University of Alberta

Towards a Positive Theory of Rational Choice: From Substantive to Procedural Rationality

By

Ashish Kumar Sinha ©

A thesis submitted to the Faculty of Graduate Studies and Research in partial fulfillment of the requirements for the degree of DOCTOR OF PHILOSOPHY

IN

MARKETING

FACULTY OF BUSINESS

EDMONTON, ALBERTA

SPRING 1997

0-612-21637-3

University of Alberta

Library Release Form

Name of Author : Ashish Kumar Sinha

Title of Thesis : Towards a Positive Theory of Rational Choice: From Substantive to Procedural Rationality

Degree : Doctor of Philosophy

Year this Degree Granted : 1997

A. Sinha

Ashish Kumar Sinha
Department of Marketing and International Management
University of Waikato, Hamilton, NZ

Date 30-12-96

University of Alberta

Faculty of Graduate Studies and Research

The Undersigned certify that they have read, and recommended to the Faculty of Graduate Studies and Research for acceptance, a thesis entitled TOWARDS A POSITIVE THEORY OF RATIONAL CHOICE: FROM SUBSTANTIVE TO PROCEDURAL RATIONALITY submitted by ASHISH KUMAR SINHA in partial fulfillment of the requirements for the degree of DOCTOR OF PHILOSOPHY in MARKETING.

Terry Elrod

Dr. Terry Elrod
(Supervisor)

Dr. Harry Timmermans

Dr. Peter Popkowski Leszczyc

Denise Young

Dr. Denise Young

A. D. Shocker

Dr. Allan D. Shocker

Dated December 23, 1996

Dedication

To Terry Elrod, for believing in me more than I did in myself.

Towards a Positive Theory of Rational Choice:
From Substantive to Procedural Rationality.

The last two decades of marketing has seen an exponential growth of academic articles in the area of consumer choice models. A casual perusal of any leading marketing journal will show that choice is the dominant research agenda in marketing. Though a number of different models have been proposed in the literature, most of them assume that consumer use all the available relevant information for making decisions. Recently, a few researchers, using cost-benefit analysis, have shown that this assumption is often violated. Earlier taxonomies that have been proposed in the literature also identify this area of consumer choice as not having received its due attention. The inability of models, that assume computational limitations of a consumer, to capture dynamic behavior was a major stumbling block for research in this area. Attempts that have been made to capture dynamic behavior have found models to be computationally burdensome.

I propose a dynamic brand choice model that explicitly assumes computational limitations of a decision maker. Borrowing concepts from "bounded rationality" this work argues that due to computational limitations the amount of information that a person can attend to is often limited. The basic premise of this research is that attention is a scarce resource. A "probabilistic activation function" is used to implement the construct of "attention" into formal models of choice. Scanner data is used to understand and predict the effects of recent purchase behavior, brand preferences and marketing mix variables on attention. The author also shows that the "probabilistic activation function" can help in capturing the effects of consideration sets and choice heuristic.

Acknowledgements

I would like to acknowledge the guidance, assistance and help of a number of people who have played an instrumental role in the completion of this work. In particular, I thank Terry Elrod for his continual guidance since the inception of this thesis. He has often led me through deep tunnels of darkness and despair towards the bright light of knowledge. He has been both a friend and a mentor. I would also like to thank David Taylor for buying me one of the most powerful computing systems that has allowed me to run my immensely complicated simulations and programs, without which the task would have been impossible. Also, I would like to express my deep appreciation and gratitude for the support, understanding, friendship and guidance extended by Peter Popkowski. Finally, I would like to thank my wife Anna for being a perennial source of inspiration and motivation.

Table of Contents

List of Tables

Chapter One

Introduction

The last two decades of marketing research have seen an exponential growth of academic articles in the area of consumer choice models (Meyer and Kahn 1991). A casual perusal of any leading marketing journal will show that choice is the dominant research agenda in marketing. To emphasize the importance of choice models in marketing, Meyer and Kahn (1991) write:

> Since 1980, for example, there have been over 200 articles published on the subject in the literature of marketing, and this reflects a small fraction of the total literature that has appeared across disciplines. Indeed the work has proliferated to such a degree that the field of quantitative models of behavior is probably best thought of as a discipline in its own right. (p 85)

Various taxonomies have been proposed in the literature (Cortsjen and Gatushi 1983; and Mcfadden 1986) to understand the complex area of individual choice models - the most comprehensive of them being the one proposed by Meyer and Kahn (1991). They divide the area of choice models into two intellectual traditions (1) utility maximizing models, which assume that an individual makes decisions by considering all relevant information (Mcfadden 1986; Erdem 1993; and Guadagni and Little 1983), and (2) heuristic elimination models which view individuals as inherently limited in their ability to process information and, hence, make choices using simplified heuristics (Restle 1959; Tversky 1972; Tversky and Sattah 1979).

In this thesis, I borrow concepts from bounded rationality (Simon 1955) to describe these two streams of research. Substantive rational models emphasize their pre-occupation with the result of rational choice, whereas procedural rational models are concerned with the choice process. Central to this paradigm is the assumption of rationality as the product and process of thought rather than simply that of the outcome. This view of rationality allows individuals to be satisficers[1] (Simon 1955) rather than optimizers.

Most choice models in marketing can be described as substantive rational models. Recently, a few researchers (e.g. Shugan 1980), using cost-benefit analysis, have shown that the assumptions of substantive rational models are often violated. The inability of models, that assume computational limitations of a consumer, to capture dynamic behavior was a major stumbling block for research in this area. Attempts that have been made to capture dynamic behavior have found models to be computationally burdensome. Hence much of the growth in the area of procedural models has centered around static models.

A notable limitation of these static procedural models is the *a priori* assumption of the heuristic (Kahn, Moore and Glazer 1987) that will be used by consumer to make decisions. A consumer has a host of heuristics to choose from and it is not certain how consumers decide which rule to use to make decisions (Meyer and Kahn 1991). A common conjecture among researchers is that, as adaptive decision maker, a consumer would use cost-benefit analysis to make this decision (Bettman, Johnson and Payne 1991). A normative solution to this problem has been

[1] A "satisficer" is an individual who is ready to accept a satisfactory solution rather than the best solution. Simon (1955) argues that obtaining the best solution is beyond human computational abilities due to a number of cognitive constraints such as the limitations of short term memory. Due to these cognitive limitations consumers have to make decisions by applying some heroic approximations. This makes a consumer a satisficer.

provided by Johnson and Payne (1985), with normative solutions for optimal elimination strategies offered by Grether and Wilde (1984). There are two limitations to this viewpoint: (1) How do consumers make decisions in an uncertain environment where costs and benefits of using alternative rules are not known. (2) The cost-benefit analysis is not applicable to scanner panel data as the process of making brand choice is not observed by the researcher. The only information available to him is that of the brand bought by a consumer.

In this thesis, a different viewpoint than the one described in the paragraph above, is proposed. This work argues that only certain bits of information will become active in the working memory, and hence, will be used by consumers to make decisions. This approach can also be found in the works of Anderson (1976) and Smolesky (1986). It assumes information to be a luxury, and attention to be the scarce resource (Simon 1982). Both Simon (1982) and Anderson (1986) argued that though we are constantly bombarded with information the amount of information that an individual can attend to, due to computational limitations, is often limited. Attention is conceived of as being a very limited mental resource. Trying to emphasize the implication of this human limitation Anderson (1986) argued that

> ....all information gets into sensory memory, but to be retained, each unit of information must be attended to and transformed into some permanent form. Given that attention has limited capacity, all elements in the sensory cannot be attended to before they are lost. (p 53)

Consumer theorists such as Bettman (1979) have also emphasized the importance of attention in consumer choice decision making. Despite the importance of attention in consumer decision process, it has never been incorporated into formal models of choice. This research is an attempt to model the effect of attention on consumer choice. Common to the work of Anderson (1993), it argues that consumers try to optimize their actions in order to adapt to the local environment. In this work, the construct of attention is implemented as a probabilistic threshold activation function.

There are two advantages in using the activation function. First, the need for assumptions about the cost and benefit of different choice heuristics is circumvented. Hence using activation functions, procedural models, that assume computational limitations of consumers, can be implemented on normally available econometric (panel) data. Also, researchers in cognitive psychology (e. g. Rumelhart and McCleland 1986) have demonstrated the ability of the activation mechanism to capture different heuristics. Thus a single activation model may subsume several heuristics. A second advantage to using an activation function is that it provides a more plausible description of how consumers make choices in excessively rich environments. Extant procedural models, such as cost-benefit models, assume that consumers first choose a heuristic appropriate to the task before implementing it (Huber and Klein 1989). While this allows consumers to use a simpler brand choice stage, it also complicates the choice process by adding a decision making stage prior to the choice of a brand. The probabilistic activation function, can on the other hand, explain what types of information are more likely to be used in choosing a brand without recourse to an additional pre-choice decision task (Anderson 1983).

In brief, this dissertation (1) proposes a new taxonomy for choice models, (2) identifies the limitations of existing choice models, (3) introduces an explicit attention construct in a formal model of consumer choice, (4) uses the activation mechanism to unify disparate areas of consumer choice such as consideration sets, memory, expectations, perceived risk and choice strategy and, finally, (5) sheds insight upon the relationship among brand equity, brand awareness, loyalty and perceived risk by estimating and empirically testing these models on scanner panel data.

The remainder of the dissertation is organized as follows: in Chapter two I review the relevant literature. In Chapter three I provide a conceptual framework for the theoretical foundations proposed in this thesis. In Chapter four I use the conceptual framework developed in Chapter three to derive mathematical models for testing the theory. Chapter five provides a brief description of the method employed to test the theory. The data is described in Chapter six and the estimates are provided in Chapter seven. Finally, Chapter eight discusses the conclusion to be drawn from this work and identifies a program of further research.

Chapter Two

Literature Review

In this chapter I will provide the necessary background to the problem identified in the previous chapter by reviewing the relevant literature. Since my review is not limited to marketing alone, but draws on many different streams of research such as artificial intelligence, computing science, information processing theory and economics, it is not meant to be exhaustive. Sections that are pertinent to my research problem have been included in this chapter.

2.1 Substantive Vs Procedural Rationality

As my research problem is grounded in many different areas in marketing, I have used a set of common characteristics to compare contributions. Most models that are being reviewed in this chapter have been described in Table 1 in terms of six characteristics. These six characteristics are as follows: (1) consumer limitations; (2) consideration sets; (3) consumer expectations; (4) consumer heterogeneity; (5) nature of data; and (6) purchase decision.

The models differ the most in terms of the assumption they make about the computational limitation of the consumer. Using the terms which Simon (1978) proposed, I employ substantive rational for those models that view consumers to be utility maximizers and procedural rational models for those that assume them to have computational limitations.

Insert Table 1 about here

Table 2 shows the contribution of models based on the criterion of substantive and procedural rationality. Substantive rational models are limited to their preoccupation with the result of rational choice. They are grounded in the expected utility maximizing paradigm and assume that an individual makes decisions by considering all relevant information (Mcfadden 1986; Erdem 1993; and Guadagni and Little 1983).

