2004

Exploring online brand choice at the SKU level : the effects of internet-specific attributes

Yanan WANG

Terms of Use

EXPLORING ONLINE BRAND CHOICE AT THE SKU LEVEL:

THE EFFECTS OF INTERNET-SPECIFIC ATTRIBUTES

WANG YANAN

MPHIL

LINGNAN UNIVERSITY

2004

EXPLORING ONLINE BRAND CHOICE AT THE SKU LEVEL:

THE EFFECTS OF INTERNET-SPECIFIC ATTRIBUTES

by

WANG Yanan

A thesis
submitted in partial fulfillment
of the requirements for the Degree of
Master of Philosophy

Lingnan University

2004

ABSTRACT

Exploring Online Brand Choice at the SKU Level:
The Effects of Internet-Specific Attributes

by

WANG Yanan

Master of Philosophy

E-Commerce research shows that existing studies on online consumer choice behavior has focused on comparative studies of channel or store choice (online or offline), or online store choice (different e-tailers). Relatively less effort has been devoted to consumers' online brand choice behavior within a single e-tailer. The goal of this research is to model online brand choice, including generating loyalty variables, setting up base model, and exploring the effects of Internet-specific attributes, i.e., order delivery, webpage display and order confirmation, on online brand choice at the SKU level. Specifically, this research adopts the Multinomial Logit Model (MNL) as the estimation methods. To minimize the model bias, the refined smoothing constants for loyalty variables (brand loyalty, size loyalty, and SKU loyalty) are generated using the Nonlinear Estimation Algorithm (NEA). The findings suggest that SKU loyalty is a better predictor of online brand choice than brand loyalty and size loyalty. While webpage display has little effect on the brand choice, order delivery has positive effect on the choice. Online order confirmation turns out to be helpful in choice estimation. Moreover, online consumers are not sensitive to net price of the alternatives, but quite sensitive to price promotion. These results have meaningful implications for marketing promotions in the online environment and suggestions for future research.

Key words: Multinomial Logit model; Brand choice; Internet-specific attributes; Smoothing constant

I declare that this thesis《Exploring Online Brand Choice at the SKU Level: The Effects of Internet-Specific Attributes》is an original work based primarily on my own research, and I warrant that all citations of previous research, published or unpublished, have been duly acknowledged.

WANG Yanan
10th Sept. 2004

CONTENTS

Acknowledgements

I would like to express my heartfelt gratitude particularly to my chief supervisor, Dr. CUI Geng, who has given me professional opinions and close supervision throughout my study. I would also like to express my sincere gratitude to my co-supervisor, Professor CHAN Tsang-Sing, for his kind guidance and support in assisting the completion of my thesis.

My heartfelt gratitude also goes to Dr. LUI Hon-Kwong, Dr. ZHOU Lianxi at Lingnan University, together with Prof. ZHOU Nan in City University of Hong Kong, for their professional advice on my thesis. Also, I wish to extend my appreciation to Dr. YIM Chi Kin in Hong Kong University and Dr. Bruce G.S. Hardie in London Business School, for their assistance and advice on certain aspects of my thesis.

I am thankful also to Dr. ZHANG Yongqing, the CEO of Eguo.com, for his support on collecting the data and sharing with me his managerial experience. Also, great thanks to WANG Yang and Albert, for their help on the data processing and programming.

My special thanks also go to Dr. ZHANG Guichang and Miss LI Lin, for their kind assistance on mathematical aspect of my thesis, together with their sincere friendship to me. I would also like to thank Miss LU Jiafang, WANG Liyan, ZHANG Fang and other friends, for their great encouragement and sharing the happiness and sadness during my research. Also great thanks to Linda and Kimmy, the department secretaries, for their considerate help.

Last but not the least, I would like to give the deepest gratitude to my family, for their love, encouragement and patience. Without their support and understanding, I would never begin my postgraduate study again and finish this thesis.

1. Introduction

Existing research on online consumer choice behavior has focused on comparative studies of channel or store choice (online or offline) or online store choice (different e-tailers) (Tonegawa, 2002). Relatively less effort has been devoted to the online brand choice behavior of consumers within a single e-tailer.

There are several reasons for the lack of attention to this topic. First, due to the turbulence of the Internet economy, research on business to consumer (B2C) e-commerce has focused on how a retailer can survive in a different market environment. As a result, many comparative empirical studies have appeared during the Internet hype period. Second, brand choice has been a major area of marketing research. The theories and choice models are well developed, thus it is difficult to make significant improvement. Third, how to collect and analyze online scanner panel data is a great challenge. Online panel data, which are in a format different from the ordinary offline panel data, are rich and voluminous, yet noisy and poorly structured, hidden in the large amount of clickstream data or page views.

Existing research on online customer loyalty illustrates that due to high switching cost and consumer inertia, a registered consumer of an e-tailer will stay with a website offering satisfactory service, neglecting price dispersion and trivial service discrimination (Ancarani and Shankai, 2002; Janssen and Moraga, 2001; Ward, 1999). Recently, the development of B2C e-commerce has been more stable, in other words, more rational, after the market shakeout of 2000. The e-tailers that survived the Internet hype have accumulated a certain number of loyalists or routine buyers, who exhibit more rational purchase behavior, including forming consideration set and making brand choice (Srinivasan et al., 2002).

However, to date, we have little knowledge how Internet-specific variables affect online brand choice behavior. Brand choice is a well-developed topic in the literature of marketing models. Previous research on brand choice has been based on the data from the offline purchasing environment. Degeratu *et al* (2000) find that there are systematic differences in consumer choice behavior between online and regular (offline) stores. Certain attributes are unique to the online stores but not relevant in the offline environment, such as webpage display and order delivery. How do these Internet-specific attributes affect the brand choice of online consumers? How to model the brand choice behavior in online environment? Identifying and understanding the effects of these Internet-specific attributes are important for formulating marketing strategies for online marketers.

Internet-specific variables in this research refer to the variables that are unique or specific to online environment, for instance, due to data collection process (e.g., online order confirmation), or online choice attributes (e.g., delivery). Therefore, this research attempts to explore the effects of the Internet-specific attributes on online brand choice, using online panel data for a frequently purchased non-durable category (cola) from eguo.com, an e-tailer headquartered in Beijing, China.

Specifically, this research adopts the Multinomial Logit Model (MNL) as the estimation methods. To minimize the model bias, the refined smoothing constant for loyalty variables (brand loyalty, size loyalty and SKU loyalty) is generated. Loyalty variables with less bias are critical for building the base model in this research. Thus, the goal of this research is to develop a complete model of online brand choice, from generating loyalty variables, setting up base model to exploring the effects of Internet-specific attributes, including order delivery, webpage display and order confirmation, on online brand choice at the SKU level.

The results suggest that the smoothing constant of loyalty variables for the product (cola) is lower than the common smoothing constant found in the offline environment. Furthermore, SKU loyalty is a better predictor of online brand choice than brand loyalty and size loyalty. As for Internet-specific attributes, webpage display has little effect on brand choice, while faster order delivery has positive effect. Online order confirmation turns out to be helpful in choice estimation. Overall, incorporating the Internet-specific attributes contributes to more accurate models of online brand choice.

This research has meaningful managerial implications. It goes without doubt that market share is an aggregation of individual customer choices. If researchers can understand how and why online consumers choose one product over another, they can gain insight into the reasons for a product's success or failure in the online environment and develop more effective strategies to influence consumer choices

The remainder of this thesis is organized as follows. Chapter 2 presents the literature review on consumer choice models and previous research on online brand choice. Chapter 3 delineates the modeling approach for the MNL model. Chapter 4 describes the data and Chapter 5 the variables, including normal marketing mix variables and Internet-specific variables, and hypotheses. Chapter 6 presents the estimation results. Chapter 7 discusses the findings, implications, limitations and suggestions for further research.

2. Literature Review

2.1 Review of Brand Choice Models

A primary goal of marketing science is to describe, model, and finally predict the behavior of the consumers and their attitudes towards the products that form the market. (Matsatsinis and Samaras, 2000). Consumer behavior is one of the most important, dynamic research areas in market science (Matsatsinis and More, 2000).

The goal of the investigation of consumer behavior is to discover patterns of consumers' attitudes in their decision to buy or to ignore a product. Brand choice models have originated from research in the fields of marketing and management in the early 1950s. The main motive is the attempt to find answers to two intricate questions: (1) why changes occur in the market shares of the products, i.e. why do consumers shift from one brand to another? and (2) in what fashion are these changes taking place, i.e. how do consumers shift from one brand to another? The goal of brand choice models (or consumer personal choice models) is to model the purchasing behavior of consumers and more specifically to model the process which consumers follow when making decisions (Matsatsinis and More, 2000).

2.1.1 Brief introduction of choice models

Up to today, a number of consumer choice models has been developed, which can be used for the evaluation of important market data, such as market shares and a brand's purchase probability. These choice models largely fall into two categories: stochastic models of consumer choice and process-oriented models. Stochastic models of consumer behavior are often classified according to the type of behavior they attempt to describe and the models are being applied at such occasions that consumer's choice process is stochastic. Process-oriented models are adopted when consumer choice process is not stochastic but oriented towards one of the five stages of purchase cycle. Here is a summary of stochastic models and process-oriented models.

Table Ⅰ: A brief summary of stochastic models of brand choice

Model group	**Models**	**Brief description**
Zero-order models	Heterogeneous Bernoulli Model (Ehrenberg (1972))	A zero-order model applies when current and future purchasing behavior does not depend on past purchase history.
	Simple multi-brand model	
Markov models	Stationary first order Markov model	Markov models assume that only the last brand chosen affects the current purchase.
Learning models	Simple linear model (Kuehn (1962), Lilien (1974))	At the individual level, each purchase of a given brand enhances the likelihood of future purchases of the brand.
Variety seeking models	1st order (Jeuland (1979), Massy & Montgomery & Morrison (1970))	After buying a brand, the repeat-purchase probability will decrease.
	2nd order (Givon & Horsky (1978,1979), Kuehn (1962))	
Reinforcement/ Variety-seeking models	2nd models (Keon (1983))	One purchase of a brand increases its repurchase probability, but two consecutive purchases decrease that probability.
Combining brand choice and purchase timing	Multinomial / Dirichlet model (Jeuland, Bass&Wright (1980), Zufryden(1978), Dalal, Lee&Sabavala(1984))	Interpurchase times have an Erlang distribution of order, and the brand choice is a zero-order process.
Incorporating explanatory variables	Incorporating marketing variables in Stochastic brand choice model (Kuehn(1978), Givon&Horsky (1990))	The models attempt to incorporate explanatory variables in the framework of a linear learning model.

(Source: Lilien, Kotler & Moorthy, 1992 "Marketing Models" and Matsatsinis & More, 2002)

In many situations, particularly for low-involvement products where little conscious decision making takes place, a stochastic model---concentrating on the random nature of the choice process rather than on a deterministic explanation---may be more appropriate. Low involvement product means that with the basic purchase motivation, consumers involve little effort in choosing an adequate product to their satisfaction. A low-involvement product is normally frequently purchased, inexpensive, and the purchase decision requires little information and is less risky.

Stochastic models of brand choice can usually be distinguished by how they deal with (1) population heterogeneity, (2) purchase-event feedback, and (3) exogenous market factors. The mixing distribution is the most popular approach for moving from many models of individual behavior to one overall model for the population as a whole. It deals with population heterogeneity by having each individual in the population make a random choice from some distribution of responses. Although conceptually difficult, this approach has a distinct advantage: the identification and the measurement of specific discriminating characteristics of households need not be done explicitly, and most models are developed using this approach.

Stochastic models of brand choice also differ according to how they deal with purchase-event feedback, the influence of present purchase behavior on future purchase probabilities (Lilien, Kotler & Moorthy, 1992). Table I provides the details for stochastic models of brand choice (Table I).

In the dataset of this research, the product belongs to the frequently purchased consumer packaged goods. They are low-involvement items and have a large quantity of brand-switching data. Much of the focus of stochastic choice modeling on brand choice has centered on such products. In this research, modeling the brand choice with marketing variables is suitable to explore the Internet-specific attributes. The logit model, including its two important generalizations, the nested logit and the probit models, is often adopted as the modeling approach.

Stochastic models have the advantage of parsimony, which allows models of individual consumers to be aggregated and enables researchers to describe the characteristics of the population as a whole. As the complexity of the purchase situation increases, researchers may wish to include more phenomena into their models and allow consumers to vary across a wider range of characteristics (Lilien, Kotler & Moorthy, 1992). Thus, researchers have developed a series of models to describe the consumer behavior identified in the five stages of the framework (Table II).

The five stages from need arousal to the later purchase decision all have relationship with the brand choice. In this research, online brand choice focuses on the purchase stage of the purchase cycle. To set up the choice model that applies to the purchase stage, this research will consider the models mentioned above (Table II). It is illustrated that the multinomial logit model and the nested logit model are perhaps the most frequently used choice models in purchase stage models (Leeflang, et. al., 2000). In fact, logit and probit models are overlapping among stochastic models and process-oriented models when dealing with online brand choice for low-involvement product in this research.

