from direct broadcast satellites and the competition with cable TV. *Econometrica*, 72 (2), 351-381.

Greene, W.H. (2008). *Econometric analysis (6th edition)*. Prentice Hall, New York.

Gruber, J. & Owings, M. (1996). Physician financial incentives and cesarean section. *The RAND Journal of Economics*, 27, 99-123.

Grunert-Beckmann, S.C., Gronhoj, A., Pieters, R. & van Dam, Y. (1997). The environmental commitment of consumer organizations in Denmark, the United Kingdom, the Netherlands, and Belgium. *Journal of Consumer Policy*, 20 (1), 45-67.

Gupta, S. & Ogden, D.T. (2009). To buy or not to buy? A social dilemma perspective on green buying. *Journal of Consumer Marketing*, 26 (6), 376-391.

Hanssen, O.J. (1999). Sustainable product systems – experiences based on case projects in sustainable product development. *Journal of Cleaner Production*, 7 (1), 27-41.

Hauser, J.R. & Wernerfelt, B. (1989). The Competitive Implications of Relevant-Set/Response Analysis. *Journal of Marketing Research*, 26, 391-405.

Hausman, J. & McFadden, D. (1984). Specification test for the multinomial logit model. *Econometrica*, 52 (5), 1219-1240.

Heckman, J. (1978). Dummy endogenous variables in a simultaneous equation system. *Econometrica*, 46, 931-959.

Hewett, P.C., Mensch, B.S., de A Ribeiro, M.C.S., Jones, H.E., Lippman, S.A., Montgomery, M.R. & van de Wijgert, J.H.H.M. (2008). Using sexually transmitted infectio biomarkers to validate reporting of sexual behavior within a randomized, experimental evaluation of interviewing methods. *American Journal of Epidemiology*, 168, 202-211.

Holly, A., Gardiol, L., Domenighetti, G. & Bisig, B. (1998). An econometric model of health care utilization and health insurance in Switzerland. *European Economic Review*, 42, 513-522.

Hu, M. Y., Zhang, G., Jiang, C.X. & Patuwo, B.E. (1999). A cross-validation analysis of neural network out-of-sample performance in exchange rate forecasting. *Decision Sciences*, 30 (1), 197-216.

Huber, J. & Train, K. (2001). On the Similarity of Classical and Bayesian Estimates of Individual Mean Partworths. *Marketing Letter*, 12 (3), 259–269.

Jaffe, L.J. & Senft, H. (1966). The roles of husbands and wives in purchasing decision. in L. Adler and I. Crespi, eds., *Attitude Research as Sea*. Chicago: American Markeing Association, pp. 95.110.

Jeong, G. (2008). *Development and Application of Choice Models with Simultaneity and Endogeneity*. Ph.D. dissertation, Seoul National University, Seoul.

Jepen, C. (2008). Multinomial probit estimates of college completion at 2-year and 4-year schools. *Economics Letters*, 98, 155-160.

Kardes, F.R., Gurumurthy, K., Murali, C. & Ronald, J. (1993). Brand Retrieval,

consideration set composition, consumer choice, and the pioneering advantage. *Journal of Consumer Research*, 20 (1), 62-75.

Kass, G.V. (1980). An exploratory technique for investigating large quantities of categorical data. *Applied Statistics*, 29 (2), 119-127.

Kawatkar, A.A. & Nichol, M.B. (2009). Estimation of causal effects of physical activity on obesity by a recursive bivariate probit model. *Value in Health*, 12, A131-A132.

Kellgren, C.A. & Wood, W. (1986). Access to attitude-relevant information in memory as a determinant of attitude-behavior consistency. *Journal of Experimental Psychology*, 22, 328-338.

Kim, J., Allenby, G.M. & Rossi, P. (2002). Modeling consumer demand for variety. *Marketing Science*, .21 (3), 229-250.

Kim, Y., Lee, J.D. & Heo, E. (2003). Bayesian estimation of multinomial probit models of work trip choice. *Transportation*, 30, 351-365.

Kim, W.J., Lee, J.D. & Kim, T.Y. (2005). Demand forecasting for multigenerational products combining discrete choice and dynamics of diffusion under technological trajectories. *Technological Forecasting & Social Change*, 72, 825-849.

Kim, Y., Park, Y., Lee, J.D. & Lee, J. (2006). Using stated-preference data to measure the inconvenience cost of spam among Korean E-mail users. *Applied Economics Letters*, 13, 795-800.