Insert Table 2 about here

On the contrary, procedural rational models are concerned with the process of choice rather than with the outcome of choice. Central to this paradigm is the assumption of rationality as the product and process of thought. These models view an individual as a satisficer and not as an optimizer. In its extreme form (Currim, Meyer and Nhan 1988; Simon 1955; Anderson 1983; Newell 1990; and Grossberg and Gutowski 1987), it reject the existence of utility functions. By using differencing and differential equations, researchers (Anderson 1983 and Newell 1990; Grossberg and Gutowski 1987) have used computer programs that emulate human decision processing. The middle ground for extreme viewpoints of procedural and substantive rationality is found in the works of Bettman, Johnson and Payne (1991), Tversky and Kahnemann (1972), and Tversky and Sattah (1979). Though these models share the same ethos as those of heuristic models, they can at best be called a generalization of the expected utility models (Grossberg and Gutowski 1987).

Most models in marketing fall into the category of substantive rational models. Meyer and Kahn (1991) term them as simply scalable choice models that assume errors to be independent of the consideration sets. These models are described as variants of the original model proposed by Luce (1959), and implemented by Mcfadden (1973) as the multinomial logit model. Recently, Chintagunta (1993), Kamakura and Russell (1989), and Gonul and Srinivasan (1993) have shown that not accounting for individual level heterogeneity can bias parameter estimates. In addition, several researchers have accounted for dynamics of consumer choice (Meyer and Sathi 1985; Erdem 1993; and Keane 1995). Erdem (1993), in an important application of structural equations to consumer choice, showed that learning effect of attributes may cause a temporal or a structural state dependence in brand choice. Keane (1995) used a probit model to estimate a factor analytic covariance structure model and incorporates dynamics of choice by allowing errors to be correlated across time.

On the contrary, most of the models proposed in the area of bounded rationality have been dominated by Tversky's (1972) elimination by aspect (EBA) model. Process tracing studies have shown that the characterization of the choice process assumed by substantive models of choice (Bettman 1971; and Russo and Dosher 1983), even in highly simplistic settings, is often violated. The process can more accurately be described as a sequence of discrete elimination heuristics in which only limited information is used by a consumer to make decisions.

Unfortunately, not too many models have been developed that assume computational limitations of consumers. There are three limitations to this approach. The major drawback is the computational complexity involved in the estimation of the model i.e. even for moderate set size the number of possible combination for an

EBA model makes them virtually inestimable. Tversky and Sattah (1979) proposed a computationally tractable version of this choice heuristic (PRETREE) by imposing a prior known hierarchical structure of elimination. Researchers (Kahn, Moore and Glazer 1987; and Lehmann and Moore 1985) who have applied PRETREE to a marketing setting have identified certain problems such as the difficulties of identifying the correct decision tree, and the inability of models to capture individual level heterogeneity in decision rule within a sample.

The second limitation to this approach is the lack of understanding of how consumers decide how to decide. A common conjecture among researchers (Hogarth 1980; Payne, Bettman and Johnson 1988) is that consumers use an intuitive cost-benefit calculation to decide which heuristic to use i. e. having a host of heuristics to choose from, a consumer would select the rule that yields the highest expected outcome at the lowest cognitive effort. Normative treatments of this problem have been provided by Johnson and Payne (1985) and Shugan (1980) and normative solutions to this problem provided by Grether and Wilde (1984) and Huber and Klein (1989). Much of the work in this area has been carried out in stable controlled environments. Virtually little or no work has been done to identify the effect of a dynamic environment on changing choice strategies (Meyer and Kahn 1991). Applying heuristic rules to normally available econometric scanner panel data poses insurmountable problems for the researchers as there are a multitude of different possible heuristics being used by consumers which might vary across consumers and time periods. Therefore, with the exception of Andrew and Srinivasan (1995), no known study has applied these models to panel data. The intractability of the static models is compounded when trying to capture dynamics of choice over time (Andrew and Srinivasan 1995). Meyer and Kahn (1991) have identified this area as one that has not received enough attention due to the lack of tools that can be used to implement these models on normally available econometric data.

The third limitation to this approach, probably the most fundamental of all, is the use of parsimony for deciding the most plausible hierarchical sequential decision process. Anderson (1993), though not the first theorist, claims that although behavioral data can be used to study the steps of mind at the algorithmic level (Bettman 1971; and Currim, Meyer and Le (1988)), they lack identification at the implementation level. Anderson (1993) argues that proceduralists are in search of a function that can map input to output, and that there are innumerable possible functions. Though parsimony can be used to settle this dispute, it stretches credulity beyond reasonable bounds to assume that nature chose the most parsimonious design for the human mind.

Though attempts have been made and normative models have been developed to explain the effects of computational limitations of the consumer on brand choice, they are largely at an embryonic stage. Even in the simplistic possible situation, where only a few brands exist in the environment, the set of possible combinations of the brands can make the problem intractable (Meyer and Kahn 1991), particularly for models that try to explain choice over time.

2.2. Consideration Sets

The second set of characteristics on which models differ is in their ability to capture the phenomenon of consideration sets. The fact that consumers, while making decisions, do not consider all available information is not new to marketing (Andrew and Srinivasan 1995). Of the brands that consumers are aware of, only a few of them will be considered at any given time (Shocker et al. 1991). This concept of consideration sets, or the set of alternatives being considered by a consumer at any given time, has been of interest to marketers (Roberts and Lattin 1991; Hauser and Wernerfelt 1990; Nendungadi 1990). Theoretical rationale for the concept of consideration sets can be found in both economics and psychology. Borrowing arguments from information economics, research in the area of consideration sets is based on the premise that a consumer will continue to search for information as long as the expected returns from information search exceeds the marginal cost of future search. Hauser and Wernerfelt (1990) and Ratchford (1980) provide a normative treatment for this problem, while Roberts and Lattin (1991) applied this construct on data for the choice of breakfast cereal.

Despite recent interest in the idea of consideration sets, dynamic models of consideration set theory are sparse. An obstacle of research in this area is the inability of models to specify consideration sets at the individual level and across time periods. Andrew and Srinivasan (1995), in an attempt to study the dynamics of the consideration set, estimate a probabilistic model on panel data and conclude from their tests that their model does better than a multinomial logit model. The success of this model comes at a great price as the computational complexity increases exponentially with the increase in the size of the consideration set. To simplify the computational requirements of the models, Andrew and Srinivasan (1995) kept the analysis to the brand level and within brand considerations, such as size, were not investigated.

2.3 Expectations as a Means of Capturing the Process of Choice.

The third characteristic on which models differ is the incorporation of expectations in models of choice. Ever since Kahneman and Tversky (1979) proposed an alternative to the frame of "reference", the concept of reference price has been used to explain the effects of promotion and price reduction on brand switching (Winer 1986; and Winer 1985). The underlying assumption in this literature is that positive value of (p^o-p^r) is perceived negatively, while negative values of (p^o-p^r) are viewed positively, where p^o is the observed retail price and p^r is the individual's internal reference price,. Reference price (p^r) is defined as the price used by consumers to evaluate the price for alternatives available at any given time. It can be broadly defined as internal or external reference price. External reference price is one that exists in the environment, for example a price display, used by consumers to assess the value of an item. On the other hand, internal reference price is stored in a consumer's memory but may also serve to evaluate external reference price. Various types of reference price which have been suggested in the literature include aspiration, market and historical prices (Klein and Oglethorpe 1987); lowest and highest prices (Monroe 1990); fair price (Thaler 1985); average market price (Emory 1970); expected future price (Jacobson and Obermiller 1990); and lowest market price (Biswas and Blair 1991).

Expectation models have dominated the economics literature for the last two decades. However, the applications of these models to marketing have been sparse. Russell Winer has played an important role in introducing the ideas of rational expectations to marketing. In his pioneering work Winer (1985) proposed that future expectations were an important component of consumer decision making. He believed that the future expectation of price of a consumer durable depends on the past reference prices and expected prices, general economic conditions, both current and anticipated price expectations, future price signals and household specific variables. Due to inadequate data, a simplified model had to be estimated and the results provided preliminary support for the more general model. In his subsequent work Winer (1986) used a rational expectations formulation to test reference price theory, though the crucial element of forward expectations was missing from the model.

Jacobson and Obermiller (1990) were the first to test rational expectations. They argued that not incorporating future price as an important dimension of reference pricing was a violation of neoclassical economic theory. In an experimental study, explicit measures of future expectations were obtained over a period of eight weeks and tested for unbiasedness and efficiency requirement of the rational expectations paradigm. The rational expectations hypothesis was rejected at the aggregate level but the serial correlation model, which accounted for the influence of unobserved variables, was most consistent with the data.

Kalwani et al. (1990) developed and calibrated a price expectations model of consumer brand choice. A two-stage model was used to estimate and study (1) the formation of expectations, and (2) the effect of this price expectation on consumer brand choice. The important distinction they made between expectations and reference prices was that the latter was a weighted function of past prices while the former was not only a function of past prices but also a function of other economic variables. The expectation of price of a brand was believed to be affected by the frequency of promotions, deal proneness of consumers, past prices, and market trends. Consistent with reference price theories, the asymmetry about price expectations was also obtained.

Kalwani and Yim (1992) in a controlled experiment studied the impact of price promotions on consumers' price expectations. They investigate the effects of price promotion depth and frequency on price expectations and tested the effects of price promotions on brand choice. In a study similar to that of Jacobson and Obermiller (1990), they elicited price expectations directly from the respondents and did not use surrogate variables as a proxy for the latent expectation construct. The results of the experimental study were consistent with the findings of Kalyanaraman and Little (1989) as they found that both frequency and size of the discount had a significant impact on a brand's expected price.

Though internal reference price has been incorporated in many applications of reference price theories to scanner panel data, external reference prices, which also play an important role in current price expectation, have largely been ignored, for example Kalwani et al. (1990) and Lattin and Bucklin (1989). In contrast, most experimental studies related to reference price have focused on the effects of external reference price. In fact, only one experimental study has examined the effects of internal reference price (Biswas, Wilson and Licata 1993). There is considerable support for the notion that consumers' current expectations of price are affected by internal and external reference price (Biswas and Blair 1991). Biswas and Blair (1991) showed that advertising and the store for which the external reference price was being advertised could affect current expectations of price. Besides this, Urbany, Bearden and Weilbaker (1988) showed that advertisements that have a plausible reference price raised subjects' estimates of the advertiser's regular price and the perceived offered price. In general, there is strong support for the effects of contextual variables such as store type, brand familiarity and advertisements on a consumer's current expectation of price.

Consistent with the view of rational expectations, the current expectations of price (combination of the external and internal reference price) are formed by incorporating all available information. Urbany, Bearden and Weilbaker (1988) proposed that a consumer first evaluates the credibility of the external reference price and then either assimilates it, causing a shift in the internal reference price towards the external reference price, or completely rejects the external reference price resulting in no change of the internal reference price. Biswas and Blair (1991) propose that the effect of the external reference price is in two dimensions: direction and magnitude. This will be dependent upon: (1) a difference between the external reference price and internal reference price; (2) consumers' confidence in his prior beliefs, and; (3) the credibility of the external reference price.

A major limitation of this stream of research is its inability to capture the uncertainty in the environment. The need to form expectations arises primarily because consumers, in deciding which brand to choose from, are faced with an uncertain future environment. They can only speculate about the future and form expectations about the future course of a variable, for example that of price. Therefore, future expectations play an important role in most consumer decisions. Consumers are concerned with the implications of their current actions on their future (Oliver and Winer 1986). Katona (1960), who spent a large portion of his career researching in the area of future expectations of a consumer, believed that though a consumer's ability to pay is important to purchase decision making, willingness to pay is also important. He believed that consumers' willingness to pay depends on their expectations of the future economic environment.