Table Ⅱ: A brief summary of process-oriented models of consumer choice processes

Stages on choice process	Models	Brief description
Need arousal	Binary choice models of need arousal *(Hauser (1986)* Binary probit Binary logit Linear probability model	Need arousal corresponds to the category purchase decision. The models will examine which specific product or brand the consumer chooses. The discrete choice models at both stages are similar: the choice of when to buy and the choice of what to buy.
Information search	Models of brand awareness *(Blattberg & Jeuland (1981)* Models of consideration set formation *(Roberts & Lattin 1991)* Information integration (Hagerty & Aaker 1984)	The process of consumers gathering information about potentially suitable brands prior to evaluation is grouped into three steps: awareness, consideration and information integration.
Evaluation	Perceptual-evaluation (Moore & Winer 1987)	The models infer dimensions that discriminate between consumers' evaluations of different products based on brand interrelationship.
	Models of attitude and preference formation (Fishbein 1963)	The models take explicit attribute data and distill them into underlying dimensions or factors.
Purchase	Multinomial choice models (Luce's Model, Lesourne model (1977), Multinomial logit model, Low reinforcement model, Nested logit model) Markov Models	To forecast purchase by using preference measures, the models introduce an additional source of variation on the form of measurement error and changes in the consumer's preferences between the time of measurement and the time of purchase.
Postgraduate and purchase feedback	Models of variety seeking (Lattin & McAlister 1985)	The models the effect of current choice on future behavior by understanding the deterministic influences of choice.

(Source: Lilien, Kotler & Moorthy, 1992 "Marketing Models" and Matsatsinis & More, 2002)

2.1.2 Multinomial unordered discrete choice models

Brand choice modeling has grown to be a very substantial area in marketing research over the past decades. Since this research intends to explore how marketing variables affect the multinomial choice and the choice among different alternatives in this research is clearly unordered, multinomial models with discrete dependent variables are appropriate in exploring the effects of the Internet-specific variables.

Among multinomial unordered discrete choice models, there are several basic models: multinomial logit, multinomial probit, nested logit, and mixed logit (Greene, 2000; Lilien et al., 1992). Here, multinomial logit model denotes the conditional logit model. The conditional logit model is largely based on the work of McFadden (1973). This model is often (and confusingly) referred to as a Multinomial Logit model (Bowen and Wiersema, 2004). The key distinction between MNL and other models is that the variables used to explain the choices in the conditional logit model are the characteristics of the choices themselves rather than characteristics of the individual decision makers. In this research, the independent variables are the choice attributes. The exact name for the model is the conditional logit model. However, this study adopts the conventional name of Multinomial Logit Model as the majority of researchers have done.

(1) Multinomial Logit Model (MNL)

In a multinomial choice model, an individual chooses among several alternatives. The MNL model was developed by McFadden (1973) and has become perhaps the most frequently used choice model in marketing (as well as in other disciplines) (Leeflang *et al.*, 2000). The MNL model (in fact the conditional logit model) is defined by Warren Kuhfeld of the SAS Institute (2004) as follows:

"… the x-variables differ across alternatives but the estimated coefficients are the same

for each alternative. This is a common model in marketing for conjoint analysis. In a choice set the attributes of the alternatives are given as x-variables. The coefficients for these attributes -- the attribute "importance weights" -- are assumed to be the same for all alternatives (and respondents)."

An important assumption of the MNL model is that the odds of one choice versus another choice do not depend on the number of choice alternatives available. In other words, adding choices to the existing set of choices does not affect the odds between any two alternatives. This feature of the MNL model is derived from the formal equation for the odds in the model and is called the Independence of Irrelevant Alternatives (IIA)(McFadden, 1973). The practical advice often given is that when the alternatives are close substitutes, the IIA assumption may be violated and the MNL model may not give reasonable results (Bowen, Wiersema, 2004).

(2) Nested Logit Model

The Nested Logit model partially relaxes the IIA assumption by using a tree structure for the decisions that can be characterized as a set of branches and twigs (Greene, 2000). Each branch is a set of first level choices while each twig along a given branch represents a final choice. In a Nested Multinomial Logit (NMNL) model, consumer choice may follow a hierarchy of differentiating characteristics. The NMNL model has been used to model choices in product categories such as soft drinks, coffee and peanut butter (Leeflang, *at el*., 2000). In estimating the Nested Logit model, one can test the assumption of separating the decisions into branches and twigs or if the model can instead be collapsed into a standard MNL model of choice among all twigs (Greene, 2000).

(3) Multinomial Probit (MNP) Model

The Multinomial Probit (MNP) model primarily combines simulation with estimation. As a random utility model, the MNP model offers a highly desirable flexibility in substitution among alternatives that a MNL model fails to process. The unrestricted character of the variance matrix in the multivariate normal distribution that underlies the probit models cannot be produced by logit, even in its generalized extreme value forms (Ruud, 1996). That is, the disturbances of the random utility are assumed to follow a multivariate normal distribution. This distribution allows the utilities of alternatives to be correlated, so that the IIA assumption can be relaxed.

(4) Mixed Logit model

The Mixed Logit model augments the conditional logit model by including variables on decision maker characteristic (Bowen, Wiersema, 2004). The independent variables in Mixed Logit model include both choice attribute variables and choice maker characteristics. It should be noted that the Mixed Logit model also assume Independence of Irrelevant Alternatives (IIA) and therefore this assumption should be tested to assess model adequacy. Apparently, modeling both the choice attributes and the consumer characteristics are far more complicated and challenging. Since the data do not include variables of choice maker characteristics, the mixed model is excluded from the discussion.

2.2 Review of Former Research on Online Brand Choice

Up to date, only a few researchers of e-tailing have explored the choice behavior involving the purchase stage in a single B2C website. Two of them are empirical studies, one on choice behavior comparison between online and offline supermarkets (Degeratu, Rangaswamy and Wu, 2000) and the other on consideration set formation (Wu and Rangaswamy, 2003). Degeratu, Rangaswamy and Wu (2000) examine whether a consumer's choice behavior differs across online and offline transactions. They developed a brand choice model based on panel data from both online and offline environment. Their empirical results indicate that factual, non-sensory information affects online choice more strongly than sensory cues, and brand names become more important online in some categories than in others depending on the extent of information available to consumers -- brand names are more valuable when information on fewer attributes is available online. Their research also include personal list, an Internet-specific attribute, and explores its effect on price sensitivity. Their results indicate that online website's personal lists restrict consumers' consideration sets and may contribute to lower price sensitivity online than offline.

Wu and Rangaswamy (2003) propose a fuzzy set model of consideration set formation for choice making, calibrated on data from an online supermarket. They track consumers' decision process by analyzing clickstream data, and therefore, are able to characterize the consideration set formation mechanism. Their results suggest that consumers not only rely on their internal memory, but also sometimes engage in external information search to reduce the fuzziness of their consideration sets. They also find there is heterogeneity in consumers' capability to process external information. For some consumers, searching the external information (in the online store) dramatically increases the sizes of their consideration sets, but for others, external search does not impact the consideration set much at all. They suggest that online stores should be designed to provide easy-to-use mechanisms for personalization. Actually, their research explores consumer's choice behavior during evaluation stage. Their research also considers some Internet specific attributes, such as online shopping environment and personalized application.

Besides these two studies, Haubl and Trifts (2001) explore the effects of interactive decision aids on consumer decision-making in online shopping environments, by a controlled experiment using a simulated online store. The interactive decision aid is a unique characteristic of online shopping environments. These aids allow vendors to create retail interfaces that include highly interactive features. The authors develop a set of hypotheses pertaining to the effects of the interactive decision aids on various aspects of consumer decision-making. In particular, they focus on how the interactive decision aids affect consumers' search for product information, the size and quality of their consideration sets, and the quality of their purchase decisions in an online shopping environment. Their findings suggest that interactive decision aids designed to assist consumers in the initial screening of available alternatives and to facilitate in-depth comparisons among selected alternatives have strong favorable effects on both the quality and the efficiency of consumers' purchase decisions in online shopping environments — shoppers make much better decisions while expending substantially less effort.

2.3 Summary

Existing research has compared the online and offline shopping environments. Some of the studies also explored the effects of several Internet-specific attributes in the Internet purchasing environment and some specific decision tools. Overall, the existing studies have provided some preliminary evidence about the effect of Internet-specific variables on online brand choice behavior. Since online choice is not a separate process from offline behavior, it is plausible to propose that Internet-specific attributes can exert a significant effect on consumers' brand choice in the online shopping environment. Yet the effects of other Internet-specific attributes, especially those choice and product attributes, on the actual brand choices have not been explored and need to be examined with empirical data. Therefore, this research is largely exploratory in nature and aims at examining several basic issues associated with consumer online brand choice and focus on the effects of Internet-specific choice attributes. Knowing exactly how each Internet-specific variable affects brand choice, e-tailers can induce customers to choose the brands that have higher margin contributions to the e-tailer's revenue by facilitating or influencing the purchase process and consumer choices. Moreover, e-tailers can also adapt website design to facilitate the consumers' choice decision process. More importantly, e-tailers can be more efficient in inventory management.

This research will focus on a product category that belongs to the category of frequently purchased consumer packaged goods. Such products are often low-involvement items and have a large quantity of brand-switching data. Much of the focus of stochastic choice modeling on brand choice has centered on such products. Likewise, this research focuses on modeling brand choice using marketing decision variables that represent the effects of Internet-specific attributes. Thus, the MNL model is suitable for modeling online brand choice behavior. Chapter three will further explain why comparing with the nested logit and probit model, the MNL model is appropriate for modeling online brand choice.

3. Modeling Approach

An adequate marketing model should be chosen first before modeling online brand choice. Over the last twenty years, many marketing researchers working with household-level scanner data have used the Multinomial Logit model (MNL) (MacFadden, 1974) to study choice decisions involving multiple alternatives. The products in the scanner data are normally low-involvement items. This model has been very helpful for understanding and predicting brand choice behavior (Guadagni and Little, 1983), studying the effects of marketing mix and demographic variables on households' choice probabilities of brands, in particular the responsiveness to promotions (Papatla and Krishnamurthi, 1996), variety-seeking (Lattin, 1987), and advertising (Mela, Gupta and Lehmann, 1997).

The MNL model or conditional choice model is also adopted as the estimation method in this thesis, because of (1) its conceptual appeal being grounded in econometric theory (Jain, Vilcassim and Chintagunta, 1994), (2) its analytical tractability and ease of econometric estimation, namely, parsimony, and (3) its excellent empirical performance as measured by model fit and other criteria (Guadagni and Little, 1983).

A logit model uses the maximum likelihood method to estimate the parameters. The MNL model computes the probability of choosing an alternative as a function of the attributes of all the alternatives available. MNL model has the appeal of being stochastic and yet admitting decision variables (Guadagni and Little, 1983). These decision variables, such as price, promotional price discount, advertising or brand loyalty, constitute the deterministic part of the utility function that is used to compute the probability of choosing a specific product among several alternatives (Bentz and Merunka, 2000).

A major assumption of the model is that a consumer will choose the alternative that gives him/her maximal utility on each purchase occasion. Following the literature, it is assumed that online consumer derives a certain amount of utility from each brand and chooses the brand that provides him or her maximum utility. The utility of a brand is written as a function of marketing mix variables that normally include price, display, promotion, or the interaction terms of these marketing mix variables (Guadagni and Little, 1983). In addition, because brand utility is also a function of consumer loyalty, a loyalty variable is included in the model.

3.1 Maximum Utility

The basic idea is that the utility of a brand *i* is written as:

$$\mathbf{U} = \mathbf{Vi} + \varepsilon_i \qquad (1)$$

Where Vi is the deterministic component and ε_i the random. The consumer chooses the alternative for which utility is maximal. Thus the observed choice variable y_i is defined as:

$$\mathrm{y_i} = \begin{cases} 1\, if\ U_i > \mathrm{U_r}\ for\ all\ r \neq i, r = 1,\dots,n, \\ 0\ otherwise \end{cases} \qquad (2)$$

In the hypothetical case that Vi contains perfect information about the determinants of utility, a consumer would simply choose the product with the maximal Vi (Baltas,1998). The random component ε_i in (1) represents the inability of the modeler to include all the determinants of choice and introduces uncertainty regarding the outcome of consumer decision.

3.1.1 Linear utility

The deterministic component of a customer's utility for alternative *i* will be expressed as a linear function of observed variables, called the attributes of *i* in a specific choice occasion. In this thesis, they will be the attributes of a product in a specific choice occasion, including such exogenous variables as price, promotion and other marketing mix variables.