Kim, Y., Jeong, G., Ahn, J. & Lee, J.D. (2007). Consumer preferences for alternative fuel

vehicles in South Korea. *International Journal of Automotive Technology and Management*, 7, 327-342.

Kilbourne W.E. & Beckmann, S.C. (1998). Review and critical assessment of research on marketing and the environment. *Journal of Marketing Management*, 14 (6), 513-532.

Kohli, R., Devaraj, S. & Mahmood, M.A. (2004). Understanding determinants of online consumer satisfaction: A decision process perspective. *Journal of Management Information Systems*, 21 (1), 115-136.

Koop, G. (2003). *Bayesian Econometrics*. John Wiley and Sons.

Koop, G., Poiriet, D.J. & Tobias, J.L. (2007). *Bayesian Econometric Methods*. Cambridge University Press, New York.

Kupek, E. (2006). Beyond logistic regression: structural equations modeling for binary variables and its application to investigating unobserved confounders. *BMC Medical Research Methodology*, 6 (13), 1-10.

Latif, E. (2009). The impact of diabetes on employment in Canada. *Health Economics*, 18, 577-589.

Layton, D.F. (2000). Random coefficient models for stated preference surveys. *Journal of Environmental Economics and Management*, 40 (1), 21-36.

Lee, C.K. & Kwak, S.J. (2007). Valuing drinking water quality improvement using a Bayesian analysis of a multinomial probit model. *Applied Economics Letters*, 14, 255-259.

Lee, J., & Cho, S. (2009). Demand forecasting of diesel passenger car considering consumer preference and government regulation in South Korea. *Transportation Research: Part A*, 43, 420–429.

Lee, Y. (2001). Consumption patterns of elderly households: are they different between yourger and older elderly in Korea? How do they differ between Korea and the US?. *Consumer Interest Annual*, 47, 1-11.

Lehmann, D.R. & Pan, Y. (1991). The Impact of New Brand Entry on Consideration Set. *working paper*, Marketing Department, Columbia University, New York.

Lehmann, D.R. & Moore, W.L. (1986). Two approaches to estimating hierarchical models of choice. *working paper*, Graduate School of Business, Columbia University.

Li, K. (1998). Bayesian inference in a simultaneous equation model with limited dependent variables. *Journal of Econometrics*, 85, 387-400.

Lussier, D. & Olshavsky, R. (1979). Task complexity and contingent processing in brand choice. *Journal of Consumer Research*, 6, 154-165.

Maddala, G.S. (1983). *Limited Dependent and Qualitative Variables in Econometrics*. Cambridge University Press, Cambridge.

Maitra, P. & Ranjan, R. (2006). Household expenditure patterns and resource pooling: evidence of changes in post-apartheid South Africa. *Review of Economics of the Household*, 4, 325-347.

Manchanda, P., Ansari, A. & Gupta, S. (1999). The “Shopping Basket” : A model for multicategory purchase incidence decisions. *Marketing Science*, 18 (2), 95-114.

Marra, G. & Radice, R. (2011). Estimation of a semiparametric recursive bivariate probit model in the presence of endogeneity. *The Canadian Journal of Statistics*, 39 (2), 259-279.

Matsukawa, I. & Ito, N. (1998). Household ownership of electric room air conditioners. *Energy Economics*, 20, 375-387.

McCulloch, R. & Rossi, P.E. (1994). An exact likelihood analysis of the multinomial probit model. *Journal of Econometrics*, 64, 207-240.

McCulloch, R., Polson, N. & Rossi, P.E. (2000). A Bayesian analysis of the multinomial probit model with fully identified parameters. *Journal of Econometrics*, 99 (1), 173-193.

McFadden, D. (1974). *Conditional logit analysis of qualitative choice behavior*. in P. Zarembka (ed.), Frontiers of econometrics, New York: Academic Press, 150–142.

McFadden, D. (1986). The choice theory approach to marketing research. *Marketing Science*, 5, 275-297.

McFadden, D. (2000). *Disaggregate Behavioral Travel Demand's RUM side: A 30-Year Retrospective*. Prepared for a presentation at the International Association of Travel Behavior analysts.

Mueller, M.L., Park, Y., Lee, J. & Kim, T.Y. (2006). Digital identity: How users value the attributes of online identifiers. *Information Economics and Policy*, 18, 405-422.