Most choice models, which have accounted for the effects of uncertainty of product characteristics and imperfect information, have not incorporated the mediating roles of price expectations on consumer choice decisions. Notable exceptions are the work of Meyer and Assuncao (1990) and Krishna (1992). Meyer and Assuncao (1990), in an experimental setting, studied the effects of future expectations of price on brand choice decisions. From their results they concluded that subjects, even when provided with future price distribution, did not use this information optimally according to the dynamic programming algorithm. Krishna (1992) used a variant of the Golabi model to study the effects of future price expectations on brand choice and stock piling. Using a Monte Carlo market simulation, she concluded that both brand choice and stock piling behavior depend on not only the future expectations of price but also on the future price of the competing brand.

Other choice models that have incorporated uncertainty of product characteristics (Meyer and Sathi 1985; Eckstein, Horsky and Raban 1988; Roberts and Urban 1988; and Erdem 1993) have repeatedly shown that the learning effect of these attributes may cause a temporal or a structural state dependence in brand choice. With the exception of Eckstein, Horsky and Raban (1988) and Erdem (1993), dynamic brand choice models have tended to be backward looking and have not incorporated the effects of consumer search on choice dynamics. Erdem (1993) in a structural model framework shows the interdependencies in consumer risk, information search, brand choice, brand values and a firm's marketing mix decisions under imperfect information. As equations of the model are structural or behavioral, the parameter estimates are policy invariant and allow her to carry out certain policy experiments.

2.4 Consumer Heterogeneity

The fourth characteristic on which models differ is unobserved heterogeneity. Heterogeneity attempts to account for the difference in behavior due to observed and unobserved variables. Observed heterogeneity can be a result of the effect of income, family size and geographical location on choice. However, there are variations across consumers that cannot be observed by researchers. These variations are known as unobserved heterogeneity.

Heterogeneity across households has been characterized in different ways in the literature. Gupta (1988), Guadagni and Little (1983), Krishnamurthi and Raj (1988), Chiang (1991), and Bucklin and Lattin (1991) used prior household history and income to account for heterogeneity. Though in the literature this implementation has been defined as observed heterogeneity, this implementation has two limitations. First, incorporating heterogeneity based on prior household behavior can confound loyalty and heterogeneity and these two different aspects of consumer behavior cannot be disentangled (Keane 1995). Secondly, as researchers often have to make inferences about certain unobservable variables, for example, inventory (Gupta 1988; and Bucklin and Lattin 1991), there is uncertainty around the inferences which are required to be integrated out of the likelihood function. To the best of my knowledge this has never been done in marketing.

Unobserved heterogeneity has been incorporated into models of choice in three different ways. The simplest way of handling heterogeneity is to estimate a fixed effect model (Jones and Landwehr 1988). A major limitation to this approach is the lack of availability of a large number of observations at the individual level. The non-availability of data at the individual level can cause substantial estimation error that can lead to overstating the true population heterogeneity (Elrod 1991).

A second way of incorporating unobserved heterogeneity into models of choice is to specify a functional form for the distribution of heterogeneity. The parameters of the underlying distribution for heterogeneity can be estimated directly from the data using empirical bayes (EB) or bayesian techniques (Elrod 1988; Keane 1995; Gonul and Srinivasan 1993; Elrod and Keane 1995; and Allenby and Lenk 1994). Allenby and Lenk (1994) use a hierarchical bayesian method, while Keane (1995) and Elrod and Keane (1995) use an empirical bayes method, to estimate the parameters of the model. Generally, monte carlo integration is used for estimating these models.

A major limitation of EB models is that a large amount of data at the individual level is required to estimate parameters for the distribution of heterogeneity and for model convergence. Though Gibbs-sampling techniques mitigate this limitation of the EB models, they cannot be used to estimate nonlinear models.

A third way of incorporating heterogeneity into formal models of choice is by using a latent class approach (Kamakura and Russell 1989 and Chintagunta 1993). In a latent class approach the population of consumers is assumed to consist of a finite number of segments, each segment having its own parameters. Often, fit statistics, such as AIC (Akaike Information Criterion) or BIC (Bayesian Information Criterion), are used to determine the number of segments in the market.

Of the three methods, latent class analysis is often considered the most robust as the mis-specification of the underlying distribution for heterogeneity will not bias the results. Although this is true, latent class models often understate variability of the true population (Elrod 1991).

2.5 Nature of Data and Purchase Decision

The last two criteria on which the models differ are the nature of the data and the purchase decision being studied. With the exception of Krishna (1992), most marketing studies can broadly be classified into two groups. The first group concerns models that are analyzed on single source scanner panel data (Chintagunta 1993; Keane 1995; and Erdem 1993). The second group concerns models that use experimental data (Assuncao and Meyer 1990). Single source data, though widely used in marketing, has many limitations (Koslow 1990), four of which are discussed in the paragraphs below. A first limitation of the scanner panel data is that as the external variables assumed to affect choice are not in the control of researchers, the effect of highly collinear variables, for example price reductions and promotions, cannot be disentangled. A second limitation of scanner panel data research is that often researchers infer variables (e.g. inventory level, reference price, brand loyalty) without accounting for the measurement errors of their inferencing (Bucklin and Lattin 1991). No accounting for measurement errors can lead to biased and inconsistent parameter estimates. Third, Koslow (1990) has argued and shown that aggregate model of choice (Guadagni and Little 1983; and Gupta 1988) cannot be used to detect casuality using panel data. And fourth, most studies use data that pertain to only those store visits in which purchases are made in the product category of interest (Guadagni and Little 1983; and Kamakura and Russell 1989). This has the effect of overstating the effect of marketing mix variable (Chintagunta 1993) as the analysis ignores store visits on which purchase was not made. Some studies (Gupta 1988, 1991) have assumed that consumers visit the store every week, thereby imputing store visits using only purchase data. The availability of single source data has shown that the assumption of consumer visiting a store only once a week is often violated.

Subject responses obtained using experimental design is often considered as a better means of understanding casuality (Koslow 1990), as the external environment is largely under the control of the researcher. A large part of research in marketing is based on data from experimental studies (Meyer and Sathi 1985; Assuncao and Meyer 1990; Jacobson and Obermiller 1990; and Kalwani and Yim 1992). Problems with this form of data have largely been under-acknowledged in the literature. McCleland (1955) argued that most studies conducted in closed static experimental conditions are often divorced from the environment in which the decision is being made. Although experimental study might help in understanding casuality, it lacks external validity. The lack of external validity poses a problem for marketers who are interested in identifying variables that affect consumer choice in the purchase environment. A researcher understands the effect of certain variables on behavior by making subjects attend to a few variables that a researcher feels are relevant to his study and by controlling or eliminating the effects of others. As the real environment provides the consumer with a multitude of cues, an important question that remains to be answered is whether the variables under study will have the same effect on consumer choice as what is predicted by experimental studies. The lack of external validity is a major limitation of experimental studies (Newell and Simon 1972).

The next section proposes a framework which draws on the various streams of literature previously discussed. The models described in the next chapter link the following concepts: (1) attention as a scarce resource with computational limitations of consumers; (2) purchase incidence; and (3) brand choice decision. These three concepts form a framework that explains the process of choice. In particular this framework attempts to incorporate the effect of price expectation on choice.

Table 1

Categorization of Consumer Choice Models

Reference	Decision Studied	Nature of Data	Unobserved Heterogeneity	Consumer Limitations	Consideration Sets	Expectations
Guadagni and Little (1983)	Brand Choice	Panel	No	No	No	No
Neslin, Henderson and Quelch (1985)	P. Timing P. Quantity	Panel	No	No	No	No
Krishnamurthi and Raj (1988)	Brand Choice P. Quantity	Panel	No	No	No	No
Tellis (1988)	Brand Choice P. Quantity	Panel	No	No	No	No
Jones and Landwehr (1988)	Brand Choice	Panel	Yes	No	No	No

Table 1 (Cont'd)

Categorization of Consumer Choice Models

Reference	Decision Studied	Nature of Data	Unobserved Heterogeneity	Consumer Limitations	Consideration Sets	Expectations
Gupta (1988)	Brand Choice P. Quantity P. Timing	Panel	No	No	No	No
Gupta (1991)	P. Timing	Panel	Yes	No	No	No
Bucklin and Lattin (1991)	Brand Choice P. Incidence	Panel	No	No	No	No
Chiang (1991)	Brand Choice P. Quantity	Panel	No	No	No	No
Kamakura and Russell (1989)	Brand Choice	Panel	Yes	No	No	No
Helsen and Schmittlein (1990)	P. Timing	Panel	No	No	No	No

Table 1 (Cont'd)

Categorization of Consumer Choice Models

Reference	Decision Studied	Nature of Data	Unobserved Heterogeneity	Consumer Limitations	Consideration Sets	Expectations
Erdem (1993)	Brand Choice P. Incidence	Panel	No	No	Yes	Yes
Andrew and Srinivasan (1995)	Brand Choice	Panel	No	Yes	Yes	No
Winer (1986)	Brand Choice	Panel	No	No	No	Yes
Jacobson and Obermiller (1990)	—————	Quasi-Experimental	No	No	No	Yes
Kalwani et. al. (1990)	Brand Choice	Panel	No	No	No	Yes
Kalwani and Yim (1992)	—————	Experimental	No	No	No	Yes

Table 1 (Cont'd)

Categorization of Consumer Choice Models

Reference	Decision Studied	Nature of Data	Unobserved Heterogeneity	Consumer Limitations	Consideration Sets	Expectations
Chintagunta (1993)	Brand Choice P. Incidence P. Quantity	Panel	Yes	No	No	Yes
Meyer and Assuncao (19900	———	Experimental	No	No	No	Yes
Krishna (1992)	———	Simulation	No	No	No	Yes
Lattin and Bucklin (1989)	Brand Choice	Panel	No	No	No	Yes
Sinha (1996)	Brand Choice P. Incidence	Panel	Yes	Yes	Yes	Yes

TABLE 2

Substantive Vs Procedural Models

	Substantive Rational	Procedural Rational	
		Expected Utility Models	Extreme Version
Static	Bass (1974) Bass, Jeuland and Wright (1976) Bass and Pilon (1980)	Kahneman and Tversky (1979) Tversky (1973) Tversky and Sattah (1979) Kahn, Moore and Glazer (1987) Moore, Lehmann and Pessemier (1986) Lehmann and Moore (1975)	Currim, Meyer and Le (1988)
Dynamic	Erdem (1993) Keane (1995) Guadagni and Little (1983) Krishna (1992)	Sinha (1996)	No Existing Models

Chapter 3

Consumer as Limited Information Processor: An Activation Approach

In this section the conceptual framework, required to develop formal models of choice grounded in the literature of bounded rationality, is outlined. In particular my endeavor will be in providing a framework that can be used to first, distinguish between psychological and technological limitation, and second, capture the psychological limitation of the consumer.

A major assumption of this research concerns the computational limitations of the consumer (Simon 1955). Bettman (1979) and Payne, Bettman and Johnson (1988) viewed consumers as limited information processors who are incapable of making optimal decisions, as assumed by economic models of choice. Even in the simplest of choice situations, axioms of normative utility theory are often violated at the individual level (Bettman, Johnson and Payne 1990). Though various attempts have been made to incorporate the limitation of the consumer, they can broadly be classified into two categories: (1) mind as a scarce resource (Payne, Bettman and Johnson 1988), and (2) information as a positive good[2] (Roberts and Lattin 1991; Hauser and Wernerfelt 1990).