In addition, heterogeneity has traditionally been an important issue in brand choice models. Not accounting for heterogeneity when estimating logit models on panel data may lead to biased parameter estimates and more severely biased choice probability estimates (Jones and Landwehr, 1988). Thus a variable, which captures the effects of consumers' unobserved heterogeneity, should be included in the linear function. Loyalty variables are normally included in the online linear utility function to account for such effects. Adding brand loyalty or size loyalty has been well demonstrated in the former research (Guadagni and Little, 1983). Moreover, it goes without doubt that inadequate estimation of a nonlinear parameter can affect other model coefficients. So it is quite important to determine the best value of smoothing constant for the loyalty variables. In this research, smoothing constant is a major precondition for subsequent implementation of the choice models.

In general, the deterministic component is

$$V_i^k = \sum \beta_{ji} x_{ji}^k \tag{3}$$

where x_{ji}^k = observed value of attribute *j* of alternative *i* for customer *k*

β_{ji} = utility weight of attribute j of alternative i

To model consumer choice in the online environment, besides the alternative specific constants (ASCs) that are unique to the alternatives, the other attributes can be segmented into two classes.

(1) Tc={attributes common to online and offline environment}. These variables exist both in the online and offline store environment and have minor differences between these two environments, such as regular price, price promotion. These variables supposedly have the similar effect on brand choice decision-making but may have different weights.

(2) Ts={Internet-specific attributes}. These variables might have two categories. The first group of variables is similar to those offline, yet with the different definition and measurement due to the different marketing environment, such as webpage display. The other group consists of those variables that are comparably specific to online retailing environment, such as delivery and online order confirmation.

Hence, this thesis develops a choice model that is linear with a vector of observable brand or choice attributes including the Internet-specific attributes as well as unobservable individual/brand-specific effects.

3.1.2 Multinomial logit model

Logit models are the natural extensions of the regression models when the regressor is not a continuous variable but a state that may or may not be obtained, or a category in a given classification (Bentz and Merunka, 2000). MNL model is a conditional logit model, hence the model parameters are constant across all choices, and the explanatory variables are allowed to vary across both choice outcomes and individuals. This is different from a "standard" logit model, in which the parameters are allowed to vary across choices, but the explanatory variables can only vary across individuals (Jones and Landwehr, 1988). Here, the attributes of the conditional model allow researchers to compare the parameters and discern whether the variables are significant predictors of the dependent variables by the likelihood ratio test and the Chi-square test.

The probability that a customer *k* will choose brand *i* at occasion *t* is captured by the following equation:

$$P_{it}^{k} = \frac{exp(v_{it}^{k})}{\sum_{j} exp(v_{ji}^{t})} \tag{4}$$

where v_{it}^{k} is the utility determinant part mentioned above, and

i is an index for brands in the choice set.

3.1.3 Quality of Fit

Measures of quality of fit and parameter estimation guide model specification and help appraise the success of the calibration (Guadagni & Little, 1983). The fit measures for MNL models include the following.

(1) t-values for coefficients. T-values indicate whether the parameter coefficients of the variables are statistically significant.

(2) U^2 for model. Linear regression models offer residuals and R^2 as indicators of fit, while a logit model predicts only probabilities, which must then be compared to actual choices. For a logit model, U^2 is measured to evaluate the models in terms of

calibrated model and prior probabilities of choice. U^2 is somewhat analogous to R^2 in that they both have a range of 0 to 1 and indicate degree of variability explained. U^2 equals McFadden's (1974) likelihood ratio index, ρ^2. Therefore, U^2 equals to 1 minus $L(X)/L_0$, where $L(X)$ is the log likelihood of the calibrated model with explanatory variables, X, and L_0 is the log likelihood of the null model. If $L(X)$ of the calibrated model does not improve on L_0, then $L(X)=L_0$ and $U^2=0$. If the model is perfect, i.e., the predicted probabilities are all zeros and ones and correct, then the likelihood equals 1, $L(X)=0$ and $U^2=1$.

(3) Chi-square tests of model significance. If one model, say A, can be formulated as a restriction of the parameters of the tested model, say B, then L=2 log [likelihood ratio of model B to model A] is chi-square distributed with degrees of freedom equal to the difference in degrees of freedom between model B and model A. This test helps determine whether adding a parameter or set of parameters is worthwhile.

3.2 Detection of Independence of Irrelevant Alternatives

The MNL, however, imposes the restriction that the distribution of the random error terms is independent and identical over alternatives. This restriction leads to the Independence of Irrelevant Alternatives (IIA) property (Wen and Koppelman, 2000). IIA states that the odds of choosing one alternative over another is constant regardless of whichever other alternatives are present (Leeflang, et al., 2000). So some researchers adopt the nested MNL (NMNL) model as it relaxes the restrictive patterns of the inter-brand substitution imposed by the simple MNL (Baltas, 1998). Other researchers chose the probit model, the integrated Luce's choice model, or the McFadden's random utility function to make the error terms cross-sectionally correlated to avoid the independence of IIA (Allenby and Lenk, 1995).

For several reasons, this research adopts the MNL as the estimation method, neglecting the most commonly applied hierarchical model in marketing (Lilien, Kotler and Moorthy, 1992), the NMNL. First, since the online panel data tend to be noisy and parsimonious, low-complexity models should be preferred. Second, as the product category in the data is low-involvement non-durable commodity, the choice process is supposed to be simple. It is not necessary, as the nested MNL, to cluster alternatives into a hierarchical tree structure according to their similarities and the underlying consumer choice process. Third, though some models, such as NMNL and the MNP model, can deal with the problems that arise from the IIA assumption, including statistical tests of IIA, they often suffer from computational complexity. Thus, previous research on brand choice of low-involvement products has largely been based on the MNL model.

It is widely acknowledged that when using MNL model for discrete choice behavior, researchers should test for the IIA property. The violation of the IIA will cause systematic errors in predicted choice probabilities. The study will test IIA to determine if there is any violation of this assumption.

4. Data Description

Online panel data are needed to model the choice behavior via MNL model in this thesis. Panel data present histories of purchases for a sample of households or buyers. Traditionally, a cooperating household displays an identification card at the checkout point. The store clerk keys in the household number into the cash register, thereby causing the purchase record to be stored. Over time this creates a longitudinal customer history. While in the online environment, collecting purchase data seems quite easy due to the characteristics of the Internet and the database management.

As for online panel data, since application servers can keep track of a user's login mechanisms or cookies, it is easy to associate individual page views with a particular visitor (Kohavi *et al*., 2000). Also, the visitor's page views, purchase records and registration information can be extracted from the warehouse server. It is quite easy to retrieve the purchase data of an individual customer at the item level. Thus, online panel data can be mined from the website's server, through the noisy and miscellaneous data sets. Comparing with the traditional panel data, online panel data are easier and cheaper to collect. How to make the online data sets "speak" is the key task for the data analyst. Online panel data provide a good opportunity for research on brand choice and other consumer behavior.

The SKU-level online panel data in this research are from www.eguo.com, a well-known e-tailer headquartered in Beijing. The data consist of choice-specific attributes, without the individual-specific characteristics, in the purchase data of cola products. The sample consists of more than 2,000 panelists who bought cola products at least twice between Feb 2001 and Aug 2003. Among all the panelists, there is small percentage (only around 1%) of business buyers.

Initially, there are totally 17 SKUs in the dataset. MNL model will have a large number of parameters for modeling choice behavior for all these SKUs. Moreover, a model with nearly 20 SKUs is likely to violate the IIA property of the logit model. Thus, only the top seven SKUs (Coca-Cola 2L, Coca-Cola 355ml, Coca-Cola 355ml in full box, diet Coca Cola 355ml, Pepsi 2L, Pepsi 355ml, Pepsi 355ml in full box) are included because each of the rest of the SKUs has less than 1.5% share of the e-tailer's sales in the cola category. The SKUs included in the analysis account for 88.4% of the total number of purchases. Moreover, to simplify the modeling task, the study begins by assuming that each customer k has a fixed choice set (only the 7 SKUs), instead of the practically unbalanced choice set.

This research focuses on cola, a non-durable product commodity, as the product category for the choice model, because in online panel data, frequent purchases of the product can be tracked. Non-durable goods give more observations to calibrate the brand choice model. Meanwhile, cola has good market penetration and enough alternatives for comparing different choices. Moreover, cola is one of the best selling products in eguo.com and has more frequent promotions than other products.

For consumers of eguo.com, on some purchase occasions, an individual may purchase multiple units of SKUs. In the data used in the estimation, the study does not consider the quantity decision but only the SKU choice on each purchase occasion. Some researchers choose to randomly select one of the purchased brands when there are very few purchase occasions where multiple brands are bought. (e.g., Chintagunta, Kyriazidou and Perktold, 2001). However, in our online dataset, the multiple-choice is not the extreme minority (approximately 8% in the complete data set). In the case where multiple brands (say, N) were bought within one order in our online dataset, the common practice is to segment the single order as N purchases with zero intercept time according to their orders being chosen. Thus, no multiple purchase records are deleted randomly, keeping the dataset with full information about the consumers.

5. Variables and Hypotheses

5.1 Online Order Processing

Before model calibration, it is necessary to describe the online order processing and fulfillment in eguo.com. A registered customer first views the webpage for a specific category and may click on the aimed products that he/she intends to purchase. After he/she has chosen all the intended goods for this order, the customer submits the order online, together with the delivery address. When eguo.com receives the order, the staff in the customer service center calls the customer to confirm the order. If the customer confirms the order via the phone, the delivery staff in the nearest distribution center will provide free delivery service to the customer within designated hours. As for the payment method in eguo.com, the majority is cash by delivery, while some consumers purchase by bankcard. However, in China, credit card is still rare for online payment, due to payment safety and low credit card usage. In fact, order processing and delivery procedures at eguo.com are similar to those of other e-tailers in China. The payment method for e-tailers in China is quite different from e-tailers in more developed countries such as USA or European countries.

5.2 Modeling at the SKU Level

Most contemporary choice modelers use the brand of a product as the fundamental unit of analysis. However, Guadagni and Little's (1983) preliminary examination of switching behavior yields no evidence to suggest that customer choice is hierarchical on either brand or size. In fact, different sizes of the same brand are clearly different products from both retailer's and customer's points of view. Customers may show distinct size loyalty and retailers also promote sizes separately. Therefore, researchers working with panel data in fact model the brand-sizes. Overall, brand is obviously an important component of this decision, but brand choice is rarely a final decision by itself. Other components cannot be ignored in the choice process. Therefore, SKU choice is a more fitting description of the overall decision process and has been adopted in some brand choice models with panel data.

Specifically, the SKU information includes a product's brand name and other attributes of the unit. In addition, when a customer accesses an e-tailer's website and opens the category webpages, the brand becomes a product line comprising many SKUs. Consumers typically choose among SKUs on the basis of a set of product attributes, which tend to be discrete and tangible. Moreover, from the managerial aspect of e-tailers, modeling at the SKU level can generate more meaningful implications for promotion and inventory management. Thus, unlike some other choice models in the marketing literature, this thesis models the online choice at the level of stock-keeping units (SKUs) level.

5.3 Variables and the Measurements

This research is largely data driven, and the Internet-specific attributes are limited to the variables that are available in the online panel data. The followings are the definitions of the variables in the linear utility function and their measurements at the SKU level.

Price. Price for a SKU is the actual shelf price per unit, net of all discounts or deals offered by the e-tailer. Since cola has different product sizes for all the alternatives, each product price is converted to a per unit basis, based on the actual package size purchased. Thus, if the shelf price of a 355ml Coca-Cola is 2 RMB, and a consumer purchases 2 cans of this Coca-Cola, the net price is 0.56 RMB per 100ml. The net price enables the study of brand competition across all package sizes.

Previous research argues whether the price variables should be operationalized as shelf price, excluding coupons, or paid price, which is shelf price minus coupon value (Papatla and Srishnamurthi, 1996). The e-tailer in this research, eguo.com, does not offer e-coupons except for only a short period of time on a trial basis. So this thesis excludes the effect of coupons. It is assumed that the net price is what the consumers use to choose among alternatives. So in this thesis, price sensitivity is accordingly interpreted as net price sensitivity.

Sales Promotion. In general, there are two types of promotions, i.e. price promotion and non-price promotion. Price promotion focuses on price discount, which is operationalized as a dummy variable, assuming a value of 1 if an alternative has price discount and 0 otherwise. In eguo.com, the practice is to provide popular products with low prices. This strategy was used frequently. For instance, there was a price discount for Coke Cola products in 2001, selling 355ml Coca Cola for 1 RMB per can (50% discount of the original price of 2 RMBs per can).

For non-price promotion, the dataset has only one-year information on online advertisement including banner advertisement in the homepage, flashing hyperlink in the homepage, a product being listed on the “hot product” item, and floating banner in the homepage. Eguo.com has close cooperation with Coca-Cola, so eguo.com only made online non-price promotion for Coca-Cola products. Non-price promotion is not available for Pepsi products. Thus, as far as the online non-price promotion is concerned, this variable is only limited to several specific SKUs. Thus, including non-price promotions will lead to a selection bias.