Nedungadi, P. (1990). Recall and Consumer Consideration Sets: Influencing Choice without Altering Brand Evaluations. *Journal of Consumer Research*, 17, 263-276.

Nevo, A. (2001). Measuring market power in the ready-to-eat cereal industry. *Econometrica*, 69 (2), 307-342.

Nobile, A. (2000). Comment: Bayesian multinomial probit models with a normalization constraint. *Journal of Econometrics*, 99 (2), 335-345.

Olshavsky, R. (1979). Task complexity and contingent processing in decision making: a replication and extension. *Organizational Behavior and Human Performance*, 24, 300-316.

Ottman, J. (1992). Sometimes, consumers will pay more to go green. *Marketing News*, 26 (6), 16.

Payme, J.W. (1976). Task complexity and contingent processing in decision making: an information search and protocol analysis. *Organizational Behavior and Human Performance*, 16, 366-387.

Parks, R.W. (1969). Systems of demand equation: and empirical comparison of alternative functional forms. *Econometrica*, 37, 629-650.

Park, J.H., Seo, K.K. & Jang, D.S. (1999). Recycling cell formation using group technology for disposal products. In: Proceedings of the First International Symposium on Environmentally Conscious Design and Inverse Manufacturing. IEEE, 830-835.

Park, Y., Kim, M. & Lee, G. (2011). *Empirical study on usage behavior of smart devices*. Korea Information Society Development Institute. (In Korean)

Paulin, G. (2008). Expenditure patterns of young single adults: two recent generations

compared. *Monthly Labor Review*, 131 (19), 19-50.

Peattie, K. (1995). *Environmental Marketing Management: Meeting the Green Challege*. Pitman Publishing, London, UK.

Pollak, R.A. & Wales, T.J. (1969). Estimation of the linear expenditure system. *Econometrica: Journal of the Econometric Society*, 37, 611-628.

Pollak, R.A. & Wales, T.J. (1978). Estimation of complete demand systems form household budget data: The linear and quadratic expenditure systems. *American Economic Review*, 68 (3), 348-359.

Prendergast, G.P. & Thompson, E.R. (1998). Cynical segmentation of environmental markets: the product of marketers' dispositions or corporate dements?. *Journal of Euromarketing*, 6 (4), 17-34

Quinlan, J.R. (1986). Introduction of decision trees. *Machine Learning*, 1 (1), 81-106.

Rao, V.R. & Winter, F.W. (1978). An Application of the Multivariate Probit Model to Market Segmentation and Product Design. *Journal of Marketing Research*, 15 (3), 361-368.

Ratneshwar, S. & Shocker, A.D. (1991). Substitution in Use and the Role of Usage Context in Product Category Structures. *Journal of Marketing Research*, 28, 281-295.

Roberts, J. & Nedungadi, P. (1995). Studying consideration in the consumer decision process: Progress and challenges. *International Journal of Research in Marketing*, 12 (1), 3-7.

Roberts, J.H. & Lattin, J.M. (1991). Development and Testing of a Model of Consideration Set Composition. *Journal of Marketing Research*, 28, 429-440.

Rolle, J.D. (1997). Estimation of Swiss railway demand with computation of elasticities. *Transportation Research Part E: Logistics and Transportation Review*, 33 (2), 117-127.

Rombouts, J.P. (1998). A knowledge-based system for ranking DfE-options. In: Proceedings of the 1998 IEEE International Symposium on Electronics and the Environment. IEEE, 287-291.

Rose, C., Beiter, K. & Ishii, K. (1999). Determining end-of-life strategies as a part of product definition. In: Proceedings of the 1999 IEEE International Symposium on Electronics and the Environment. IEEE, 219-224.

Schlossberg, H. (1991). Americans passionate about the environment? Critic says that. *Marketing News*, 25, 8-10.

Seetharaman, P.B., Chib, S., Ainslie, A., Boatwright, P., Chan, T., Gupta, S., Mehta, N., Rao, V. & Strijnev, A. (2005). Models of multi-category choice behavior. *Marketing Letters*, 16 (3), 239-254.

Shin, J., Hong, J., Jeong, G. & Lee, J. (2012). Impact of electric vehicles on existing car usage: A mixed multiple discrete-continuous extreme value model approach. *Transportation Research Part D: Transport and Environment*, 17 (2), 138-144.