One way of capturing the concepts of bounded rationality is to view the mental processing capacity as a scarce resource (Payne, Bettman and Johnson 1988). Payne, Bettman and Johnson (1988) show that consumers will optimize on the amount of processing that they will indulge in. They argue that as global optimization is beyond the realms of most consumers, these consumers will use simple rules of thumb or heuristics to make decisions. In addition, these researchers claim that the

[2] A good is a "positive" good if its acquisition is beneficial to the individual.

characteristic of the problem environment, person, and social context will influence the heuristic used by the consumer to make a decision.

Another way of capturing the concepts of bounded rationality is to assume that in a world of imperfect information an individual has to search for alternative courses of action (brands for consumers). Stigler (1961) argued, using an example of the information search for a second hand automobile, that an individual will search till the point where the marginal utility of search is equal to the marginal cost of search. Researchers in choice models (Hauser and Wernerfelt 1990; and Roberts and Lattin 1991) have used this concept to bring human bounded rationality within the compass of rational optimization.

There are two limitations to the models described above. The next section details these limitations.

3.1 Computational Limitation as a Technological Limitation

Both these methods (mind as a scarce resource and information as a positive good) are at odds with the nature of the concept that they try to capture. Rather than simplifying the problem at hand, the complexity of the problem is compounded by the introduction of these mechanisms. Not only do consumers have to choose a brand but also they have to make decisions about the amount of search or the heuristic that has to be used to choose a brand (Simon 1978). Rather than simplifying the choice problem by assuming sub-goal identification (Simon 1978) and satisficing (Simon 1955), which is the intent of bounded rationality, these methods add to the computational complexity, making the problem more difficult than before.

Criticizing the search theories proposed in economics (Stigler 1961), Simon (1978) writes

> Limits and costs of information are introduced, not as psychological characteristic of the decision maker, but as part of his technological environment. Hence, the new theories do nothing to alleviate the computational complexities facing the decision maker-do not see him coping with them by heroic approximation, simplifying and satisficing, but simply magnify and multiply them. Now he needs to compute not merely the shape of his supply and demand curves, but, in addition, the costs and benefits of computing those shapes with greater accuracy as well. (p 485)

This criticism, levelled at economics of information, can also be applied to mind as a scarce resource. Scarcity of mind should not be viewed as an optimization problem, with the assumption that individuals are using intuitive cost/benefit analysis to decide the heuristic to be used, but as psychological limitations that forces individuals to use heuristics. Unfortunately, viewing scarcity of mental computing as a psychological limitation poses insurmountable problems of identification (Anderson 1993). There are a host of input-output functions that can capture the same process with no test for identifying the correct heuristic.

3.2 Information as a Scarce Resource

The literature on search and information transfer in economics (Stigler 1961) and marketing (Hauser and Wernerfelt 1990) views information as positive good. Information is assumed to be a scarce resource and, hence, has to be optimally searched for. A number of researchers (Hogarth 1980; Bettman 1979; and Simon 1982) provide empirical examples to show that this assumption is often violated. In an information scarce world, information is a positive good. On the contrary, in an information rich world, information is a luxury which sometimes might direct attention away from what is important (Simon 1978). In an empirical study of individual choice, Kunreuther (1978) showed that the best predictor of choice of flood insurance was not the constituents of the utility maximization equation but that of focus of attention. Neoclassical economics assume that an individual will use cost/benefit analysis to make these decision. The empirical data are at odds with this assumption as it appears that the decision of purchase by individuals was made on the basis of prior experience, more or less independent of the cost/benefit thesis. Finally, Van Raaij's (1977) eye movement studies of consumer choice indicate that consumers select some pieces of information and ignore others.

The view that attention is a scarce resource is not new to marketing and is found in the works of Bettman (1979) and Van Raaij (1977). Hogarth (1980) writes that a consumer can perceive only 1/70 of what is present in his visual field. Since consumers are completely inundated with information, their computational abilities only allows them to focus on a few bits of information at any given time (Simon 1982). Borrowing concepts from Theil (1954), Simon (1982) has provided a normative solution to this problem. A major limitation to the framework provided by Simon (1982) is the impossibility of applying this approach to empirical data as the measures required for implementing this analytical technique is not available (Simon 1982). Secondly, this approach belies the very nature of the concept, bounded rationality, that a researcher sets out to incorporate into formal models of choice. Not only do consumers now have to make brand choice decisions, but also they have to make decisions about allocation of their scarce resource. Hence the problem is now more complex than before.

3.3 Limited Computational Ability as a Psychological Limitation

Attempts that have been made to capture the psychological limitation of human decision making fall into two categories: (1) modeling the limitations of the short term memory (Newell and Simon 1972); and (2) using probability functions to activate only a few bits of information at any time (Anderson 1976; and Anderson 1983).

Newell and Simon (1972) allow only seven chunks of information to exist in the short term memory at any given time. Having seven small bits of information in the short term memory does not allow researchers to simulate complex human behavior (Anderson 1976). A way around this problem is to use the system of chunking. Chunks are learned configurations of symbols which act as a single symbol. This view of short term memory being able to retain seven chunks of information was popularized by Miller (1956) and has extensively been used by Newell (1991) and Newell and Simon (1972). The advantage of using chunks of information is that long strings of symbols will only occupy one slot in the short term memory. So, for example, while we can hold only seven random letters in the short term memory, we can hold seven ten letter words in our memory.

It is apparent from the discussion in the previous paragraph that a major limitation of modeling the computational limitation in this manner is the arbitrary definition of "chunks of information". Though chunking allows researchers to simulate complex human behavior, it contradicts the notion of modeling the limitation of short term memory (Anderson 1976) as it endows consumers with information that might be well beyond their ability.

A second approach found in the literature is the use of the activation function to capture the limitation of short term memory (Anderson 1976). Using a system of differential equations, Anderson (1983; 1976) models computational limitation of human decision processing by making certain bits of information more active than others. Hence, there is a higher likelihood of the active information being used in making a particular decisions. At a conceptual level the framework provided by Anderson (1983) differs from Newell and Simon (1972) as it assumes parallel distributed processing. In addition, this framework did not view memory to be compartmentalized into short and long term memory. Anderson (1983) assumes that the portion of long term memory that is active at any given time to be the working memory.

3.4 Activation as a Means of Capturing Attention and Psychological Limitation

Perception of information is not comprehensive but selective (Hogarth 1980). One of the key factors in human intelligence is the ability to identify and utilize information that is relevant to a particular problem (Anderson 1983). Activation plays a major role in that facility. A piece of information or knowledge will become active to the extent that it is related to the current decision being made. Thus, activation identifies and favors the processing of information that is most pertinent to the immediate context.

Activation measures the likelihood that a particular piece of information will be useful at any particular time. It can be viewed as a heuristic that tells individuals the relevance of information. For example, when an individual is making a decision about cars, attributes of orange juice that are important to him will rarely come to mind as he knows that the factors which influence his decision about which juice to buy are not related to the factors which influence his decision about which car to buy.

Though the principal of association can be traced back to Aristotle (Anderson and Bower 1973), the work of Quillian (1969) was important to the resurgence of work on spreading activation, particularly through the models and theories of parallel distributed processing (Rumelhart and McCleland 1986; Grossberg and Gutowski 1987; and Anderson 1983). Quillian (1969) argued that this mechanism eliminates or reduces the costly knowledge search processes that can be the pitfall of any artificial intelligent system with a large database. In addition, Anderson (1983) claimed that it is relatively cheap for the brain to spread activation but expensive for it to perform symbolic manipulations.

However, in order to select information, it is necessary for the decision maker to know what to select. Anticipation plays an important role in this selection process as physical and motivational reasons account for most of this selective procedure (Hogarth 1980). Bruner (1957) showed that the more complex or ambiguous the stimuli is, the more the perception is determined by what is "in" the subject rather than what is "in" the stimulus. Simon (1976) argued that both motivational and cognitive mechanisms mingle in the selective process. Selective attention to a part of stimuli may reflect: (1) deliberate ignoring of other stimuli as not relevant to the goals of the mechanism; and (2) a learned response stemming from the past history. A serious gap in much of the theorizing is that little or no work has been done in developing models that can be implemented on empirical data.

Anderson (1993) showed that, under certain assumptions of independence, observed data can be used to predict pieces of information most likely to be used by consumers to make decisions. Using principles of adaptive rationality and associations, he argued that history and cues in the external environment will determine the use of a certain piece of information for making decisions. History manifests itself in the frequency and recency of use, while contextual and environmental variables are external stimuli that are provided to the decision maker.

In the next chapter I propose a modeling framework that allows me to capture the different aspects of the theory proposed in this chapter. Not only does it allow me to capture the psychological limitation but also the information that a consumer is attending to.

Chapter Four

Model Development

In this section the dynamics of consumer choice is modeled by considering the impact of brand familiarity, loyalty, external cues and brand attributes on choice decision. The formal models of choice developed in this chapter are grounded in the theory of bounded rationality and attempt to capture the computational limitation of the consumer as a psychological and not a technological limitation. Borrowing modeling concepts from Anderson (1983) and Rumelhart and McCleland (1986) these models incorporate attention as an activation function.

Like the models in recent marketing literature (Thaler 1985; Hardie, Johnson and Fader 1993), the models proposed here attempt to understand the process rather than that of the outcome of choice. As the models try to capture the inner mechanism of the system in the form of activation mechanism and psychological limitation, it represents a framework that is characteristic of structural models of choice (Simon 1982). It is important to note that the definition of "structuralism" used in this work is different from the one that is generally used by researchers in marketing (Erdem 1993) and in economics (Rust 1987). An important property of structural models, as defined by Erdem (1993) and Rust (1987), is that these models treat uncertainty and time explicitly. The structural models are based on consumer utility maximization and their parameters are parameters of consumer utility functions and constraints. The decision maker is assumed to have a well defined objective function, which is dependent upon both information set, future expectations and exogenous variables. This objective function allows him to make sequential decisions.

A notable limitation of the framework provided by Erdem (1993) is that it endows human beings with information and computational abilities far beyond what their mental faculties would ever allow (Simon 1982). In addition, these models capture the structure of the decision environment and not that of the decision maker. It is only when a model captures the inner limitation of the consumer that it can be called a structural model (Simon 1982).

Unlike the framework described in the paragraphs above, Gould (1980) and Margolis (1987) view "structuralism" as a concept that attempts to understand and model the inner limitation of the system. It can be characterized as a framework that captures and tries to understand the inner mechanism of the mind of the consumer (Simon 1978; Gould 1980; Margolis 1987). Newell and Simon (1972) argue that to the extent that the behavior of a consumer is precisely what is called for by the situation at hand, it will give us information about the task environment. It is only when behavior departs from rationality that one learn about the inner mechanism of the system.

A major focus of the models that I propose in this section is that they distinguish between the parameters of the inner mechanism and that of the task environment. These models nest the rational models within models of bounded/ adaptive rationality. As models of rational behavior assume that all information available to a consumer will be used to make decisions, rational models can be specified as limiting cases of the adaptive model. This is an important advantage of the framework as it allows the use of fit criteria e.g. BIC (Bayesian Information Criteria) and likelihood tests for comparing competing models.[3]

[3]Some researchers might argue that the models being proposed in this section are not capturing bounded rationality as their estimation is based on the maximization of the derived utility functions. Though this is a limitation of the framework, it is also its strength. A

4.1 Model Assumptions

The models developed in this section are based on the following assumptions:

1. A consumer's utility function for a brand can be approximated by a compensatory model.

2. There is a higher likelihood of pertinent information being active in the working memory.

3. Anticipation plays an important role in a brand's utility function.

4. Consumers are assumed to purchase only one bottle of ketchup on any given purchase occasion.

5. Consumers can obtain information about competing brands at no cost. Hence, consumers are assumed to be making decisions among the brands available in one store at any particular time.