Previous studies of brand choice have examines non-price promotion for the whole market as an independent variable (e.g., Guadagni and Little, 1983; Kalwani, 1990). Thus, the ideal non-price promotion will include online and offline promotion activities. This dataset is limited to promotion information for the two brands and their SKUs at a single e-tailer. Thus, the online non-price promotion is not suitable for inclusion into the utility function for the seven SKUs since it is too exogenous to the dependent variable. Therefore, this study focuses on price promotion only as part of the utility function.

Delivery. Eguo.com, like other e-tailers in China, promises two kinds of delivery time, one is short (within one hour) and the other is long (within 8 hours or even longer). Delivery is a dummy variable, with a value of 1 if an alternative's delivery time is short and 0 otherwise.

Loyalty Variables. A key issue in implementing the MNL model is proper control for preference heterogeneity across households (the individuals in our online panel data). In this dataset, each purchase will be treated as an observation so that the cross-sectional data and time-series data are combined. This makes the loyalty variables particularly important since they carry not only much of the cross-sectional heterogeneity but also a good part of the purchase-to-purchase dynamics.

It has been well-known for quite some time that incorrect assumptions regarding household-level heterogeneity can lead to biased estimates of the impact of marketing mix elements upon brand choice (Gonul and Srinivasan, 1993). This variable is operationalized using the entire purchase history of a customer either as the most recent purchase (Jones and Landwehr, 1988) or as an exponentially weighted sum of all previous choices made by the customer (Guadagni and Little, 1983), or as a proportion of purchases (Krishnamurthi and Raj 1991).

Guadagni and Little (1983)’s approach to measuring brand loyalty is one of the popoluar methods. An exponentially smoothed average of past purchases will be included in the model specification. This so-called loyalty variable is designed to correct for both cross-sectional heterogeneity and for non-stationarity in preferences over time. However, Feinberg and Russell (2003) denoted that, despite their active use, loyalty variables lack a rigorous theoretical basis, and thus little is known regarding the appropriate circumstances under which they might improve the estimates or introduce biases. For Guadagni and Little’s approach, loyalty variables are based upon observed choices. Many researchers have argued that loyalty variables are contaminated by marketing mix activity and so cannot be regarded as proper estimates of a household’s true brand preferences; for example, during a period of frequent promotional activity, loyalty variables will tend to favor the promoted brand.

Since the above endogeneity problem is beyond the scope of this study, Guadagni and Little’s approach to measuring the brand loyalty is still adopted. Size loyalty is analogous. Moreover, SKU loyalty will also be measured and included. Some researchers have noted that the consumers are loyal to the brand and size, but not the specific SKU (Fader and Lattin, 1992). As stated earlier, SKUs include both brand and size information. Thus, loyalty can be tracked at the SKU level thus provides richer information loyalty behavior.

Furthermore, eguo.com promotes the cola products at SKU levels, and delivery time also differs across SKUs. Thus, in addition to brand and size loyalty, loyalty at the SKU level is also calibrated and tested.

The brand and size loyalty variables are defined as follows:

$$x_k^i(n) = \alpha_b x_k^i(n-1) + (1-\alpha_b)\begin{cases} 1 & \text{if customer i bought brand (size) of alternative k at purchase occasion (n-1)} \\ 0 & \text{otherwise} \end{cases} \quad \ldots(5)$$

α_b is the carry-over constant, or smoothing constant for brand or size loyalty.

The SKU loyalty is defined as follows

$$x'^i_k(n) = \alpha_{b'} x_k^i(n-1) + (1-\alpha_{b'})\begin{cases} 1 & \text{if customer i bought SKU alternative k at purchase occasion (n-1)} \\ 0 & \text{otherwise} \end{cases} \quad (6)$$

$\alpha_{b'}$ is the carry-over constant, or smoothing constant for SKU loyalty.

It goes without doubt that inadequate estimation of a non-linear parameter can affect other model coefficients. So it is quite important to determine the best value of smoothing constant for the loyalty variables. For less-biased loyalty estimation, the values of exponential smoothing constants need to be determined first.

In this thesis, the Nonlinear Estimating Algorithm (NEA) by Fader and Lattin (1992) is adopted to obtain the maximum likelihood estimates for smoothing constants for brand loyalty, size loyalty and SKU loyalty (Appendix One). As for the initialization of brand loyalty, Guadagni and Little's set $x_k^i(1)$ to be α_b (the carry-over constant or smoothing constant) if the brand of alternative *k* was the first purchase in the data history of customer *i*, otherwise (1-α_b)/(number of brands-1), thus insuring that the sum of loyalties across brands always equals 1 for a customer. This research, however, adopts Fader and Hardie's approach to initialize brand loyalty, namely, $x_k^i(1)$ will be 1/number of alternatives, no matter whether alternative *K* was the first purchase in the data history of customer *i* or not. This approach is simpler in the value initialization for loyalty variables than Guadagni and Little's method.

[Insert Table One here]

The results for online brand loyalty and size loyalty suggest that the smoothing constant value on online environment is much lower than the normal values found in previous studies of the offline environment. The convenient value for smoothing constant is usually between 0.6~0.9 for research in the offline environment (Fader, Lattin and Little, 1992). For the smoothing constant in this study, the value for cola from online environment is between 0.4 and 0.5 (Table One). Furthermore, comparing with the smoothing constant of brand and size loyalty in offline environment, it shows the offline consumers have higher loyalty to brand or size than online consumers. By comparison, the online consumer exhibits less loyalty to brand or size.

As shown in Table One, the decay rates (namely, the value of smoothing constant) for brand and size loyalty are comparable. The brand coefficients decay slightly more slowly (higher α), suggesting somewhat higher loyalty to brand than size. This results support the finding of previous research (e.g., Guadagni and Little, 1983).

[Insert Table Two here]

To further explore such differences, the smoothing constant for ground coffee in the database from eguo.com is computed to see whether the lower smoothing constant for cola is due to the product category. This dataset includes data on ground coffee purchases for nearly two years. The smoothing constant for coffee brand loyalty is 0.66 (Table Two). This is lower than the smoothing constant for coffee found in Guadagni and Little's (1983) study. Thus, for the same product category, the smoothing constant in the loyalty functions is lower higher in the online environment than in the offline environment. However, the smoothing constant of 0.66 for ground coffee is much higher than cola in the online environment. The smoothing constant for coffee size loyalty is 0.37 (Table Two), also slightly higher than cola. Thus, coffee indeed has higher smoothing constant values for brand and size loyalty than cola in online environment. Thus, it may be the product category in the online environment in our dataset that lead to the differences in the values of the smoothing constants for brand loyalty. Overall, the decay rate for online brand and size loyalty is higher than offline environment. There might exist some factors that contribute to systematic differences in the smoothing constants of loyalty variable between the online and offline environments, and such factors need to be further explored in future research.

[Insert Table Three here]

Besides brand loyalty and size loyalty, the study also adopts the NEA to generate the smoothing constant for SKU loyalty. The value is 0.446 (Table Three), also smaller than the normal convenient value. In the data analysis stage, all three loyalty variables are included into the utility function of the MNL to explore which variable produces better results.

Lagged purchase variable. Loyalty variables measure the cross-sectional heterogeneity among consumers, while lagged purchase variable considers purchase-to-purchase heterogeneity. This variable is operationalized as a dummy variable, which assumes a value of 1 if a consumer purchased the brand on the previous purchase occasion, and 0 otherwise. This variable is an interaction with alternative specific constants. For discrete choice model in this research, most of the variables are attributes of the choices that are measured for the choice. But for lagged purchase variable, it is measured for the individual. Such a variable can only be incorporated in the discrete choice model by using the equivalent of dummy variable interaction terms. Otherwise, the variable is the same for all choices, and its coefficient cannot be estimated.

Order confirmation: This is another important Internet-specific variable. An individual might submit an online order first. Then when the staff in customer service center in eguo.com calls the individual for order confirmation, he/she might cancel the specific online order. This record is kept in the data warehouse. Order confirmation is also operationalized as a dummy variable, which assumes a value of 1 if the order is confirmed and 0 otherwise. Also, order confirmation is measured at individual level, but not at the choice level.

Display: Once a consumer opens the specific webpage for a certain product category, one may see some of the SKUs in the first page, or one would click "to the next page" to see the other SKUs. For certain product category in eguo.com, the maximum webpage to display all the alternatives is seven, and the minimum is one. As for the cola product category in this research, there are normally three webpages in total to show out all the alternatives. Webpage display is coded as a dummy variable, as1 if the SKU is on the first webpage for that category or 0 otherwise. It is expected that SKUs on the first webpage of a category are purchased more frequently.

5.4 Utility Function

Therefore, the variables of net price, price discount, delivery, loyalty variables, webpage display, lagged purchase and order confirmation together specify the following function for the utility of a brand to an online consumer:

$$U_i^{kt} = \beta_{0i} + \beta_p.Price_i^t + \beta_d.Discount_i^t + \beta_{dl}.Delivery_i^t + \beta_l.Loyalty_i^t \\ + \beta_{lg}.Lagged_i^t + \beta_{dp}.Display_i^t + \beta_{oc}.Order\ Confirmation_i^t + \varepsilon_i^t \quad (7)$$

where *i, k, t* index SKU alternative, online consumer, and purchase occasion, respectively, and ,

$Price_i^t$ = net price of SKU alternative *i* at purchase occasion *t*,

$Discount_i^t$ = price discount (1 if SKU alternative *i* is on a discount at purchase occasion *t*, 0 otherwise),

$Delivery_i^t$ = delivery for online order fulfillment (1 if SKU alternative *i* has short delivery time at purchase occasion *t*, 0 otherwise),

$Loyalty_i^t$ = loyalty variables,

$Lagged_i^t$ = lagged purchase (1 if a consumer *k* purchased the SKU alternative *i* at the previous purchase occasion *t-1* , and 0 otherwise),

$Display_i^t$ =webpage display (1 if SKU alternative *i* is on the first webpage for that category at purchase occasion *t*, 0 otherwise),

$Order\ Confirmation_i^t$ = choice order confirmation (1 if a consumer *k* made an order confirmation at purchase occasion *t*, and 0 otherwise),

β_{0i} = alternative specific constant for SKU alternative *i*,,

β_p = price sensitivity parameter,

β_d = response to a price promotion,

β_{dl} = effect of delivery,

β_l = effect of loyalty variables on utility,

β_{lg} = effect of lagged purchase,

β_{dp} = effect of webpage display,

β_{oc} = effect of order confirmation, and,

ε_i^t = random error in the utility of i .

It is an ideal utility function if all the variables are statistically significant to the utility of SKU alternative *i.* However, the base model in this research will include only loyalty variables and other marketing mix variables. Before setting the base model, loyalty variables should be chosen first. How to set a less biased base model is essential during the modeling process in this research. Other nested models exploring the effects of Internet-specific variables would be based on the base model. Beyond the base model, this research tests the effects of webpage display, delivery and online order confirmation as Internet-specific attributes.

5.5 Hypotheses about the Internet-specific Attributes

A major objective of this is to explore the effects of Internet-specific attributes on brand choice. Since the effects of other normal marketing mix variables have been elaborately discussed in the former research on brand choice, this research will not state any hypotheses about those variables. There are three Internet-specific attributes in this dataset, i.e., webpage display, delivery, and order confirmation, based on the utility function and managerial experience from the e-tailer. Their effects on consumers' brand choice in the online shopping environment are elaborated as follows.

This study explores whether webpage display positions of a product item, such as on the first page for that product category, affect the choice of SKUs. The effect of webpage display is quite similar to that of shelf display in the supermarket, in that better display positions such as close to the entrance and at the eye level contribute to greater product sales. Likewise, the alternative on the first webpage for the product category is more obvious to the online consumers. When the consumers open the category page, the alternative will be at his/her sight at once. The research assumes that this point-of-purchase display can have positive effect to the choice of the alternatives on the first webpage.

H1: First webpage display will have larger impact on choices than non-first webpage display.

B-to-C e-commerce operators often provide customers with choices in delivery time based on product availability, the location of the distribution center, or the cost of delivery. In this case, the e-tailer eguo.com offer two types of delivery delay: within one hour and within 8 hours. The difference in these two kinds of delivery time is quite significant. Since convenience and speedy services are among the often-mentioned benefits of online shopping, it is plausible to assume that customers prefer short delivery delay and naturally the product choices associated with short delivery delays. Thus short delivery time can enhance the choice of the specific alternatives.

H2: Short delivery time will have larger impact on choices than long delivery time.