Shocker, A., Ben-Akiva, M., Boccaro, B. & Nedungadi, P. (1991). Consideration Set Influences on Consumer Decision-Making and Choice: Issues, Models and

Suggestions. *Marketing Letters*, 2 (3), 181-197.

Simonson, I. & Tversky, A. (1992). Choice in Context: Tradeoff Contrast and Extremeness Aversion. *Journal of Marketing Research*, 29, 281-295.

Sonquist, J.A., Baker, E.L. & Morgan, J.N. (1971). *Searching for structure (alias-AID-III): An approach to analysis fo substantial bodies of micor-data and documentation for a computer program (successor to the automatic interaction detector program)*. Institute for Social Research, University of Michigan.

Sudhir, K. (2001). Competitive pricing behavior in the auto market: A structural analysis. *Marketing Science*, 20 (1), 42-60.

Suehiro, S., Komiyama, R., Matsuo, Y., Nagatomo, Y. & Morita, Y. (2010). *Cost-effectiveness analysis of CO2 reduction in the automobile sector*. Institute of Energy Economics Japan (IEEJ)

Syriopoulos, T. (2002). Risk aversion and portfolio allocation to mutual fund classes. *International Review of Economics & Finance*, 11 (4), 427-447.

Teo, T.S.H. & Yeong, Y.D. (2003). Assessing the consumer decision process in the digital marketplace. *Omega*, 31 (5), 349-363.

Tiezzi, S. (2002). Environmental defensive expenditures and households behavior in Italy. *Applied Economics*, 34, 2053-2061.

Train, K. (2003). *Discrete choice method with simulation*. Cambridge University Press, Cambridge.

Train, K. & Winston, C. (2007). Vehicle choice behavior and the declining market share

of US automakers. *International Economics Review*, 48 (4), 1469-1496.

Tversky, A. & Sattah, S. (1979). Preference Trees. *Psychological Review*, 86, 542-573.

Verbeke, W. & Ward, R.W. (2001). A fresh meat almost ideal demand system incorporating negative TV press and advertising impact. *Agricultural Economics*, 25 (2-3), 359-374.

Webb, E. & Forster, J. (2008). Bayesian model determination for multivariate ordinal and binary data. *Computational Statistics & Data Analysis*, 52 (5), 2632-2649.

Wei, Y., Liu, B. & Liu, X. (2005). Entry modes of foreign direct investment in China: a multinomial logit approach. *Journal of Business Research*, 58, 1495-1505.

Wilde, J. (2000). Identification of multiple equation probit models with endogenous dummy regressors. *Economics Letters*, 69, 309-312.

Ye., X. Pendyala, R.M. & Gottardi, G. (2007). An exploration of the relationship between mode choice and complexity of trip chaining patterns. *Transportation Research Part B: Methodological*, 41, 96-113.

Yusof, S.A. & Duasa, J. (2010). Consumption patterns and income elasticities in Malaysia. *Malaysian Journal of Economic Studies*, 47 (2), 91-106.

Appendix A: Conditional Multivariate Normal Distribution

We assumed that X has an n-dimension and that all subsets of X have a normal distribution. Therefore, X has the normal distribution $N(\mu, \Sigma)$. If X is partitioned as follows:

$$\text{(A-1)} \quad \underset{(p\times 1)}{X} = \begin{bmatrix} \underset{(q\times 1)}{X_1} \\ \underset{((p-q)\times 1)}{X_2} \end{bmatrix}$$

Then, the mean of X, μ, and the covariance matrix of X, Σ, are defined as follows:

$$\text{(A-2)} \quad \underset{(p\times 1)}{\mu} = \begin{bmatrix} \underset{(q\times 1)}{\mu_1} \\ \underset{((p-q)\times 1)}{\mu_2} \end{bmatrix} \text{ and } \underset{(p\times p)}{\Sigma} = \begin{bmatrix} \underset{(q\times q)}{\Sigma_{11}} & \underset{(q\times (p-q))}{\Sigma_{12}} \\ \underset{((p-q)\times q)}{\Sigma_{21}} & \underset{((p-q)\times (p-q))}{\Sigma_{22}} \end{bmatrix}$$

According to Eaton (1983), the conditional distribution of X_2 on X_1 has a multivariate normal with a mean of $\bar{\mu}$ and a covariance matrix of $\bar{\Sigma}$. The mean $\bar{\mu}$ and covariance matrix $\bar{\Sigma}$ are defined as follows:

(A-3) $X_2 \left| X_1 \sim MTN(\overline{\mu}, \overline{\Sigma}) \right.$

$$where, \begin{cases} \overline{\mu} = \mu_{X_2} + \Sigma_{21}\Sigma_{11}^{-1}\left[X_1 - \mu_1\right] \\ \overline{\Sigma} = \Sigma_{22} - \Sigma_{21}\Sigma_{11}^{-1}\Sigma_{12} \end{cases}$$

Appendix B: The Results of the Simulation Study

The results of the simulation study for the case 1 extended model (MVP–multiple continuous–MVP) are shown in the following two tables.

Table B1. Estimation Results of the Parameters in the Simulation Study

(Case 1 Extended Model)

Model	Parameter		True value	Estimated Value	
				beta	sd
MVP (2)	Y1	θ_{10}/σ_{11}	1	1.1309	0.0811
		θ_{11}/σ_{11}	-1.5	-1.5449	0.1033
		θ_{12}/σ_{11}	1.5	1.5165	0.1029
	Y2	θ_{20}/σ_{22}	-1	-0.8766	0.0593
		θ_{21}/σ_{22}	1	0.8626	0.0645
		θ_{22}/σ_{22}	-1	-1.0581	0.0707
Multiple continuous (1)	Q1	α_{11}	1	1.0082	0.0384
		α_{12}	-1	-0.9612	0.0378
		δ_{11}	1	1.1077	0.0680
		δ_{12}	-1.5	-1.5684	0.1179
	Q2	α_{21}	1.5	1.5000	0.0312
		α_{22}	-1.5	-1.4207	0.0302
		δ_{21}	-1.5	-1.3466	0.0587
		δ_{22}	0.5	0.4106	0.0980
MVP (3)	H1	β_{11}/σ_{55}	$0.5/\sqrt{3}$ =0.2886	0.2722	0.0378
		β_{12}/σ_{55}	$-1/\sqrt{3}$ =-0.5773	-0.6099	0.0432
		δ'_{11}/σ_{55}	$0.3/\sqrt{3}$ =0.1732	0.3334	0.0896
		δ'_{12}/σ_{55}	$-0.9/\sqrt{3}$ =-0.5196	-0.5730	0.0830

H2	β_{21}	-0.9	-0.8469	0.0640
	β_{22}	1.5	1.5993	0.0924
	δ'_{21}	-0.6	-0.5793	0.1146
	δ'_{22}	1	1.1571	0.1132
RMSD of MVP-multiple continuous-MNP model = 0.0905				

Table B2. Estimation Results of the Variance–Covariance Matrix in the Simulation Study (Case 1 Extended Model)

Variance-Covariance matrix	True value	Estimated Value	
		beta	sd
σ^2_{33}	2	2.0852	0.0964
σ^2_{44}	1.5	1.4803	0.0688
ρ	0.5	0.5054	0.0241

The results of the simulation study for the case 2 base model (MVP–multiple continuous–MNP–single continuous) are shown in the following two tables.

Table B3. Estimation Results of the Parameters in the Simulation Study

(Case 2 Base Model)

Model	Parameter		True value	Estimated Value	
				beta	sd
MVP (3)	Y1	θ_{10}/σ_{11}	1	0.9888	0.0697
		θ_{11}/σ_{11}	-1.5	-1.5749	0.0924
		θ_{12}/σ_{11}	1.5	1.5260	0.0894
	Y2	θ_{20}/σ_{22}	-1	-0.9210	0.0577
		θ_{21}/σ_{22}	1	0.9677	0.0641
		θ_{22}/σ_{22}	-1	-0.9483	0.0614
Multiple continuous (1)	Q1	α_{11}	1	0.9920	0.0377
		α_{12}	-1	-1.0216	0.0352
		δ_{11}	1	0.9297	0.0643
		δ_{12}	-1.5	-1.5058	0.1086
	Q2	α_{21}	1.5	1.4769	0.0295
		α_{22}	-1.5	-1.4658	0.0302
		δ_{21}	-1.5	-1.5233	0.0535
		δ_{22}	1.3	1.2817	0.0923
MNP (4)	U1	β_{11}	0.5	0.5659	0.0736
		β_{12}	-1	-1.1307	0.1092
		δ'_{11}	0.5	0.4839	0.1299
		δ'_{12}	-0.9	-1.1041	0.2013
	U2	β_{21}	-0.9	-0.7494	0.0614
		β_{22}	1.5	1.4104	0.0890
		δ'_{21}	-0.6	-0.4911	0.1322
		δ'_{22}	1	1.1483	0.1112
Single continuous (2)	M1	λ_{11}	0.3	0.2863	0.0300
		λ_{12}	-0.5	-0.5378	0.0296
		δ''_{11}	-0.2	-0.2083	0.0774
		δ''_{12}	0.4	0.4110	0.0438