6. Consumers are assumed to be exposed to brand promotional activity in every time period.

4.2 Adaptive Rational Choice Model.

In this section the dynamics of the consumer choice are modeled by considering the framework provided in chapter three. For easy exposition, the entire model has been divided into three sub-sections:

(1) Deep Parameters - parameters that are used to model the psychological limits of the consumer.

---

limitation of the extreme form of bounded rationality (Anderson 1983) is that the system is under identified, that is, there are a number of different input-output functions with no one universal measure of testing the competing theories and mechanism. By viewing consumers to be adaptive rational (Anderson 1993), the researcher allows some structure in decision theory, thus making theory testing much simpler.

(2) Expectations - parameters that are used to model price expectations.

(3) Task environment- parameters that capture the effect of advertising, promotion and price.

4.2.1 Consumer Expected Utility

As the parameters of the utility function are used as variables in the activation function and price expectations, I will first define the most general functional form of the utility function and define the indexes used in the model. The functional form for the most general utility function can be expressed as:

$$\begin{aligned} U_{sijt} &= \alpha_{oijt} + B_{spijt}\alpha_p p_{jt} + B_{spijt}\alpha_R (p^e_{ijt} - p_{jt}) \\ &+ (1 - B_{spijt})\alpha_P p^e_{ijt} + B_{saijt}\alpha_a a_{jt} + \alpha_{std} d_{ijt^*-1} + \varepsilon_{ijt} \end{aligned} \tag{4.1}$$

where,

d_{ijt} =1, if consumer i choose j-th brand in the t-th time period

=0, otherwise

ε_{ijt} is assumed to have a double extreme Gumble distribution.

α's are the parameters to be estimated,

p is Price,

a is Advertising,

i.....indexes for the consumer (i=1,......n),

j....indexes for the brand (j=1,.....J), t.....indexes for the time period of store visits.[4] (t=1,.....T),

t*...indexes for store visits on which one of the brands was bought,

k.... indexes for attributes, price and advertising, (k=1,.....K)

s....indexes for possible state space, that is, the information that could have been used by consumers to make decisions, (s=1,.....S)

$B_{spijt} = 0$ if the price of the j-th brand is not being attended to by consumer i in the t-th time period given that the consumer is in the s-th state.

$= 1$ otherwise,

$B_{saijt} = 0$ if consumer i is not attending to the advertisement of the j-th brand in the t-th time period given that the consumer is in the s-th state.

$= 1$ otherwise,

and,

p^e_{ijt} is the price expectations of the i-th consumer for the j-th brand at the t-th time period.

The utility function defined in equation 4.1 warrants an explanation. This equation specifies that consumers are not using all available information to make decisions in each time period. This is captured by the activation function (B_{spijt} and B_{saijt}), which can take a value of 1, if the attribute of a brand is being attended to and conversely 0, if it is not. An implicit assumption of the utility specification is that advertising has no impact if it is not being attended to. The assumption does not hold

[4] I use time periods and store visits interchangeably.

for price as consumers can infer prices from prior experience and use this inference for making decisions. Also, the utility equation assumes that both absolute and relative price have an impact on consumers' utility. If consumers are attending to price then not only the absolute but also the relative price is having an impact on utility. On the other hand, if consumers are not attending to price, then they are making inferences about price (not assuming it be 0) and using this inference to make decisions.

4.2.2 Psychological Limitations of the Consumer

In the structural model the consumer psychological limitations are modeled as two separate constructs. One tries to distinguish between action and inaction, and the other captures the limited information processing ability of the consumer.

Discriminating between Action and Inaction:- A major limitation of utility maximization theory is its inability to distinguish between "action" and "inaction". The consumer is pictured as always in a state of action. No action is a particular way of doing something not distinguishable from other forms of action (Simon 1977). Hence, in the dynamic case, the observation of a consumer not making a purchase in the product category may in turn be due to two factors: (1) consumers decide not to make a decision and hence, do not buy a brand in the market; or (2) consumers decide to make a decision, processes information, and then decide not to buy one of the brands in the market. Most models in marketing (Chintagunta 1993; Gupta 1988; Guadagni and Little 1983; Erdem 1993; and Keane 1995) belong to either of these two characterizations. Guadagni and Little (1983) and Keane (1995) estimate a brand choice model assuming that the store visits when no brand was bought in the product category are irrelevant and should not be a part of the data analyzed to estimate the models. However, Erdem (1993) and Chintagunta (1993) assume that consumers are always in a state of action i. e. the choice of not buying a brand in the market is no different from choosing a brand in the market.

The opposing approaches manifested in these two models represent extreme viewpoints. The consumer decision process lies somewhere in the middle. Hence, on a given store visit on which no purchase was made in the product category, a consumer could have been in either of these two states. This requires the likelihood function to be evaluated at both these states. A uniform prior is assumed for the two states, that is, the probability that consumer i is making a decision on the t-th store visit on which no brand was bought is given as follows:

$$P(\Delta_{it} = 1) = .5 \tag{4.2.1}$$

where,

$\Delta_{it} = 1$, if consumer i is making a decision on the t-th store visit,

$= 0$, otherwise.

Consumer Information Processing :-At one extreme, it can be assumed that a consumer will always use all available information to make decisions. Hence, for this state

$$B_{skijt} = 1, \text{ for all k, j and s,}$$

where K attributes are used by consumers to make decisions. Variants of this specification are found in the works of Guadagni and Little (1983), Chintagunta

(1993), and Keane (1995). Another extreme is to assume that only certain bits of information about the attributes of a brand will be used by consumers to make decisions. For example, a specification of the sort given below:

$$B_{saijt} = 0 \text{, and}$$
$$B_{spijt} = 1 \text{, for all j}$$

would mean that consumers use only price and price expectations to make decisions.

Most models in marketing are based on the assumption that the underlying state space, that is, the information being used by consumers to make decisions, is known to researchers. Having knowledge of the state space in which the consumer is making a decision, models can be estimated on both experimental and dynamic panel data. Unfortunately, this extreme viewpoint is not able to adequately capture the process of consumer decision making. Decision making is not only influenced by the environment but also by the inherent limitations of consumers' mental capacities. Depending on the external environment in which decision makers find themselves, they will use different information to make their decisions (Payne, Bettman and Johnson 1988).

Activation functions can be used to capture the interaction of environment and consumer limitation. In common with the work of Rumelhart and McCleland (1986) and Anderson (1976), the psychological limitation of the human mind is captured using activation functions. There is some probability of information being used by consumers. As consumers are overwhelmed by the amount of information available to them, only a few bits of information can be brought into their working memory at any given time. Activation plays an important role in determining the different pieces of information that will coexist in the working memory at any given time. The activation mechanism is assumed to bring the information pertinent to a particular situation into working memory. Viewing it normatively, every piece of information is competing with others for attention and, thus there is some likelihood of it being attended.

An important advantage of using the activation function is that assumption of the state, that is, the information being used by consumers to make decisions need not be made *a priori*. The probability that a certain state will occur can be calculated as follows:

$$P(s_{it}) = \prod_{j}\prod_{k}\{\pi_{kijt}\}^{B_{skijt}} * (1-\pi_{kijt})^{(1-B_{skijt})} \tag{4.2.2}$$

where,

s is one of the the possible 2^(J X K) states,

J is the total number of brands in the market, and

K is the total number of attributes that a consumer will use to make decisions.

$P(s_{it})$ is the probability of consumer i being in the s-th state in the t-th time period,

B_{skijt} = 0 if consumer i is not attending to the k-th attribute of the j-th brand in the t-th time period given that the consumer is in the s-th state.

= 1 otherwise,

and,

π_{kijt} , the activation function, is the probability that consumer i is attending to the k-th attribute of the j-th brand at t-th time period.

It is modeled as a binary variable with the expectation that the activation function may be expressed as a logistic function:

$$\pi_{kijt} = \frac{\exp(z_{kijt}\gamma_k)}{1+\exp(z_{kijt}\gamma_k)}, \tag{4.2.3}$$

k.....indexes for the attributes,

z_{kijt} .... is a vector of 1 x m values of variables for the i-th individual on which the activation of k-th attribute of the j-th brand is dependent upon at the t-th time period, and

γ_k is a m x 1 vector of parameters that needs to be estimated for the k th attribute.

Anderson (1993) showed that the likelihood that a piece of information will be used in a particular situation is dependent on two factors: (1) history factors and (2) context factors. A history factor is the record of all the times information was used, such as: (1) frequency of use, that is, how many times in the past has it been used (2) recency of use, that is, how recently has it been used (3) spacing of use, that is, time between use. On the other hand, context factors are the external cues in the environment that will activate a particular information node.

Based on the work of Anderson (1993), Burrell (1980,1985) and Stritter (1977), it can be shown that under the assumptions of conditional independence, the likelihood of an attribute "k" being used in the choice decision is given by

$$P(k_j / H_j, Q_j) = f(H_j, Q_j), \tag{4.3}$$

where,

P( ) is the probability that the k-th attribute of the j-th brand is attended to,

j is the brand index,

H is the history, and

Q is the cue present in the environment.

It is hypothesized that there are two variables that will affect activation. These are as follows:

(a) History of the brand. Prior history of the brand will manifest itself in two variables: (1) consumer specific, brand specific intercepts[5], and (2) the time since the last brand was purchased by a consumer.

(b) Cues in the environment -for this analysis only the effect of promotion on activation will be investigated. A positive relationship is expected between promotion and activation. The promotional activity of any brand drives attention to the gain in price, that is, gain becomes an important component of the decision making. Thus, if a brand has been promoted, it is likely that the price gain of that brand will be noticed by the consumers (Dickson and Sawyer 1990; and Gijsbrechts 1993).

Hence, from the discussion the activation function in its most general form can be written as:

$$\pi_{kijt} = \frac{\exp(\gamma_{0kj} + \gamma_{Ikj}\alpha_{oijt} + \gamma_{Tkj}(Date_{it} - Date_{it^{*}-1}) + \gamma_{Dkj}D_{jt})}{1 + \exp(\gamma_{0kj} + \gamma_{Ikj}\alpha_{oijt} + \gamma_{Tkj}(Date_{it} - Date_{it^{*}-1}) + \gamma_{Dkj}D_{jt})}, \qquad (4.4)$$

where,

k is the k-th attribute,

D_{jt} =1 if j th brand is on promotion and is displayed as an ad feature or a display at the t-th store visit,

=0 otherwise.

For each attribute, four parameters have to be estimated - (suppressing subscripts) $\gamma_0, \gamma_I, \gamma_T$ and γ_D. γ_I can take either positive or negative values.

---

[5]frequency of brand purchase and individual level brand intercept will be highly correlated.

$\gamma_l > 0$, means that consumers are predisposed to use attributes of brands that they are loyal to in order to make decisions. This specification captures the impact of consideration set on brand choice. If the brand specific activation functions have large positive values and the intrinsic brand preference has large positive values, then probability that a brand is a part of the consideration set is large. On the other hand, a negative γ_l means that consumers use only unobserved common attributes (such as taste, color etc.) of the brands that they are loyal to, to make decisions. The observed attributes of other brands, such as price, that they are not loyal to are used to make decision.