Order confirmation is a significant Internet-specific attribute. Strictly, speaking, online order confirmation has no explicit meaning about the choices and has no impact on the specific choices that the consumers have already made. However, online order confirmation is a significant step in the consumers' purchase process. Unlike offline supermarket, after submitting the online order, the consumer cannot reach the order instantly. There is time lag between order submission and delivery. Moreover, most consumers have to pay cash on delivery since credit card is not popular in China. This time lag and payment method may affect the confirmation of orders. Similar to lagged purchase, order confirmation is measured for individual level, but not choice level. The coefficient for these two variables is different for each SKU. Furthermore, once a consumer places an order online, he or she may cancel the order or cancel any specific choices that are included in the order. From a modeling viewpoint, it is plausible to assume that if order confirmation is included in the linear utility function, the estimation results can be improved, comparing with the function without this variable.

H3: Order confirmation (confirmed orders) will have a positive effect on estimated the choice possibility for a certain alternative.

6. Estimation Results

The study uses the discrete choice procedure in LIMDEP software program (version 7.0) to estimate the parameters of the MNL. Before estimation, a procedure was performed with LIMDEP to detect any violations of the IIA assumption and found no evidence of such problem. The following sections are the results on the respective variables from the MNL models.

6.1 Results of Explanatory Variables

6.1.1 Loyalty variables

As explained in Chapter Five, it is assumed that online customers may be loyal at the SKU level. Besides brand loyalty and size loyalty, SKU loyalty for each individual during data processing stage is also computed. In order to define the base model, it is necessary to compare and determine which loyalty variable(s), brand loyalty and size loyalty, or SKU loyalty, can better explain the purchase behavior across individuals.

[Insert Table Four here]

To indicate the relative contribution of different loyalty variables and to investigate the stability of the coefficients against changes in model specification, S1, which contains only the alternative specific constant, is estimated first. The effect of S1 is to make each individual's purchase probability for a SKU the same and equal to that SKU's share of total purchase (Guadagni and Little, 1983). S1 is the null model with U^2 equals to 0.14899 (Table Four). The study then follows the changes in U^2 to check the effect of loyalty variables, the amount of uncertainty explained by the model when the loyalty variables are added to the null model.

(1) Brand and size loyalty

In S2, the addition of the brand and size loyalty variables produces a jump in U^2 (0.19462, Table Four). It demonstrates, as previous research indicated, that the brand and size loyalty can explain purchase behavior across individuals. However, unlike the brand loyalty with a high t-value, size loyalty, though it has expected positive sign for the parameter, has low t-statistic and turns out to be insignificant. This finding on the size loyalty is not different from the previous research on brand choice for offline environment.

For size loyalty in S2, the size includes 355ml, 2L, 355ml (full box). Eguo.com has a long-term promotion on 355ml Coke for the full box with low price. So the study makes 355ml cola in full box as a separate SKU in our choice set. Actually, for the SKUs of cola, the size only includes 355ml and 2L in our dataset. It is doubtable whether 355ml and 355ml (full box) has affected the t-value for size loyalty since they are actually the same product but sold in different units. Then we introduced another size loyalty variable (only 355ml and 2L) with smoothing constant equals to 0.6. This time, the U^2 is 0.3042 and the parameter for brand loyalty is still stable with value equal to 1.863 and a high t-value. Yet the parameter for the new size loyalty is negative and with a high t-value. The unhealthy sign for the new size loyalty is easy to explain. Since the new size loyalty only includes two sizes and the brand loyalty only includes two brands, together with a SKU with “diet” attribute, the size, brand and diet attribute can only constitute 5 alternatives. So in our dataset, the 7 SKUs cannot be totally explained by the brand loyalty and the new size loyalty variables.

Thus, the original size loyalty variable in S2 is properly constructed.

The low t-statistics for size loyalty in S2 might partly be attributed to the full box promotion in eguo.com. The long-term promotion made the size loyalty quite fragile and contaminated the size loyalty for consumer's online choice behavior. Even an individual is quite fond of 2L cola, when the price for full box 355ml cola is quite attractive, he would like to transfer to buy 355ml cola in full box. It is not surprising that consumers tend to buy products in larger quantities when they are sold are more attractive prices. This proves to some extent the success of marketing promotions. The low t-statistics for size loyalty indicate that online consumers are not quite loyal to specific size for cola products at all. Thus in S2, only brand loyalty, the "classical" loyalty variable explains consumer's choice behavior.

(2) SKU loyalty

Hence, specification S3 introduces SKU loyalty into the model and the U^2 jumps to 0.3899 (Table Four), which is higher than the U^2 in S2. And the parameter for SKU loyalty is 3.355 with a high t-statistic. The higher U^2 in S3 indicates that S3 can account for much more amount of uncertainty explained by the model than S2. In S4, SKU loyalty, brand loyalty and size loyalty are all included in the model. The U^2 declined slightly for S4. The parameter for SKU loyalty stays relatively stable with high t-statistics, yet the parameter for brand loyalty turns negative. The negative sign indicates that collinearity seems to be a serious issue in S4. The correlation between SKU loyalty and brand loyalty turns out to be 0.450, which suggest that the collinearity might introduce biased estimates.

The U^2 of S3 is higher than that of S2, indicating that SKU loyalty provides better explanation for purchase behavior than brand-size loyalty. It also suggests that at least for online environment, consumers are more loyal to SKUs for a specific brand, than to brand-size alternatives when purchasing cola products. Namely, consumer's loyalty to cola products is more specific to a SKU. Thus, the study finally adopts SKU loyalty to measure the unobserved heterogeneity in the subsequent models, excluding brand loyalty and size loyalty.

6.1.2 Net price and price discount

After the loyalty variable is decided, marketing mix variables are introduced to the discrete choice model to examine their effects to the choice behavior. S5 introduces net price and price discount variables and gains a slightly increase in U^2, from 0.38990 in S3 to 0.39015 (Table Four). The parameter of price has the normal negative sign S5, showing that high price results in less purchase and low price results in more purchase. But with very low t-statistics, it means that net price has little effect on consumer's choice behavior. In other words, online consumers are not sensitive to net price. The parameter of price discount is positive and statistically significant at the 0.05 level, showing that price discount promotion can facilitate choice behavior. Clearly price discount is not moving share around the way SKU loyalty does. However, a chi-squared test of S5 relative to S3 shows an improvement that is statistically significant at the 0.010 level.

6.1.3 Display

Webpage display and delivery are among the most Internet-specific attributes for e-tailers. S6, which is S5 plus webpage display, shows nearly no improvement in U^2, comparing with S5. The coefficient of display has the expected sign, yet the t-value is low.

Moreover, the chi-squared test of S6 relative to S5 shows no statistical significance. Thus, H1 cannot be accepted given the estimation results.

6.1.4 Delivery

S7, which is S5 plus delivery, show nearly no improvement in U^2, comparing with S5. The coefficient of delivery is positive and statistically significant, which means that short delivery time has positive effect to the choice decision. Thus, H2 is supported by the estimation results. The chi-squared test of S7 to S5 shows statistical significance at 0.010 level.

6.1.5 Order confirmation

Again, order confirmation is not a choice attribute but an individual-level variable. Thus, before estimating the effect of order confirmation and lagged purchase, the study includes SKU loyalty, price discount and delivery in the new base model, S8. The independent variables in S8 indicate their statistically significance in S1 to S7. S8 shows the stability of coefficients for the independent variables, with U^2 equals to 0.39044 (Table Four).

Then, S9 introduces the order confirmation variable to S8 and U^2 gains an increment from 0.39044 to 0.40190 (Table Four). The coefficients for order confirmation for the SKUs are all positive and significant at 0.001 level. Thus, the estimation results provide support for H3. The chi-squared test of S9 to S8 shows statistical significance at the 0.001 level, with a degree of freedom of 6.

6.1.6 Lagged purchase

S10 introduces the lagged purchase variable to S8 and U^2 gains a very slight jump from 0.39044 to 0.39633 (Table Four). For Coca-cola 2L, Coca-cola 355ml and Coca-cola 355ml in full box, the coefficients of lagged purchase are positive and statistically significant at 0.001 or 0.05 level. Except the Pepsi 355ml in full box, which is zero by definition to violate the singularity problem, the other SKUs have negative coefficients and low t-statistics.

6.2 Further Analysis

Although some variables are not statistically significant, all the coefficients have the algebraic signs that would be expected. By examining the variable's contribution to U^2 and the magnitudes of t-statistic, one can understand the relative importance of the variables in explaining the online choice behavior. Although most of the findings are straightforward, several issues deserve more in-depth discussion to gain a better understanding of their effects.

6.2.1 Loyalty

SKU loyalty has a larger coefficient and t-value than brand and size loyalty and also contributes more improvement to U^2. The coefficient of SKU loyalty, once introduced, tend to be rather stable throughout the various specifications. In this research, three different SKU attributes (brand, size, and flavor) are adopted to portray the cola alternatives. Since the study simply models choice behavior among seven alternatives, it is suitable to adopt SKU loyalty. Moreover, promotion activities for cola products are often conducted at the SKU level. So for our online brand choice scenario, the SKU loyalty is a reliable predictor of the

SKU level choice among individuals. Furthermore, by comparing the value of smoothing constant, no matter it is for brand and size loyalty or just SKU loyalty, the results suggest that online consumers have lower loyalty to brand and size when they shop for these products online.

As it would be expected with most panel data sets, many of the SKUs have small choice shares. Therefore, in the data processing stage, the study omitted the low-share items and consequently lost some information on SKU choice for the individuals. If all the SKUs (18 SKUs in our dataset) were included, the estimate of SKU loyalty may become problematic because of the large number of SKU-specific intercept terms. Rather than expressing the loyalty as a function of SKU-specific intercept terms, which implicitly assumes that consumers maintain preferences toward each individual SKU, this study follows Fader and Hardie's (1996) approach by modeling consumer preferences over the attributes that describe the SKUs in the product category.

6.2.2 Price sensitivity and discount

Price and price discount are the common attributes for both online and offline environment. Some recent studies show that there may be systematic differences in price sensitivity between the online and offline environments (e.g. Degeratu, Rangaswamy and Wu, 2000). Our analysis shows that online consumers are not that price-sensitive to the SKUs of cola products, yet they are quite sensitive to price discount. Since this data set does have information on the promotional activities of cola products in the offline environment, it cannot compare the online price sensitivity with that in the offline environment. It merely shows that the online price discount promotion is effective for estimating the choice probabilities.

Why price discount is statistically significant, while net price is not, for the online brand choice of cola products? Table Five provides some descriptive statistics on shelf price and net price for each of the seven SKUs.

[Insert Table Five here]

The net mean price for the seven SKUs shows that diet Coca-Cola 355ml is the most expensive among all the alternatives. For Coca-cola 355ml, Coca-Cola 355ml (full box), Pepsi 355ml and Pepsi 355ml (full box), the net mean prices are nearly the same with just several RMB cents difference. Coca-cola 2L and Pepsi 2L are the cheapest alternatives. Yet from Table six, it can be seen that the cheapest alternatives among a specific brand doesn’t dominate the percentage of choices, and the SKUs at the middle price levels are frequently chosen.

[Insert Table Six here]

Consumer search behavior has been found to be related to the level of product involvement (e.g. Kujala and Johnson, 1993). Cola, together with other packaged food, is particularly in low-involvement category. The purchase decision process for low-involvement category is routinized and habitual. Cola, with a great market penetration, has been sold in China market since 1979 and has accumulated large amount of routine buyers. The cola buyers are quite familiar with the price, package, and taste of the SKUs in his/her evoked choice set or consideration set. Therefore, consumers, at least on online environment, are loyal to cola products at the SKU level.

Cola purchasers from eguo.com are all Beijing residents. From the delivery addresses they submitted to eguo.com, includes both office and home addresses, it can be found that most of the purchasers are white-collar workers in a company, or university staff or live in a decent building. In 1999, Beijing Statistical Bureau announced that 63% Beijing residents preferred Coca-cola to Pepsi. In 2003, Horizon Group, one of the leading firms in professional research and management consulting in China, also found that Beijing consumers liked Coca-cola better than Pepsi. In 2003, Coca-cola dominated the Beijing carbonic soft drink market with a 65% market share, while Pepsi with 31%. What the data shows in this research is just a reflection of the whole Beijing market for cola.

So, it can be inferred that during the two-stage choice process, from the highly familiar decisions for brand in the consideration set, more than half of Beijing residents would choose Coca-Cola. Then in the consideration set of SKU within a brand, after generating the consideration set, consumers evaluate the choice alternatives in more depth in terms of choice criteria, perform relative comparisons across products (SKU here) on important attributes, such as price, price discount, taste, package, and then make a final purchase decision.

Since the coefficients for price and price discount in this research have the algebraic signs that would be expected, the explanatory importance could be indicated by comparing t-statistics. It is obvious that price is not included in the choice criteria for most online consumers here, while consumers are more attracted to price discount when they make choice decisions. However, the contribution to U^2 by adding price and price discount is quite small, and the small t-statistics for price discount indicate that its explanatory power is limited.