RMSD of MVP-multiple continuous-MNP-single continuous model = 0.0770

Table B4. Estimation Results of the Variance–Covariance Matrix in the Simulation Study (Case 2 Base Model)

Variance-Covariance matrix	True value	Estimated Value	
		beta	sd
σ_{33}^2	2	2.0856	0.0989
σ_{44}^2	1.5	1.4564	0.0677
σ_{55}^2	1.3	1.4152	0.2661
σ_{77}^2	1.6	1.4674	0.0666
ρ	0.5	0.4993	0.0246

The results of the simulation study for the case 2-1 extended model (MVP–multiple continuous–MVP–single continuous) are shown in the following two tables.

Table B5. Estimation Results of the Parameters in the Simulation Study

(Case 2-1 Extended Model)

Model	Parameter		True value	Estimated Value	
				beta	sd
MVP (3)	Y1	θ_{10}/σ_{11}	1	0.9912	0.0711
		θ_{11}/σ_{11}	-1.5	-1.5014	0.0933
		θ_{12}/σ_{11}	1.5	1.4578	0.0899
	Y2	θ_{20}/σ_{22}	-1	-1.0186	0.0604
		θ_{21}/σ_{22}	1	1.0021	0.0636
		θ_{22}/σ_{22}	-1	-0.9396	0.0601
Multiple continuous (1)	Q1	α_{11}	1	0.9838	0.0379
		α_{12}	-1	-0.9725	0.0379
		δ_{11}	1	0.9711	0.0647
		δ_{12}	-1.5	-1.5048	0.1153
	Q2	α_{21}	1.5	1.5129	0.0310
		α_{22}	-1.5	-1.5034	0.0318
		δ_{21}	-1.5	-1.5207	0.0508
		δ_{22}	1.3	1.2809	0.0869
MVP (4)	U1	β_{11}/σ_{55}	$0.5/\sqrt{1.3}$ =0.4385	0.4361	0.0432
		β_{12}/σ_{55}	$-1/\sqrt{1.3}$ =-0.8770	-0.8652	0.0561
		δ'_{11}/σ_{55}	$0.5/\sqrt{1.3}$ =0.4385	0.3129	0.0937
		δ'_{12}/σ_{55}	$-0.9/\sqrt{1.3}$ =-0.7893	-0.7861	0.0865
	U2	β_{21}	-0.9	-0.8998	0.0633
		β_{22}	1.5	1.4828	0.0809
		δ'_{21}	-0.6	-0.6005	0.1168
		δ'_{22}	1	1.0482	0.1015
Single continuous (2)	M1	λ_{11}	0.3	0.2580	0.0291
		λ_{12}	-0.5	-0.5328	0.0312
		δ''_{11}	-0.8	-0.8432	0.0696
		δ''_{12}	0.4	0.4223	0.0571

RMSD of MVP-multiple continuous-MVP-single continuous model = 0.0353

Table B6. Estimation Results of the Variance–Covariance Matrix in the Simulation Study (Case 2-1 Extended Model)

Variance-Covariance matrix	True value	Estimated Value	
		beta	sd
σ_{33}^2	2	2.0947	0.0998
σ_{44}^2	1.5	1.3751	0.0647
σ_{77}^2	1.6	1.4006	0.0719
ρ	0.5	0.4311	0.0267

Abstract (Korean)

경제 발전으로 인해 기업들은 시장에 무수히 많은 제품 및 서비스들을 시장에 출시하고, 제품 간의 경쟁을 통해 극소수의 제품 및 서비스만 살아남아 이윤을 내고 있다. 또한 지속적인 국가 발전을 위해 정부는 인프라 구축 및 기술개발 로드맵을 수립하고, 친환경 제품들의 확산을 지원하는 등 다양한 활동을 하고 있으나, 일부 정책만이 그 효과를 내고 있다. 이와 같이, 새로운 제품이나 정책에 대한 소비자들의 반응을 사전에 알 수 없는 시장의 불확실성으로 인해서 기업 및 정부는 효과적인 제품 계획 및 정책 제안하는데 한계가 있다. 따라서 시장 불확실성을 줄이기 위해 소비자 구매 행태 연구는 경영 전략 및 정책 방향 수립 관점에서 필수적이다.