A negative sign is expected for the parameter γ_T. γ_T captures the impact of time since the last brand was purchased on the activation function. A negative parameter captures the idea that with the increase in inter-purchase timing the probability of the impact of observed attributes, such as price and advertising, will decrease over time. An explanation for this decrease in probability comes from the ideas of regret theory (Bell 1982). Regret is typically defined as the feeling induced by comparing a given outcome or state of events with a state of foregone alternative (Bell 1982). Inman and Mcalister (1994) used regret theory to predict that expiration dates for coupons induce a second mode in the redemption pattern just prior to the expiration date. Empirical results show that the results are consistent with the predictions set out in their study. A negative coefficient, building on these empirical and experimental studies, would mean that utility gained by making a choice earlier than expected (by virtue of consumers' mean purchase timing) is a result of not regretting to pass up a current attractive offer in anticipation of an attractive offer in the future. A major reason for this regret is in part due to observable attributes. The impact of regret, for example, due to not buying a brand on promotion, will decrease over time as the inventory effects will dominate the decision to buy in the product category.

4.2.3 Price Expectations

A number of studies in marketing have shown that anticipation or expectation plays an important role in consumer choice decision (Winer 1985, 1986; and Kalwani et al. 1990). In this framework price expectations have been modeled as a geometrically decaying function of lagged prices. A notable difference between the characterization of price expectations in this models and the ones used in earlier studies (Kalwani et. al. 1990) is that this model tests for the effects of attention on price expectations. Price expectations are modeled as follows:

$$p^e_{ijt} = \overline{p}^e_{ijt} + \eta_{ij} \tag{4.5}$$

where,

$$\eta_{ij} \sim N(0, \sigma^2_\eta), \tag{4.6}$$

Equation 4.5 reflects the possibility that consumer price expectations, being a latent variable, have a possibility of error in their measurement. This error in measurement has been assumed to be IID normal, with zero mean and variance constant across time. The error term η_{ij} can be decomposed into errors due to two different effects, and can be written as follows:

$$\eta_{ij} = \eta_{1ij} + \eta_{2ij} \tag{4.7}$$

where,

$$\eta_{1ij} \sim N(0,\sigma^2_{\eta_1}),$$

and

$$\eta_{2ij} \sim N(0,\sigma^2_{\eta_2})$$

and, that

$$\sigma^2_{\eta} = \sigma^2_{\eta_1} + \sigma^2_{\eta_2} \qquad \mathrm{cov}(\sigma_{\eta_1},\sigma_{\eta_2})=0 \tag{4.8}$$

η_{1ij} captures the fact that consumers often do not have a good idea about the reference price for a product, particularly for a frequently purchased product. This might be due to a loss of memory or the effect of an inexpensive product that is being bought by a consumer (Dickson and Sawyer 1990).

η_{2ij} tries to capture the measurement errors involved in specifying the latent reference price construct, that is, though the history of past prices that might have been observed by a consumer is used as a rough estimate for reference price, this specification of reference price has errors in measurement that have to be incorporated in equation 4.7.

The process of price expectations is modeled as an adaptive mechanism which is also affected by consumers' attention to price at any given time. In common with the literature on reference price (Winer 1986; and Kalwani 1990), $\bar{p}^e_{ijt}$ is modeled as an adaptive process with decaying effects of lagged prices for all shopping trips. If attention is assumed to have no effect on price expectations, then the process of adaptive expectations is given as follows:

$$\overline{p_{ijt}^{\,e}} = \lambda p_{ijt-1} + (1-\lambda)\overline{p_{ijt-1}^{\,e}}\ , \tag{4.9}$$

where

λ is a parameter between 0 and 1.

In order to capture the effects of attention on price expectations, equation 4.9 can be rewritten as follows:

$$\overline{p_{ijt}^{\,e}} = \lambda_{\pi_{pijt}}\, p_{ijt-1} + (1-\lambda_{\pi_{pijt}})\overline{p_{ijt-1}^{\,e}}\ , \tag{4.10}$$

where,

$$\lambda_{\pi_{pijt}} = \lambda_o \pi_{pijt}\,. \tag{4.11}$$

4.2.4 Consumer Utility for No Purchase

For a model that captures both brand choice and purchase incidence, it is important to define the utility of no purchase. This has been specified as given below:

$$U_{i\bar{j}t} = \alpha_{np} + \alpha_{PI}(Date_{it} - Date_{it^*-1}) + \varepsilon_{i\bar{j}t}\,, \tag{4.12}$$

where,

$U_{i\bar{j}t}$ is the utility derived by the i-th consumer at t-th time period from not purchasing in the product category given by $\bar{j}$,

α_{np} is the intercept term for no purchase, constant across individuals and time period

$Date_{it}$ is the day on which the t-th store visit was made, and

$Date_{it^*-1}$ is the day on which last time any brand was purchased.

The term $Date_{it} - Date_{it^*-1}$ captures the time varying effect of disutility of purchase or the utility of no purchase. As the length of time since last brand was purchased increases, the disutility of purchase decreases. A rationale for this specification is that as the stock of inventory for a product goes down, the probability of purchase becomes higher. Hence, the expected sign of α_{Pt} is negative. Another interpretation of equation 4.12 is that the term $Date_{it} - Date_{it^*-1}$ (time since the last purchase) has a double exponential extreme value (weibull-gumble) distribution with a survivor, and hazard rate as given below:

$$S(t_d) = \exp\{-e^{-\mu_{t_d}(t_d - \eta_{t_d})}\} \tag{4.12.1}$$

and

$$h(t_d) = \frac{1}{\mu_{t_d}}[\exp\{\frac{(t-\eta_{t_d})}{\mu_{t_d}}\}] \tag{4.12.2}$$

where,

μ_{t_d} is a scale parameter that has been fixed to a value of 1,

and η_{t_d} is the mode of the distribution.

The hazard rate h(t) captures the likelihood of purchase, given that it has not occurred in the time interval (0,t). Hence, the purchase incidence model is nested within the brand choice model.[6]

---

[6] This specification does not control for inventory effects. The operationalization of the inventory variable used by Chintagunta (1993), Bucklin and Lattin (1993) is too approximate to be used in estimating models of purchase incidence and brand choice. As inventory is an unobserved variable, its operationalization is riddled with measurement errors. This renders the parameter estimates biased, inconsistent, and unstable.

4.2.5 Consumer Brand Choice Probabilities

In order to obtain consumer brand choice probabilities, one needs to substitute eq 4.5 in eq 4.1. The utility derived by consumer i for the j-th brand in the t-th time period, given that the consumer is in state s is given by:

$$\begin{aligned} U_{sijt} &= \alpha_{oijt} + B_{spijt}\alpha_R\eta_{ij} + (1-B_{spijt})\alpha_p\eta_{ij} + B_{spijt}\alpha_p p_{jt} + B_{spijt}\alpha_R(\overline{p^e_{ijt}} - p_{jt}) \\ &+ (1-B_{spijt})\alpha_P \overline{p^e_{ijt}} + B_{saijt}\alpha_a a_{jt} + \alpha_{std} d_{ijt^*-1} + \varepsilon_{ijt} \end{aligned} \tag{4.13}$$

where,

$\tilde{\alpha}_{0j} \sim N(\overline{\alpha_{0j}}, \sigma^2_{\alpha_{0j}})$,

and,

$\eta_{ij} \sim N(0, \sigma^2_\eta)$.

Assuming,

$$\sigma^2_{jp} = \alpha^2_R * \sigma^2_\eta = \alpha^2_p * \sigma^2_\eta$$

equation 4.13 can be written as follows:

$$\begin{aligned} U_{sijt} &= \alpha_{1oijt} + B_{spijt}\alpha_p p_{jt} + B_{spijt}\alpha_R(\overline{p^e_{ijt}} - p_{jt}) \\ &+ (1-B_{spijt})\alpha_P \overline{p^e_{ijt}} + B_{saijt}\alpha_a a_{jt} + \alpha_{std} d_{ijt^*-1} + \varepsilon_{ijt} \end{aligned} \tag{4.14}$$

where,

$$\tilde{\alpha}_{10j} \sim N(\overline{\alpha_{10j}}, \sigma^2_{\alpha_{10j}}),$$

and

$$\overline{\alpha_{10j}} = \overline{\alpha_{0j}},$$

$$\sigma^2_{\alpha_{10}} = \sigma^2_0 + \sigma^2_{jp} \qquad \qquad \operatorname{cov}(\sigma_0, \sigma_{jp}) = 0,$$

where,

$$\sigma^2_{\alpha_{10}} = \sigma^2_0 + \sigma^2_\eta(\alpha_R B_{Pijt} + (1 - B_{Pijt})\alpha_p)^2.$$

4.2.6 Estimating the Mean Levels of the Unobserved Attributes

The utility equation specified by equation 4.14 assumes that consumers use only observed attributes to make decisions. Elrod (1988), and Elrod and Keane (1995) specify a method by which mean utility of unobserved attributes can be inferred from scanner panel data. According to Elrod's (1988) specification, the utility that consumer i derives from the j-th brand at the t-th time period can be re-expressed as follows:

$$\begin{aligned} U_{sijt} &= L_j W_i + B_{spijt}\alpha_p p_{jt} + B_{spijt}\alpha_R(\overline{p^e_{ijt}} - p_{jt}) \\ &+ (1 - B_{spijt})\alpha_P \overline{p^e_{ijt}} + B_{saijt}\alpha_a a_{jt} + \alpha_{sd} d_{ijt^*-1} + \varepsilon_{ijt} \end{aligned} \tag{4.15}$$

where,

Lj is the mean level of the unobserved "latent" attribute that is being used by consumers to make decisions, and

$$W_i \sim N(\overline{W}, 1)$$

The model specified in equation 4.15 belongs to the general class of models proposed by Elrod and Keane (1995). Elrod and Keane (1995) developed a factor analytic specification for the covariance matrix of the consumer specific utility term to infer the underlying market structure from revealed preference data, such as scanner panel data. The model described by equation 4.15 can be viewed as a "principal component" specification of the covariance matrix. Hence, the formal models that are proposed in this thesis can infer underlying brand position from revealed preference data. However, most of the other models in the market structure literature (Elrod 1988; Elrod and Keane 1995) do not capture the computational limitations of the consumer.

It is important to note here that the utility equation specified in equation 4.15 subsumes a number of different models proposed in the marketing literature. An advantage of this utility formulation is that the models that assume consumers use all available information to make decisions are a limiting condition of the structural model specified in equation 4.14. For a brand choice model, if the activation function (B's) are assumed to be equal to 1, weights are assumed to be uniform, α_R is equal to zero, and a smoothing exponential function is assumed for the state dependence variable, then we obtain the Guadagni and Little (1983) model. Similarly, assuming expectations to be formed rationally, and the activation function (Bs) to be equal to 1, and weights to be uniform, we obtain the model of Winer (1986). Incorporating the effects of gains and losses and assuming uniform weights, Bs equal to 1, and expectations are formed rationally, we obtain the model of Kalwani et al. (1990).

4.3 Implication of the Model

A major implication of the structural model developed in this thesis is that it rejects the assumption of utility maximization theory that of consumers using all available information to make decisions. Consistent with the models outlined in the bounded/adaptive rationality literature (Bettman, Johnson and Payne 1990), it models the consumer decision making without making the assumption of normative decision rules. At a fundamental level, this work provides an analytical framework that can be used to incorporate principles of adaptation (bounded rationality) within the utility maximization framework without requiring measures of cost of thinking (Shugan

1980). In fact, it goes a step further than the models of bounded rationality proposed by Shugan (1980), and Payne, Bettman and Johnson (1988). A major limitation to the framework proposed by Shugan (1980), and Payne, Bettman and Johnson (1988) is a lack of understanding of how consumers decide which rule (heuristic) to use in any particular situation (Meyer and Kahn 1991). Activation theory, proposed in this thesis, provides an explanation for this, as information stored in the memory and as external stimuli are competing with each other for consumers' attention. Factors that are intrinsic to a consumer or exist in the external environment will not only determine the value of the activation function, but also information that will co-exist in the working memory at any given time. Hence, it provides an insight into choice strategy being used by consumers to make decisions.