7. Discussion

7.1 Key Findings

By comparing the smoothing constant for offline scenario, the findings of this research indicate that online consumers have lower loyalty on brand or SKU than offline. For setting a base model in this research, SKU loyalty is more effective in explaining the choice among individuals than brand or size loyalty. For the attributes common to online and offline environment, it implies that online consumers are more sensitive to price discount rather than net price, which suggests consumers may switch brand or SKU due to promotional pricing. The findings are similar to studies in the offline market, and especially the low involvement non-durable goods (e.g., Guadagni & Little, 1983, 1998). Also, lagged purchase for online purchasing can contribute to predict further brand choice, which is also confirmed by former research already (e.g., Guadagni & Little, 1983).

As for the Internet-specific attributes, display has little effect on brand choice, while short delivery and order confirmation can have positive effect on brand choice. Display is an attribute that is common to both the online and offline environments, yet it has different characteristics in the two environments. In an offline store, display such as purchase of point displays has gained great attention in brand choice research. For the online environment, the study limits display only to the webpage display, neglecting the banner advertisement, hot link, and floating icon. It turns out that the webpage display has little effect on the SKU choice for the online consumers.

It could be due to several reasons that display is not significant in the linear function for maximum utility. Most consumers, at least cola consumers, will search page by page for their favorite SKUs. Normally the maximum webpage number is 3 for cola display. So it is not very troublesome for the consumers to have a clear search. Although eguo.com will arrange the best selling SKUs on the first page sometimes, the results show no significant differences between the alternatives on the first webpage and those on other webpages.

Comparing with the offline environment, delivery is an Internet-specific attribute. Our results shows that short delivery time have positive effect on the choice behavior of online consumers. Delivery, obviously, is a significant attribute to make choice decision. Unlike other e-tailers in China who normally work with courier firms to deliver its products to consumers, Eguo.com has its own delivery staff to fulfill its customer orders. The self-managed delivery enables eguo.com to avoid the common problems that have plagued other e-tailers due to the cooperation problems between e-tailers and courier firms. Eguo.com takes advantage of the cost savings afforded by online procurement and its flexible delivery management; hence, eguo.com promotes its service by low price and quick delivery. Eguo.com’s order will arrive within one hour” is the slogan for eguo.com’s delivery service. Most of the SKUs will be delivered within one hour, while some of them will be within eight hours or even longer. Order confirmation is an Internet-specific attribute, which is mostly not relevant in the offline environment. The variable is useful in online brand choice estimation.

With lagged purchase, price discount, display and order confirmation as the explanatory variables for the choice model, it comes out the cross tabulation of actual vs. predicted choices. For each occasion, the MNL computes the predicted probability for each of the

SKUs. The choice option with the highest predicted probability is then selected to be predicted choice for that occasion. Doing this for the choice model produces the table as displayed in Table Seven. Summing the diagonal elements in and dividing by the total number of observations (=5317) gives the percentage of correctly classified choices as 48.2%, an obvious improvement from the randomly classified choice as 14.3% (1/7 alternatives). The estimation shows that the model can predict the choice behavior, since there is obvious improvement between the actual and predicted choices.

[Insert Table Seven here]

7.2 Implications

Although this study finds no significant effect from webpage display on brand choice, the importance of webpage design and the display positions of products should not be under-estimated. It may be true that webpage display is not significant in affecting brand choices given a small number of display pages or low-involvement products. For e-tailers that have a number of brands or SKUs that need to be displayed in many webpages, the display position of a brand may make a difference in consumer choices. It is also possible that webpage display alone has limited effect, but together with other marketing variables such as advertising or discount, the compounded effect could be significant.

Since delivery plays a significant role in SKU choice, an e-tailer might arrange short delivery time to the SKUs that are quite popular among the alternatives or have higher marginal profit. Internet order fulfillment is closely related to the service encounter satisfaction for online consumers. Consequently, it is not hard to explain why consumers prefer short delivery time to long delivery time when choosing SKUs among alternatives.

Since the hype era of E-commerce in late 1900s, e-tailers have been offering free (normally for short-distance) or low (normally for short or long distance) delivery to lure new customers. Free delivery or low-cost delivery has become the minimum requirement for selling on-line. Forrester Research (2000) found that the cost of shipping is a major factor in decision-making for 82 percent of online shoppers. Thus, after the e-tailer has gained word-of-mouth recommendation and certain amount of registered customers who have relatively high overall satisfaction with the online purchase service, advanced delivery service should gradually focus on how to facilitate the SKU choice.

Furthermore, there are many other activities that e-tailers can perform to influence the choice behavior of online consumers. Price discount is statistically significant on the effects of product choice and can be an effective strategy for e-tailers. Like many e-tailers, eguo.com promotes price discount as one of their competitive attributes for online buyers all the time. However, the everyday low price strategy has an effect on all the discounted items. For further analysis, e-tailers should consider the impact of a brand's price promotion frequency and the depth of promotional price discounts on the price consumers expect to pay for that brand. That is the key point for e-tailers to facilitate the effect of the price discount strategy. Kalwani and Yim (1992) found that both promotion frequency and the depth of price discount had a significant impact on price expectations. Consumer expectations of both price and promotional activities should be considered in explaining consumer brand choice behavior. Specifically, the presence of a promotional deal when one is not expected (or the reverse) may have a significant impact on consumer brand choice.

7.3 Limitations and Suggestions

A deficiency of the database is that it does not contain any explicit information on non-price promotion of all the products, including offline promotion activities. Recall eguo.com only has non-price promotion to Coca-Cola in the website and even this promotion information does not cover all the period for the panel data. Moreover, the data also lack the non-price off-line promotion in Beijing district, such as advertisement in newspapers, TV, flyers, bag stuffers during the two years for all panel data. Furthermore, the company only promotes its products in specific market area. Consequently, the study cannot examine the effect of offline promotions on the choice behavior of consumers shopping online. Ideally, a study should include non-price promotion and promotional activities in the offline environment such as newspaper advertising and promotional flyers so that the researcher can consider the impact of offline promotion on the online brand choice (Guadagni and Little, 1983). Such a model would make the results more powerful and convincing.

Another deficiency of the database is that it does not differentiate household buyers and business buyers. The two kinds of buyers may have significant and systematic differences in their purchasing behaviors. Although only 1% panelists in the database are business buyers, they normally purchases large amount of products on a single purchase occasion. Yet, for MNL, purchase amount is neglected in modeling the brand choice. However, it should be admitted that business buyers may exhibit choice behaviors that are different from the household buyers. With sufficient data for both types of buyers, future studies can compare such differences systematically.

This research adopted a conventional model of brand choice and used online panel data. Existing models may not be inadequate for fully capturing the richness of the choice processes that are increasingly feasible to observe in the online market (Wu and Rangaswamy, 2003). During the modeling process, this research has suffered the similar problems and cannot fully grasp the brand choice processes for online buyers. Research on online consumer behavior relies heavily on the data mining of transactional and clickstream data from the web server logs. However, the web server logs are commonly designed to debug web servers and the data they provide is noisy and insufficient, requiring the use of heuristics to reconstruct events (Kohavi, 2001). On the other hand, online data contain not only consumer's choice information but also the search processes prior to the choice stage. Such new data are likely to challenge traditional modeling approach. Thus, in the future research, some inter-disciplinary approach, such as machine learning or data mining, needs to be adopted to deal with the large and noisy online datasets and to explore the complicated choice behavior for online consumers. For instance, choice models based on fuzzy set theory in Wu and Rangaswamy's (2003) research may be necessary for deriving managerially relevant understanding of choice behavior in online markets.

For the future, e-commerce researchers should position their work carefully against traditional models. For example, what is the contribution of this research stream in terms of theory, substance and methodology? What theories continue to hold in the online setting? What theories do not hold and why (Mahajan & Venkatesh, 2000)? Although previous studies duly caution the limitations and constraints in applying existing modeling approaches to addressing the research problems in e-business, future research should emphasize the strengths of these new methods and the untapped opportunities they provide for better understanding online consumer behaviors.

As for further research topics, there is also a need to compare the e-commerce developed of developed countries with that in developing markets. China, which the e-tailer of this research is from, has lower credit card usage and lower computer penetration rate. E-tailers in developed countries will normally choose credit card as the payment method, and e-tailers may outsource the delivery and other logistics tasks. Given these differences between the e-tailers in various countries, how the marketing mix variables and Internet-specific attributes affect the choice of online consumers warrant investigation. If similar data can be collected from e-tailers of different countries, comparative study can potentially generate important empirical findings and managerial implications. These research problems represent an interesting area of quantitative research in modeling online consumer brand choice and other related fields for years to come.

Table One: Estimation for smoothing constant for cola products

Estimation results for **brand loyalty** smoothing constant

	Iteration:					
	1	2	3	4	5	6
Initial smoothing constant (λ 0)	0.650	0.566	0.517	0.498	0.492	0.491
Logit coefficients:						
Brand loyalty (LOY)	2.024	1.928	1.901	1.889	1.885	1.885
Brand loyalty derivative (DLOY)	-0.169	-0.095	-0.036	-0.011	-0.002	-0.0009
(std. Error for DLOY)	(-0.034)	(0.039)	(0.041)	(0.042)	(0.042)	(0.042)
Brand-specific constant:						
α -Coca-Cola	0.621	0.657	0.670	0.674	0.675	0.675
Updated smoothing constant (λ)	0.566	0.517	0.498	0.492	0.491	0.491
Log likelihood	-2429.089	-2428.608	-2430.071	-2430.741	-2430.954	-2430.990

Note: For iteration 1 to 2, all coefficients are significant at p=0.00;

From iteration 3 afterwards, with exception of the coefficients for DLOY, the left two coefficients are significant at p=0.00

Estimation results for **Size Loyalty** smoothing constant:

	Iteration:					
		2	3	4	5	6
Initial smoothing constant (λ 0)	0.500	0.408	0.364	0.345	0.336	0.333
Logit coefficients:						
Size loyalty (LOY)	3.022	2.962	2.933	2.920	2.914	2.912
Size loyalty derivative (DLOY)	-0.260	-0.131	-0.057	-0.025	-0.008	-0.003
(std. Error for DLOY)	0.043	0.047	0.049	0.049	0.049	0.049
Size-specific constant:						
α -355Ml	0.368	0.376	0.377	0.378	0.378	0.378
α - 2L	-0.134	-0.147	-0.153	-0.156	-0.157	-0.157
Updated smoothing constant (λ)	0.408	0.364	0.345	0.336	0.333	0.333
Log likelihood	-3268098	-3247.02	-3239.97	-3237.574	-3236.56	-3236.24

Note: For iteration 1 to 2, all coefficients are significant at p=0.00;

From iteration 3 afterwards, with exception of the coefficients for DLOY, the left two coefficients are significant at p=0.00

Table Two: Estimation for smoothing constant for coffee

Estimation results for **brand loyalty** smoothing constant

	Iteration:					
	1	2	3	4	5	6
Initial smoothing constant (λ 0)	0.6200	0.6362	0.6431	0.6486	0.6530	0.6600
Logit coefficients:						
Brand loyalty (LOY)	2.6748	2.7077	2.7347	2.7570	2.7754	2.7908
Brand loyalty derivative (DLOY)	0.02358	0.01868	0.01508	0.0124	0.0103	0.0086
(std. Error for DLOY)	(0.2308)	(0.0772)	(0.0762)	(0.0754)	(0.0748)	(0.0742)
Brand-specific constant:						
α -brand 1	-0.3622	-0.3608	-0.3597	-0.3589	-0.07600	-0.3576
Updated smoothing constant (λ)	0.6362	0.6431	0.6486	0.6530	0.6600	0.6600
Log likelihood	-570.4522	-570.5042	-570.5577	-570.6090	-570.6557	-570.6978

Note: For each iteration, all coefficients are significant at p=0.00, except DLOY.