다양한 제품 및 서비스에 대한 소비자 구매 행태 연구는 IT, 마케팅, 에너지 및 환경 분야 등 여러 분야에서 수행되어 왔다. 또한 소비자 구매 행태 연구를 위해 다양한 모형들이 개발되고 사용되어 왔고, 그 중 이산 선택 모형은 소비자 선호 분석에 유용한 방법론의 하나로써 단일 선택 모형, 복수 선택 모형, 소비자 이질성을 반영한 모형 등 다양하게 개발되어 왔다. 하지만,

기존의 대부분의 모형들은 단일 품목에만 관심을 가지고 분석하고, 다른 품목의 상품들에 대한 가격과 특징에 대해서는 무관심 하였다. 경제학적으로 다른 품목을 선택모형에 고려해야 현실적인 가격 정책 및 소비자 선호를 분석할 수 있다. 또한 소비자들은 단순 선택 상황을 제외하고는 1단계 이상의 의사결정 과정을 거치므로, 소비자 의사결정 단계를 고려한 선택 모형 개발이 필요하다.

이러한 점에서 소비자들의 의사결정 구조와 다른 품목을 고려한 새로운 선택 모형 개발이 필요하다. 이 때, 수요 분석 및 예측 모형에 다른 품목을 고려하는 것은 소비자 관점에서는 소득 배분 단계를 소비자들의 의사결정 단계에 고려하는 것과 유사하다. 따라서 본 논문에서는 다른 품목과 소비자들의 의사결정 단계를 고려하기 위해 소득 배분 단계를 선택 모형에 고려한 모형 개발을 목적으로 한다. 본 논문에서 제안된 모형은 관심 품목에 속하는 제품의 가격 및 속성 수준에 대한 소비자 선호 분석과 더불어 다른 품목의 예산 구조가 관심 품목에 속하는 제품의 선택에 미치는 영향 등 포괄적인 분석이 가능하다.

본 논문에서 제안된 선택 모형은 소득 배분 단계와 제품 선택 단계를

고려한 모형과 소득 배분 단계와 제품 선택 및 사용 단계를 고려한 모형으로 구분된다. 두 가지 모형을 이용하여 본 논문에서는 ICT, 가정용품, 자동차 분야의 대표 제품들에 대한 실증분석을 수행한다. 첫 번째 실증연구는 스마트 기기의 등장으로 인해 소비자들의 소비 패턴에 많은 영향을 끼치고 있는 점을 고려하여 스마트 기기 선택에 대한 수요 분석을 소득 배분 단계를 고려하여 분석한다. 두 번째 실증연구는 녹색 발전을 위해 널리 확산되어야 할 친환경 제품의 선택에 대한 수요 분석을 소득 배분 단계를 고려하여 분석한다. 세 번째 실증연구는 차세대 스마트 자동차 제품 선택 및 사용을 고려한 수요 분석을 소득 배분 단계를 고려하여 분석한다. 세 가지 실증연구들을 통하여 본 논문에서 제안된 다양한 모형들의 활용성과 도출 가능한 함의들을 살펴본다.

결과적으로 본 논문에서 제안된 모형을 바탕으로 관심 품목과 다른 품목들 간의 소득 배분 변화가 관심 품목에 속하는 재화 선택과 사용에 미치는 영향을 종합적으로 분석함으로써 좀 더 정교한 소비자 구매 행태 분석이 가능하게 된다. 이를 바탕으로 정확한 제품의 수요 예측과 제품의 가격 정책 등의 정책적 제안을 하기 위한 유용한 정보를 제공할 것으로 기대된다. 또한,

본 논문에서 제안된 모형은 소비자 의사결정 구조 분석에 대한 방법론적 함의를 제공할 수 있을 것으로 생각된다.

주요어 : 선택모형, 소비자 구매 행태, 소득 배분, 관심외 품목, 수요분석, 소비자 의사결정 단계
학 번 : 2007-20888