At a more applied level, this model can be used to understand different consumer choice concepts such as dynamic consideration sets, risk aversion and the promotional sensitivity signal. The activation function can be interpreted in a number of different ways. It can be conceived of as the degree of confidence that the preferred feature of an attribute is present. For example, Grossberg and Gutowski (1987) use the activation function to model the risk aversion behavior in the extended dynamic prospect theory. The expectation π_{ijt} in equation 4.2 can also be viewed as a dynamic pi function of prospect theory. As the activation mechanism is a function of the brand promotional activities, a positive relationship would show that an individual is more certain of a price cut only when it is accompanied by an advertisement or display. Thus, the activation mechanism provides an explanation for the empirical phenomenon of price promotion not accompanied by an ad-feature, or a display has no effect on sales (Popkowski Leszczyc 1994) or the effect of promotional signal sensitivity found by Inman and McAlister (1993).

Inman and McAlister (1993) argued that promotional sensitivity, the fact that consumers respond to a promotional signal even though there is no price cut, is a result of consumers being conditioned to respond to promotional signals. A major limitation of this explanation has been provided by Chomsky (1957;1968). Using paucity of stimuli as a counter argument to the S-R paradigm, Chomsky (1968) contends that the amount of data that an individuals receive in their lifetime is not enough to uniquely determine the response to each and every stimuli they are exposed to. Thus, the S-R explanation provided by Inman and Mcalister (1993) is not adequate for understanding the effects of promotional signals on consumer behavior.

A more plausible explanation is found in the works of Grossberg and Gutowski (1987). They proposed a dynamic model of risk aversion that provides an information processing explanation for the effects of the promotional signal. They contend that in a dynamic model of risky choice, the riskiness of an alternative has to be inferred from the environment. It can also be viewed as an interaction between short term and long term memory where the activation mechanism captures the risk associated with the attribute of an alternative and the reference price is the arousal level stored in the long term memory. The more the risk associated with the attribute of an alternative, the less will be its activation in the short term memory and, hence, the less the impact of that attribute on choice decisions. As the activation mechanism is dependent upon brand familiarity captured by intrinsic brand preference and promotional signal such as newspaper advertising and end of aisle displays, a positive relationship between activation and these two variables would signify that consumers are more confident of a price cut only when it is accompanied by a promotional signal.[7] The less risky the alternative, the more its impact on choice decisions.

By building on the arguments provided in the previous paragraphs, the π function can also be thought of as capturing the dynamic construct of consideration sets. Smaller values of the activation function would imply low probability of marketing mix variables having an impact on brand utility. If the brand specific activation functions have small positive values, and the intrinsic brand preferences have large negative values, then the probability that the brand is being considered by consumers is very small: the higher the risk associated with an alternative, captured by the activation mechanism, the lower the probability that it will be a part of the consideration set. A positive relationship between the external stimuli and activation function would increase the probability of a brand being included in the set of brands being considered at any given time.

In summation, this framework allows me to capture the psychological limitation of the consumer, adaptive rationality, consideration set and choice heuristics. In the next chapter, I provide a method by which satisficing models specified in the previous section can be implemented on scanner panel data.

[7] Extension of this model can be shown to capture and explain the effects of preference reversal. Readers are referred to Grossberg and Gutowski (1987) for further details.

Chapter Five

Estimation

The estimation of structural models is made more difficult due to the fact that the attributes consumers are using to make decisions are not known to the researcher. Thus, the likelihood function for the data has to be evaluated over all possible combinations of attributes and brands, thereby integrating the uncertainty due to these possible states out of the likelihood function. In this chapter I discuss the technical aspects of the procedure that I use to estimate these models. The models to be estimated are specified below. Utility for brand choice is specified as:

$$\begin{aligned} U_{sijt} &= L_j W_i + B_{spijt}\alpha_p p_{jt} + B_{spijt}\alpha_R(\overline{p^e_{ijt}} - p_{jt}) \\ &+ (1 - B_{spijt})\alpha_P \overline{p^e_{ijt}} + B_{saijt}\alpha_a a_{jt} + \alpha_{std} d_{ijt^*-1} + \varepsilon_{ijt} \end{aligned} \tag{5.1}$$

where,

$$W_i \sim N(\overline{W}, 1),$$

and the utility for no choice is specified as:

$$U_{i\bar{j}t} = \alpha_{np} + \alpha_{Pl}(Date_{it} - Date_{it^*-1}) + \varepsilon_{i\bar{j}t} \tag{5.2}$$

In the structural model given by equation 5.1, consumers' probability of choice is specified by a logit formulation. If v is defined as the deterministic part of the expected utility function, then v is dependent upon the attributes of different brands that are being used to make decisions. The number of different combinations of the attributes and brands that could have been used by consumers to make decisions is given by 2^(J X K), where J is the number of brands in the market, and K is the total number of attributes that could have been used by consumers to make choice decisions.

Let us assume that parameters that include price coefficient (α_p), the advertising coefficient (α_a), the relative price coefficient (α_R),the common attribute (Lj), the coefficient for state dependence (α_{sd}), are grouped into a vector given by θ_1; and the coefficient for no purchase (α_{np}), the coefficient for inter-purchase timing (α_{PI}), are grouped into another vector, given by θ_2. Similarly, parameters for the activation function that include the intercept term γ_0, the parameter that captures the impact of brand preference on activation γ_I, the coefficient for inter-purchase timing γ_T , and that for promotion γ_D, are grouped into another vector θ_3.

Using this notation described above, equation 5.1 and 5.2 can be re-expressed as follows:

$$U_{sijt} = V_{sijt}(W_i, \theta_1) + \varepsilon_{ijt} \tag{5.3}$$

$$U_{i\bar{j}t} = V_{i\bar{j}t}(\theta_2) + \varepsilon_{i\bar{j}t} \tag{5.4}$$

where the indexes are as described above.

Unfortunately, the general equation for the choice probabilities of the model cannot be specified by a common equation for the two states, that is, for: (1) store visits on which a purchase is made in the product category by consumers; (2) store visit on which a purchase is not made in the product category by consumers. The following section outlines a method by which these probabilities can be specified.

The unconditional probability that the j-th brand is chosen by consumer i on the t-th time period is given as follows:

$$\begin{aligned} P(ijt) &= d^*_{it}\ \{\sum_{S}\ [P(ijt|\{s_{it},\theta_1,\theta_2,W_i\})]*P(s_{it}|\{\theta_3\})]\} \\ &+ (1-d^*_{it})*\{\sum_{S}\ [P(ijt|\{s_{it},\theta_1,\theta_2,W_i,\Delta_{it}=1\})] \\ &* P(s_{it}|\{\theta_3,\Delta_{it}=1\})P(\Delta_{it}=1)\} \end{aligned} \tag{5.5}$$

and the unconditional probability that no brand ($\bar{j}$-th alternative) is chosen by consumer i on the t-th time period is given by:

$$\begin{aligned} P(i\bar{j}t) &= d^*_{it}*(\sum_{S}[P(i\bar{j}t|\{s_{it},\theta_1,\theta_2,W_i\})P(s_{it}|\{\theta_3\})] \\ &+ (1-d^*_{it})*\{\sum_{S}[P(i\bar{j}t|\{s_{it},\theta_1,\theta_2,W_i,\Delta_{it}=1\})P(s_{it}|\{\theta_3,\Delta_{it}=1\})P(\Delta_{it}=1)], \\ &+ P(\Delta_{it}=0)\} \end{aligned} \tag{5.6}$$

where,

$d^*_{it} = 1$, if consumer i makes a purchase in the product category in the t-th time period,

$= 0$, otherwise,

$P(s_{it})$ is the probability of consumer i being in the s-th state in the t-th time period, and,

$P(\Delta_{it})$ is the probability that consumer i is making a decision in the t-th time period.

The different possible scenarios will vary depending upon: (1) the store visit on which no brand is bought, and (2) the store visit on which one of the brands is bought from the purchase category. In order to estimate the model, probabilities specified in equations 5.5 and 5.6 need to be defined.

5.1 Store Visits on Which No Purchase is Made by a Consumer

There are two possible scenarios for a store visit on which no purchase is made in the product category: (1) consumer i makes a decision not to make a purchase in the product category, and (2) consumer i makes no decision at all. Even in the second scenario, consumer i makes no purchase in the product category. As there is complete uncertainty about these two states described above, it has been assumed that the probability of the i-th consumer making a decision at the t-th time period is as follows:

$$P(\Delta_{it} = 1) = .5 \qquad (5.7)$$

Δ_{it} =1, if consumer i is making a decision on the t-th shopping trip

=0, otherwise.

Hence, when consumers make no purchase in the product category, there is an equal probability of consumers either making a decision or that of not making a decision.

If consumers make no decisions in the product category, then a researcher knows with certainty that there is only one state possible, which is that of no information being processed, and hence, no brand being bought by a consumer. The probability of not buying in the product category, given that no decision is being made by a consumer is the product of three probabilities given below:

$$P(\Delta_{it} = 0) = .5, \tag{5.8}$$

$$P(s_{it} \mid \{\theta_3, \Delta_{it} = 0\}) = 1 \tag{5.9}$$

$$P(i\bar{j}t \,/\, \{s_{it}, \theta_1, \theta_2, W_i, \Delta_{it} = 0\}) = 1 \tag{5.10}$$

Combining equation 5.8, 5.9 and 5.10 we obtain the probability of consumer i choosing alternative $\bar{j}$ (no brand in the product category) at time t, given that a consumer is not making a decision. This probability is expressed as follows:

$$P(i\bar{j}t \mid \{\Delta_{it} = 0\}) = [P(i\bar{j}t \mid \{s_{it}, \theta_1, \theta_2, W_i, \Delta_{it} = 0\}) * P(s_{it} \mid \{\theta_3, \Delta_{it} = 0\}) * P(\Delta_{it} = 0)] = .5 \tag{5.11}$$

On the contrary, if a decision is being made in the product category and no brand is being bought in the product category, then there are 2^(J X K) different possible states, where J is the total number of brands in the market, and K is the total number of attributes being used by consumer to make decisions. The probability that a consumer is in a particular state is given as follows:

$$P\,(s_{it} \mid \{\theta_3, \Delta_{it} = 1\}) = \prod_{j=1}^{J} \prod_{k=1}^{K} (\pi_{kijt})^{B_{skijt}} * (1 - \pi_{kijt})^{(1-B_{skijt})} \qquad (5.12)$$

where,

B_{skijt} = 1, if in a given state s, consumer i is attending to the k-th attribute at t-th

time period of the j-th brand,

=0, otherwise,

and,

π_{kijt} is the probability of consumer i attending to the k-th attribute of the j-th brand at t-th time period, and is given as:

$$\pi_{kijt} = \frac{\exp(\gamma_0 + \gamma_I \alpha_{1oijt} + \gamma_T (Date_{it} - Date_{it^*-1}) + \gamma_D D_{jt})}{1 + \exp(\gamma_0 + \gamma_I \alpha_{1oijt} + \gamma_T (Date_{it} - Date_{it^*-1}) + \gamma_D D_{jt})} \qquad (5.13)$$

Given that a decision is being made in the product category, and that a consumer is in state s, the probability that $\bar{j}$-th alternative (not to purchase any brand) is chosen is as follows:

$$P(i\bar{j}t \mid \{s_{it}, \theta_1, \theta_2, W_i, \Delta_{it} = 1\}) = \frac{\exp(V_{i\bar{j}t}(\theta_2))}{\exp(V_{i\bar{j}t}(\theta_2)) + \sum_{j=1}^{J} \exp(V_{ijt}(W_i, \theta_1))} \qquad (5.14)$$

Hence, the probability that consumer i is choosing alternative $\bar{j}$ (no-choice) at t-th time period, given that a consumer is making a decision and is in state s is obtained by combining equations 5.7, 5.12 and 5.14 :

$$[P(i\bar{j}t|\{s_{it},\theta_1,\theta_2,W_i,\Delta_{it}=1\})]P(s_{it}|\{\theta_3,\Delta_{it}=1\})P(\Delta_{it}=1)] \qquad (5.15)$$

and the unconditional probability of not choosing a brand in the market by consumer i for the t-th store visit when no purchase is made in the product category is as follows:

$$P(i\bar{j}t)=\sum_S [P(i\bar{j}t|\{s_{it},\theta_1,\theta_2,W_i,\Delta_{it}=1\})P(s_{it}|\{\theta_3,\Delta_{it}=1\})P(\Delta_{it}=1)]+P(\Delta_{it}=0) \qquad (5.16)$$

where ,

S is the total number possible combinations.