Estimation results for **Size Loyalty** smoothing constant:

	Iteration:					
	1	2	3	4	5	6
Initial smoothing constant (λ 0)	0.5500	0.4765	0.4317	0.4017	0.3858	0.3716
Logit coefficients:						
Size loyalty (LOY)	2.3412	2.2294	2.1638	2.1237	2.1237	2.1189
Size loyalty derivative (DLOY)	-0.1403	-0.1026	-0.0651	-0.0350	-0.3495	-0.3490
(std. Error for DLOY)	(0.0847)	(0.0899)	(0.0930)	(0.0947)	(0.0947)	(0.0940)
Size-specific constant:						
α -small	0.4885	0.5013	0.5094	0.5143	0.5143	0.5143
α - middle	-0.09583	-0.0964	-0.0967	-0.0968	-0.0968	-0.0968
Updated smoothing constant (λ)	0.4765	0.4317	0.4017	0.3858	0.3716	0.3716
Log likelihood	-900.3844	-899.2754	-898.9706	-898.9646	-898.9464	-898.9464

Note: For iteration 1, all coefficients are significant at p=0.00;

From iteration 2 afterwards, with exception of the coefficients for DLOY, the left two coefficients are significant at p=0.00

Table Three: Estimation for smoothing constant for SKU Loyalty

Estimation results for **SKU Loyalty** smoothing constant:

	Iteration:				
	1	2	3	4	5
Initial smoothing constant (λ 0)	0.65	0.525	0.4634	0.449	0.446
Logit coefficients:					
SKU loyalty (LOY)	3.457	3.371	3.344	3.339	3.337
SKU loyalty derivative (DLOY)	-0.398	-0.173	-0.041	-0.014	-0.003
(std. Error for DLOY)	0.037	0.042	0.045	0.045	0.045
Net Price	-0.109	-0.110	-0.110	-0.110	-0.110
Price discount	0.283	0.286	0.287	0.287	0.287
Size-specific constant:					
α -sku01	0.046	0.024	0.014	0.013	0.012
α -sku02	0.620	0.653	0.659	0.659	0.659
α -sku03	-0.013	-0.032	-0.039	-0.040	-0.041
α -sku04	-0.605	-0.648	-0.662	-0.665	-0.666
α -sku05	-0.377	-0.416	-0.430	-0.431	-0.432
α -sku06	-0.682	-0.730	-0.746	-0.748	-0.749
Updated smoothing constant (λ)	0.525	0.4634	0.449	0.446	0.446
Log likelihood	-6315.156	-6311.269	-6306.758	-6306.651	-6306.562

Note:

For iteration 1 to 3, all coefficients are significant at p=0.00, except α -sku01 and α -sku03;From iteration 4 afterwards, with exception of the coefficients for DLOY, α -sku01 and α -sku03, other coefficients are significant at p=0.00

Table Four: Results of the MNL brand choice model

	Coefficient estimates[a]						
Specification	**S1**	**S2**	**S3**	**S4**	**S5**	**S6**	**S7**
U^2	0.14915	0.19509	0.39004	0.37869	0.39032	0.39035	0.39062
adjusted U^2	0.14899	0.19462	0.38990	0.37828	0.39015	0.39016	0.39043
SKU loyalty			3.355	3.529	3.355	3.354	3.353
			(63.088)*	(32.000)*	(63.023)*	(62.996)*	(62.984)*
Brand Loyalty		1.853		-0.450			
		(18.660)*		(-3.280)**			
Size Loyalty		0.057		0.037			
		(0.940)		(0.503)			
Net Price					-0.179	-0.143	-0.195
					(-0.339)	(-0.213)	(-0.370)
Price discount (Promotion)					0.180	0.206	0.177
					(2.156)***	(2.292)***	(2.126)***
Display						0.107	
						(0.778)	
Delivery							0.148
							(2.464)**
SKU constants							
Coca-Cola 2L	0.522	0.111	0.030	-0.092	-0.015	-0.057	-0.010
	(9.923)*	(1.301)	(0.474)	(0.967)	(-0.127)	(-0.428)	(-0.085)
Coca-Cola 355ml	1.362	0.926	0.654	0.724	0.490	0.452	0.502
	(29.195)*	(12.048)*	(11.744)*	(8.516)*	(4.964)*	(4.126)*	(5.080)*
Coca-Cola 355ml (full box)	-0.077	-0.780	-0.216	-0.305	-0.216	-0.335	-0.176
	(-1.290)	(-7.573)*	(-3.157)**	(-2.738)**	(-2.934)**	(-1.973)***	(-2.336)**
Diet Coca-cola 355ml	-0.664	-1.057	-0.701	-0.610	-0.674	-0.843	-0.651
	(-9.300)*	(-9.472)*	(-8.942)*	(-5.131)*	(-5.226)*	(-3.331)*	(-5.031)*
Pepsi 2L	-0.278	-0.051	-0.410	-0.290	-0.477	-0.417	-0.476
	(-4.382)*	(-0.558)	(-5.769)*	(-2.799)**	(-3.745)*	(-2.788)**	(-3.733)*
Pepsi 355ml	-0.844	-0.761	-0.776	-0.682	-0.785	-0.796	-0.779
	(-11.121)*	(-6.711)*	(-9.505)*	(-5.658)*	(-9.403)*	(-9.389)*	(-9.330)*
Log likelihood	-8804.898	-3569.554	-6312.096	-2755.351	-6309.138	-6308.834	-6306.055

[a] Numbers in parentheses are the t-statistics for the estimated coefficients.

*Significant at α =0.001 in the one-tailed asymptotic t-test

** Significant at α =0.01 in the one-tailed asymptotic t-test

*** Significant at α =0.05 in the one-tailed asymptotic t-test

Table Four Continued

	Coefficient estimates[a]				
Specification	**S8**		**S9**		**S10**
U^2	0.39062		0.40218		0.39633
adjusted U^2	0.39044		0.40190		0.39605
SKU loyalty	3.3541		3.3935		3.3207
	(63.033)*		(62.332)*		(61.937)*
Price discount (Promotion)	0.1879		0.1933		0.1851
	(2.404)***		(2.463)***		(2.362)***
Delivery	0.1480		0.1433		0.1452
	(2.459)***		(2.373)***		(2.400)***
SKU constants (ASCs)					
Coca-Cola 2L	0.0283		-1.4800*		-0.1826
	(0.453)		(-8.120)		(-2.159)
Coca-Cola 355ml	0.4887*		-1.5963*		0.0915
	(5.314)		(-9.194)		(0.862)
Coca-Cola 355ml (full box)	-0.1867**		-1.9767*		-0.3509*
	(-2.656)		(-7.702)		(-3.602)
Diet Coca-cola 355ml	-0.6889*		-1.3586)*		-0.6816*
	(-8.745)		(-6.476)		(-6.827)
Pepsi 2L	-0.4377*		-0.9625*		-0.4166*
	(-6.040)		(-5.513)		(-4.545)
Pepsi 355ml	-0.7850*		-1.5462*		-0.7024*
	(-9.591)		(-7.056)		(-6.965)
Interactions with ASCs					
Order Confirmation		1_order confirm	1.7492*	1_purchase lag	0.4522*
			(9.021)		(3.676)
Or		2_order confirm	2.3639*	2_purchase lag	0.7745*
			(14.149)		(6.986)
Purchase LAG		3_order confirm	2.0439*	3_purchase lag	0.3511***
			(7.657)		(2.569)
		4_order confirm	0.8347*	4_purchase lag	-0.0801
			(3.684)		(-0.500)
		5_order confirm	0.6632*	5_purchase lag	-0.1003
			(3.471)		(-0.695)
		6_order confirm	0.9360*	6_purchase lag	-0.3271
			(3.956)		(-1.892)
Log likelihood	-6306.123		-6185.257		-6246.979

[a] Numbers in parentheses are the t-statistics for the estimated coefficients.

*Significant at α =0.001 in the one-tailed asymptotic t-test

** Significant at α =0.01 in the one-tailed asymptotic t-test

*** Significant at α =0.05 in the one-tailed asymptotic t-test

Table Five: Descriptive Statistics of price for the SKUs

SKUs	Mean (Shelf price)	Std. D. (Shelf price)	Mean (Net price*)	Std. D. (Net price)
Coca-Cola 2L	6.65	0.2399	0.34	0.01200
Coca-Cola 355ml	1.96	0.2416	0.55	0.06805
Coca-Cola 355ml (full box)	49.75	1.1063	0.59	0.01299
Diet Coca-Cola 355ml	2.58	0.1764	0.73	0.04969
Pepsi 2L	6.52	0.1865	0.33	0.00932
Pepsi 355ml	1.99	0.0256	0.56	0.00721
Pepsi 355ml (full box**)	45.61	0.5280	0.54	0.00620

* Net Price is the price per 100ml for the SKUs, in RMB. 1 RMB approximately equals to 0.121 USD.

** Full box equals to 24 cans of 355ml coke.

Table Six: Purchase Percent for the SKUs

SKUs	Times of being chosen	Percent of being chosen
Coca-Cola 2L	972	18.3%
Coca-Cola 355ml	2253	42.4%
Coca-Cola 355ml (full box)	534	10.0%
Diet Coca-Cola 355ml	297	5.6%
Pepsi 2L	437	8.2%
Pepsi 355ml	248	4.7%
Pepsi 355ml (full box**)	577	10.8%
Total	5318	100%

** Full box equals to 24 cans of 355ml coke.

Table Seven: Cross-tabulation of actual vs. predicted choices

Actual/Predicted Choices	C-2L	C-355ml	C-355ml(B)	DietC-355	P-2L	P-355ml	P-355ml(B)	Total
C-2L	463	168	68	42	103	42	84	971
C-355ml	220	1451	121	113	112	94	142	2253
C-355ml(B)	41	183	217	21	22	18	31	534
DietC-355	43	101	24	60	25	15	30	297
P-2L	133	74	30	21	115	20	44	437
P-355ml	26	111	16	15	20	34	25	248
P-355ml(B)	45	164	58	25	40	25	220	577
Total	971	2253	534	297	437	248	577	5317

C indicates Coca-Cola

P indicates Pepsi

B indicates full box

Row indicators are actual choices and column indicators are predicted choices.

Predicted total is F(k,j,i)=Sum(i=1,...,N) P(k,j,i).

Column totals may be subject to rounding error.

Appendix: Nonlinear Estimation Algorithm for Smoothing Constant

The usual multinomial logit (MNL) model assumes that a linear combination of the attributes is linked to choice probabilities as follows:

$$p_j^h(t) = e^{v_j^h(t)} \Big/ \sum_k e^{v_k^h(t)} \quad , \qquad \text{where} \qquad (1)$$

$p_j^h(t)$ = the probability that household h chooses brand j on purchase occasion t ,

$v_j^h(t) = \sum \beta_r x_{jr}^h(t)$

= the deterministic component of utility of brand j to household h at purchase occasion t,

$x_{jr}^h(t)$ = rth explanatory variable for brand j and household h on purchase occasion t, r=1,…,R,

β_r =coefficient to be estimated.

In applications to household scanner data, the $x_{jr}^h(t)$ generally include brand-specific intercept terms and marketing mix variables such as price and difference types of promotions, and sometimes exposure to television advertising. In addition, variables may be added to capture other sources of variation across households and over time.

We first consider the case of a MNL model with any number of variables, of which one, $x_{jm}^h(t)$, is nonlinearly dependent on a single parameter α. Because α is imbedded within α $x_{jm}^h(t)$ it cannot be estimated directly as an ordinary logit coefficient. For expositional clarity, we suppress the subscripts m, h and j, and make α explicit. The notation for $x_{jm}^h(t)$ becomes $x(t, \alpha)$.

First expand $x(t, \alpha)$ in a Taylor series around a startling value α_0 :

$$x(t,\alpha) = x(t,\alpha_0) + \frac{dx(t,\alpha)}{d\alpha}\Big|_{\alpha 0}(\alpha - \alpha_0) + \sum_{n=2}^{\infty} \frac{d^n x(t,\alpha)}{d\alpha^n}\Big|_{\alpha_0} \frac{(\alpha - \alpha_0)^n}{n!} \tag{2}$$

If *x(t,* α *)* is smooth (e.g., its derivatives with respect to α are bounded) in an interval containing both α_0 and the maximum likelihood estimate (MLE) value of α, then the second and higher-order terms in (2) will approach 0 as a0 approaches its MLE value. Letting $x'(t,\alpha) = dx(t,\alpha)/d\alpha$, we have as a current approximation for *x(t,* α *),*

$$x(t,\alpha) \cong x(t,\alpha_0) + x'(t,\alpha_0)(\alpha - \alpha_0) \tag{3}$$

Which becomes exact upon convergence of α_0 to α.

Letting β be the coefficient for *x(t,* α *)* in the MNL, the contribution of *x(t,* α *)* to utility is approximately

$$\beta x(t,\alpha) \cong \beta x(t,\alpha_0) + \beta(\alpha - \alpha_0)x'(t,\alpha_0) \tag{4}$$

From (4), we see that we can better represent the contribution of β *x(t,* α *)* to utility by including $x'(t,\alpha_0)$ as well as $x(t,\alpha_0)$ among the variables in the MNL estimation. Denoting the resulting estimates by β ' and β , we end up with a contribution to utility of

$$\beta x(t,\alpha) \cong \beta x(t,\alpha_0) + \beta' x'(t,\alpha_0) \tag{5}$$

Comparing (4) and (5), we see that $\beta' \cong \beta\alpha - \alpha_0)$, or

$$\alpha \cong \alpha_0 + \beta' / \beta \tag{6}$$

Thus, we can use β ' to obtain a new, better estimate of α . Substituting this for α_0, we iterate until ($\alpha - \alpha_0$) becomes as small as desired; i.e., until $\beta' \cong 0$. Usually this requires only a few iterations.

Although each iteration makes use of the Taylor series approximation for $x(t,\alpha)$, by iteratively running the MNL linear estimation routine, these approximations converge to the exact value of the function. Thus we have our Nonlinear Estimation Algorithm (NEA):

(1) Choose a starting value of α, say α_0.

(2) Calculate $x(t,\alpha)$ and $x'(t,\alpha)$ at α_0 for all observations *t*.

(3) Include $x(t,\alpha_0)$ and $x'(t,\alpha_0)$ along with all the other variables in the logit model, and estimate coefficients in the usual manner.

(4) Update α_0 using equation (6): $\alpha_0 \leftarrow (\alpha_0 + \beta'/\beta)$.

(5) Return to step (2) and iterate until α_0 converges, i.e., until the coefficient of x' is indistinguishable from 0. Denote the final estimates by $\hat{\alpha}$, $\hat{\beta}$, and $\hat{\beta}'$.

(6) Calculate the standard error of $\hat{\alpha}$ from $SE(\hat{\alpha}) = SE(\hat{\beta})/\hat{\beta}'$

Extension to multiple nonlinear parameters is straightforward. Suppose that one of the independent variables involves a J-dimensional vector of imbedded parameters: $x(t,\alpha)$, where $\alpha = (\alpha_1, \alpha_2, \ldots \alpha_J)$. The development now uses the linear terms of a multivariate Taylor series. The iterative estimation procedure still applies. The partial derivatives of $x(t,\alpha)$ with respect to the α_J are evaluated at some starting vector $\alpha_0 = (\alpha_{10}, \alpha_{20}, \ldots, \alpha_{J0})$. All of the derivative variables are included in the logit model. The analog of (6) is

$$\alpha_J \cong \alpha_{J0} + \beta'_J / \beta \tag{7}$$

Where β'_J represents the logit coefficient for the *j*th derivative variable, and β is the logit coefficient for $x(t,\alpha)$. The update in step (4) becomes ($\alpha_{J0} \leftarrow (\alpha_{J0} + \beta_J'/\beta)$). If there are several variables containing nonlinear parameters, each variable and its parameters are treated in the same way as α.

Although the introduction of multiple nonlinear parameters requires nothing new theoretically or conceptually, the number of iterations needed for convergence tends to increase. In a case where a model with a single α might not require 2-3 runs of the linear estimation program, our experience has shown that a set of 6 α_J 's takes about 5-7 iterations.

--------------*Extracted from Fader, Peter S., Lattin, James M., Little, John D C. (1992) "Estimating nonlinear parameters in the multinomial logit model" Marketing Science, 11(4), 372-385.*

Bibliography:

Allenby, G. M., & Lenk, P. J. (1995, July). *Reassessing brand loyalty, price sensitivity, and merchandising effects on consumer brand choice*. Journal of Business and Economic Statistics, 13 (3), 281-289.

Ancarani, F., & Shankai, V. (2002). *Price levels and price dispersion on the Internet: a comparison of pure play Internet, bricks-and-mortar, and bricks-and-clicks retailers.* MIT E-business research center working paper.

Batlas, G. (1998, October). *An integrated model of category demand and brand choice.* Journal of the Market Research Society, 40(4), 295-307.

Bawa, K., & Shoemaker, R. W. (1987, November). *The effects of a direct mail coupon on brand choice behavior*. Journal of Marketing Research, 24 (4), 370-376.

Ben-Akiva, M., & Lerman, S. R. (1985). *Discrete choice analysis: theory and application to travel demand*. Cambridge, MA: The MIT press.

Bentz, Y., & Merunka, D. (2000). N*eural networks and the multinomial logit for brand choice modeling: a hybrid approach*. Journal of Forecasting, 19, 177-200.

Bowen, H., & Wiersema, M. F. (2003). *Modeling limited dependent variables: methods and guidelines for researchers in strategic management.* Vlerick Leuven Gent Working Paper Series 2003/30.

Campanelli, M. (2000). *Customer Service: Even online, strong customer service skills are a must if you want your site to be one of the survivors*. Entrepreneur's Nettpreneur, www.Entrepreneur.com.

Chintagunta, P. K., & Dube, J. (2003). *Estimating a SKU-level brand choice model combining household data and store data*. Working papers of University of Chicago

Degeratu, A., Rangaswamy, A., & Wu, J. (2000). *Consumer choice behavior in online and traditional supermarkets: the effects of brand name, price, and other search attributes*. International Journal of Research in Marketing, 17 (1), 55-78.

Erdem, T. (2001, April). T*esting for choice dynamics in panel data*. Journal of Business and Economic Statistics, 19 (2), 142-153.

Fader, P. S., Lattin, J. M., & Little, L. (1992, Fall). E*stimating nonlinear parameters in the multinomial logit model*. Marketing Science, 11 (4), 372-385.

Feinberg, F. M., & Russell, G. J. (2003). *A consistent loyalty measure for generalized logit models.* Journal of Marketing Research, forthcoming.

Fry, T. R., & Harris, M. (1998). *Testing for independence of irrelevant alternatives, some empirical results.* Sociological Methods and Research, 26 (3), 401-423.

Gonul, F., & Srinivasan, K. (1993, Summer). *Modeling multiple sources of heterogeneity in multinomial logit models: methodological and managerial issues*. Marketing Science 12 (3), 213-229.

Greene, W. H. (1995). *LIMDEP version 7.0 user's manual.* Econometric Software, Australia.

Greene, W. H. (2000). *Econometric analysis (fourth edition).* Prentice-Hall, Inc.

Guadagni, P. M., & Little, J. (1983, Summer). *A logit model of brand choice calibrated on* *scanner data.* Marketing Science, 2 (3), 203-238.

Guadagni, P. M., & Little, J. (1998, June/July). *When and what to buy: A nested logit model of coffee purchase.* Journal of Forecasting, 17 (3-4), 303-326.

Gupta, S. (1988, November). *Impact of sales promotions on when, what, and how much to buy.* Journal of Marketing Research, 25 (4), 342-355.

Haubl, G., & Trifts, V (2001). *Consumer decision making in online shopping environments: the effects of interactive decision aids*. Marketing Science, 19 (2), 4-21.

Holland, J. & Baker, S. M. (2001, Autumn). *Customer participation in creating site brand loyalty.* Journal of Interactive Marketing, 15(4), 34-45.

Honore, B., & Kyriazidou, E. (2000,July). *Panel data discrete choice models with lagged dependent variables*. Econometrics, 68 (4), 839-874.

Howard, J. A. (1989). C*onsumer behavior in marketing strategy*. Englewood cliffs, N.J.: Prentice Hall.

Hsiao, C. (1990). *Analysis of panel data.* Cambridge University Press.

Hutchinson, J. W., Kamakura, W. A., & Lynch, J. G. (2000). *Unobserved heterogeneity as an alternative explanation for "reversal" effects in behavioral research.* Journal of Consumer Research, 27, 324-344.

Jain, D.C., Vilcassim, N. J., & Chintagunta, P. K. (1994, July). *A random-coefficients logit brand-choice model applied to panel data.* Journal of Business and Economic Statistics, 12 (3), 317-328.

Janssen, M., & Moraga, J. L. (2001). *Pricing, consumer search and maturity of Internet market*. Working paper.

Jiang, J. J., Hsu, M. K., Klein, G., & Lin, B. (2000). *E-commerce user behavior model: an empirical study*. Human System Management, 19, 265-276.

Kalwani, M. U., & Yim, C. K. (1992). *Consumer Price and Promotion Expectations: An Experimental Study.* Journal of Marketing Research, 29 (1), 90-100.

Kamakura, W. A., & Russell, G, J. (1993). *Measuring brand value with scanner data.* International Journal of Research in Marketing, 10, 9-22.

Kannan, P. K., & Kopalle, P. K. (2001, Spring). *Dynamic pricing on the Internet: importance and implications for consumer behavior*. International Journal of Electronic Commerce, 5 (3), 63-83.

Keane, M. P. (1997, July). M*odeling heterogeneity and state dependence in consumer choice behavior*. Journal of Business and Economic Statistics, 15 (3), 310-327.

Kohavi, R. (2001). *Mining E-Commerce Data: The Good, the Bad, and the Ugly.* KDD 01 San Francisco CA USA.

Kohavi, R., Brodley, C. E., Frasca, B., Mason, L.,& Zheng, Z. (2000). *KDD-Cup 2000 Organizer's Report: Peeling the Onion.* SIGKDD Explorations, 2 (2), 86-93.

Krishnamurthi, L. & Raj, S. P. (1991, Spring). *An empirical analysis of the elasticity.* Marketing Science, 10, 172-183.

Kujala, J. T., & Johnson, M. D. (1993). *Price knowledge and search behavior for habitual* *low involvement food purchases*. Journal of Economic Psychology, 14, 249-265.

Kumar, V., Karande, K., & Reinartz, W. (1998). T*he impact of internal and external reference prices on brand choice: the moderating role of contextual variables*. Journal of Retailing, 74(3), 401-426.

Lattin, J. M. (1987, winter). *A model of balanced choice behavior*. Marketing Science, 6 (1), 48-65.

Leeflang, P., Wittink, D., Wedel, M., & Naert, P. A. (2000). B*uilding models for marketing decisions*. Kluwer Academic Publishers.

Lilien, G. L., & Rangaswamy, A. (2000). *Modeled to bits: decision models for the digital, networked economy*. International Journal of Research in Marketing, 17, 227-235.

Lilien, G. L., Kotler, P., & Moorthy, K. S. (1992). *Marketing Models*. Englewood cliffs, N.J.: Prentice Hall

Mahajan, V., & Venkatesh, R. (2000). *Marketing modeling for e-business*. International Journal of Research in Marketing, 17, 215-225.

Manrai, A. (1995). *Mathematical models of brand choice behavior*. European Journal of Operational Research, 82, 1-17.

Manrai, A., & Andrews, R. L. (1998). *Two-stage discrete choice models for scanner panel data: an assessment of process and assumptions*. European Journal of Operational Research 111 (1998) 193-215.

Matsatsinis, N. F., & More, M. N. (2000). *A knowledge base for brand choice model* *selection.* Technical University of Crete, working paper.

Matsatsinis, N. F., & Samaras, A. P. (2000). *Brand choice model selection based on consumers' multicriteria preferences and experts' knowledge*. Computers and Operations Research, 27, 689-707.

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

McFadden,D. (1974). *Conditional logit analysis of qualitative choice behavior*. Frontiers in Econometrics, Zarembka ed., New York: Academic Press, 105-142.

Mela, C. F., Gupta, S., & Lehmann, D. R. (1997, May). "*The long-term impact of promotion and advertising on consumer brand choice*". Journal of Marketing Research, 34(2), 248-261.

Morgan, J., & Landwehr, J. T. (1988, Winter). *Removing heterogeneity bias from logit model estimation*. Marketing Science,7 (1), 41-58.

Nowlis, S., & Simonson, I. (1996, February). *The Effect of New Product Features on Brand Choice*. Journal of Marketing Research, 33 (1), 36-46.

Papatla, P., & Krishnamurthi, L. (1996, February). *Measuring the dynamic effects of promotions on brand choice*. Journal of Marketing Research, 33 (1), 20-35.

Raju, J. S., Dhar, S. K., & Donald G. M. (1994, Spring). *The effect of package coupons on brand choice*. Marketing Science, 13(2), 145-154.

Ratchford, B., T., Talukdar, D., & Lee M.(2001, Spring). *A model of consumer choice of the Internet as an information source*. International Journal of Electronic Commerce, 5 (3), 7-21.

Russel, G. L., & Petersen, A. (2000). *Analysis of cross category dependence in market basket selection.* Journal of Retailing, 76(3), 367-392.

Ruud, P. A. (1996). *Approximation and simulation of the multinomial probit model: an analysis of covariance matrix estimation*. Working paper of University of California, Berkeley.

Seetharaman, P. B., Ainslie, A., & Chintagunta P. K. (1999, November). *Investigating household state dependence effects across categories.* Journal of Marketing Research, 36 (4), 488-500.

Shankar, V., Smith, A., & Rangaswamy, A. (2003). *Customer satisfaction and loyalty in online and offline environments.* International Journal of Research in Marketing, 20 (2), 153-175.

Srinivasan, S., Anderson, R., & Ponnavolu, K. (2002). *Customer loyalty in e-commerce: an exploration of its antecedents and consequences*. Journal of Retailing, 78, 41-50.

Tabachnick, B. G. & Fidell, L. S. (1996). *Using multivariate statistics (third edition)*. HarperCollins College Publishers.

Tonegawa, K. (2002). *A review and analysis: E-Commerce research forum 1998-2001*. Working paper for ERF at MIT.

Ward, M. R. (1999). W*ill e-commerce compete more with traditional retailing or direct marketing*? Working paper for ERF at MIT.

Wen,C.,H, & Koppelman, F. (2000). *The generalized nested logit model*. Working paper of Feng Chia University.

Wills. H. (1987). *A note on specification tests for the multinomial logit model*. Journal of Econometrics, 34, 263-274.

Wu, J., & Rangaswamy. A. (2003). *A fuzzy set model of search and consideration with an application to an online market*. Marketing Science, 22 (3), 411-434.