Also, the unconditional probability of consumer i choosing brand j on the t-th store visit on which no purchase was made in the product category is given by:-

$$P(ijt)=\sum_S [P(ijt|\{s_{it},\theta_1,\theta_2,W_i,\Delta_{it}=1\})P(s_{it}|\{\theta_3,\Delta_{it}=1\})P(\Delta_{it}=1)] \qquad (5.17)$$

5.2 Store Visit on Which a Brand is Bought from the Product Category

If on a store visit consumer i bought a brand in the product category, then a researcher knows with certainty that a consumer is making a decision. This can be expressed as follows:

$$P(\Delta_{it} = 1) = 1 \qquad (5.18).$$

As a decision is being made in the product category, there are in all 2^(J X K) different possible states, with the probability that a consumer is in any particular state given by equation 5.9 and equation 5.10.

Given that a decision is being made in the product category, and that a consumer is in state s, the probability that j-th brand is chosen on the t-th store visit is as follows:

$$P(ijt|\{s_{it},\theta_1,\theta_2,W_i,\Delta_{it}=1\}) = [\frac{\exp(V_{ijt}(W_i,\theta_1))}{\exp(V_{i\bar{j}t}(\theta_2)) + \sum_{j=1}^{J} \exp(V_{ijt}(W_i,\theta_1))}] \qquad (5.19)$$

Combining equations 5.12, 5.18 and 5.19, the unconditional probability of j-th brand being chosen on the t-th time period by consumer i is given as follows:

$$P(ijt) = \sum_{S} [P(ijt|\{s_{it},\theta_1,\theta_2,W_i,\Delta_{it}=1\})]P(s_{it}|\{\theta_3,\Delta_{it}=1\})P(\Delta_{it}=1)] \qquad (5.20)$$

As equation 5.20 is independent of Δ_{it}, it can be expressed as follows:

$$P(ijt) = \sum_{S} [P(ijt|\{s_{it},\theta_1,\theta_2,W_i\})] * P(s_{it}|\{\theta_3\})] \qquad (5.21)$$

Also, the unconditional probability that consumer i chooses not to purchase a brand in the product category on the t-th store visit on which a brand was bought is given as follows:

$$P(i\bar{j}t) = \sum_{S} [P(i\bar{j}t|\{s_{it}, \theta_1, \theta_2, W_i\})] * P(s_{it}|\{\theta_3\})] \qquad (5.22)$$

Combining equation 5.16 and 5.22, the unconditional probability that consumer i chooses brand j on the t-th store visit is expressed as follows :

$$\begin{aligned} P(i\bar{j}t) &= d^*_{it} \{\sum_{S} [P(i\bar{j}t|\{s_{it}, \theta_1, \theta_2, W_i\})] * P(s_{it}|\{\theta_3\})]\} \\ &+ (1-d^*_{it}) * \{\sum_{S} [P(i\bar{j}t|\{s_{it}, \theta_1, \theta_2, W_i, \Delta_{it}=1\})] * P(s_{it}|\{\theta_3, \Delta_{it}=1\}) P(\Delta_{it}=1)] + P(\Delta_{it}=0)\} \end{aligned} \qquad (5.23)$$

and, combining equations 5.17 and 5.21, the unconditional probabilities that consumer i chooses the j-th brand on the t-th store visit is given as follows:-

$$\begin{aligned} P(ijt) &= d^*_{it} \{\sum_{S} [P(ijt|\{s_{it}, \theta_1, \theta_2, W_i\})] * P(s_{it}|\{\theta_3\})]\} \\ &+ (1-d^*_{it}) * \{\sum_{S} [P(ijt|\{s_{it}, \theta_1, \theta_2, W_i, \Delta_{it}=1\})] * P(s_{it}|\{\theta_3, \Delta_{it}=1\}) P(\Delta_{it}=1)] \end{aligned} \qquad (5.24)$$

where,

$d^*_{it} = 1$, if consumer i makes a purchase at the t-th time period in the product category,

=0 otherwise,

and,

$\bar{j}$ = option of no choice,

j = the j-th brand in the market.

Both equation 5.23 and 5.24 have a heterogeneous logit formulation (McFadden 1991), hence, do not have a closed form solution. Variants of this model have been estimated using numerical integration (Elrod 1988), a semi-parametric approach (Chintagunta 1991), or sampling techniques (Erdem 1993). A major limitation of the numerical integration approach is that integration of more than three or four dimensions is not possible and therefore cannot be applied to large choice sets. Contrary to this, the semi-parametric approach is often considered to be the most preferred estimation procedure as the underlying distribution does not have to be specified (Elrod 1991). Rather, the heterogeneity distribution is approximated by a cumulative distribution function (Kamakura and Russell 1989). A major limitation of the semi-parametric approach is that, as a multivariate distribution function is being approximated by a few mass points, the possibility of mis-specification is very high. Though simulation techniques suffer from the drawback of having to specify the underlying distribution for heterogeneity, an important advantage of this method is that high dimensional integrals can be evaluated using very inexpensive Monte-Carlo techniques (McFadden 1989; Pakes and Pollard 1989).

Unfortunately, the model specified by equation 5.23 and 5.24, must be evaluated over discrete distributions in order to account for uncertainty due to researcher's lack of knowledge of the information being used by consumers to make decisions. Similar to the models in Monte-Carlo simulation literature (Erdem 1993; Supan and Hajivassiliou 1993), for a given state space, simulates are sampled from the prior distribution. Given the values of the simulates, state space and the parameter estimates required to evaluate the probability function provided in equation 5.14 and equation 5.19, it is possible to construct a likelihood function for the observed data.

Using an iterative procedure such as BFGS, one can obtain the maximum likelihood estimates for the problem.

If Θ are the estimates of the set of parameters on any given iteration, then the corresponding simulated log-likelihood for the data set conditional on the parameters, is given by

$$LL = \sum_{i=1}^{N}\sum_{j=1}^{J}\sum_{t=1}^{T}(d_{ijt}) * LnP*(ijt|\Theta) + \sum_{i=1}^{N}\sum_{t=1}^{T}(d_{i\bar{j}t}) * LnP*(i\bar{j}t|\Theta) \qquad (5.25)$$

where,

$d_{ijt} = 1$, if consumer i chooses the j -th brand in the t-th time period

=0, otherwise, and,

$d_{i\bar{j}t} = 1$, if consumer i chooses not to purchase in the product category in the t-th time period

=0, otherwise.

The simulated probabilities $P*(ijt|\Theta)$ and $P*(i\bar{j}t|\Theta)$ are given as follows:

$$\begin{aligned} P*(ijt) = d*_{it}\{\sum_{S}[P*(ijt|\{s_{it},\theta_1,\theta_2\})]*P(s_{it}|\{\theta_3\})]\} \\ +(1-d*_{it})*\{\sum_{S}[P*(ijt|\{s_{it},\theta_1,\theta_2,\Delta_{it}=1\})]*P(s_{it}|\{\theta_3,\Delta_{it}=1\})P(\Delta_{it}=1)] \end{aligned} \qquad (5.26)$$

and,

$$\begin{aligned} P*(i\bar{j}t) = d*_{it}\{\sum_{S}[P*(i\bar{j}t|\{s_{it},\theta_1,\theta_2\})]*P(s_{it}|\{\theta_3\})]\} \\ +(1-d*_{it})*\{\sum_{S}[P*(i\bar{j}t|\{s_{it},\theta_1,\theta_2,\Delta_{it}=1\})]*P(s_{it}|\{\theta_3,\Delta_{it}=1\})P(\Delta_{it}=1)]+P(\Delta_{it}=0)\} \end{aligned} \qquad (5.27)$$

where,

$$P^*(ijt|\{s_{it},\theta_1,\theta_2,\Delta_{it}=1\})=\frac{1}{N}\sum_{l=1}^{N} P(ijt|\{s_{it},\theta_1,\theta_2,W_{il},\Delta_{it}=1\}),$$

and

$$P^*(i\bar{j}t|\{s_{it},\theta_1,\theta_2,\Delta_{it}=1\})=\frac{1}{N}\sum_{l=1}^{N} P(i\bar{j}t|\{s_{it},\theta_1,\theta_2,W_{il},\Delta_{it}=1\}).$$

W_{il}, 1,2.....,N, are random vectors drawn from the distribution $f(W_{il}|\Theta)$ and N is the total number of draws from this distribution.

Thus, monte-carlo integration can be used to estimate the parameters of the model described by equation 5.26 and 5.27. Before discussing the estimation result (Chapter seven), I first discuss the specifics of the data set used for theory testing.

Chapter Six

Data Description

The data that I have used to test the hypotheses is a subset of the ketchup data provided by Nielsen Inc. The data base includes store and brand choice information for approximately 3000 households, from the two markets in Sioux Falls SD, and Springfield MO. Daily data is available for three years between 1986 and 1988, and a variety of exogenous variables are provided for each individual store visit. The household demographic information is also available. The ketchup consists of three major brands, Heinz, Hunts, Del-Monte, and a group of private labels and generic brands. These account for 99% of the market share.

I use scanner panel data for two reasons: (1) the model has been developed for frequently purchased non-durable goods, and (2) the theory predicts the mediating role of reference price and activation or attention on brand choice. The formation of reference price and activation mechanism requires some experience by consumers. Ketchup is a product category that is frequently purchased by consumers which will provide them with the prequisite experience to form price expectations.

Ketchup as a product category has been chosen for the following reasons:

(1) it is a frequently and regularly purchased item;

(2) the brands in this category are regularly promoted, and are therefore a source of price variability;

(3) as my model does not account for purchase quantity, it is desirable to use a product category in which consumers rarely or never purchase more than one item of the same brand; and

(4) previous research has shown (Erdem 1993) that for mature market products like ketchup or laundry detergent, preferences for the brands are well established.

As my model does not incorporate learning effects, ketchup data set was a good choice for investigation.

This data set has all the information required to estimate and test the model. The exogenous variables, for each store visit, such as price, advertising, price specials, displays, store at which the brand was bought are available. The information required to implement the probabilistic activation function is also available or can be imputed reasonably well from the data.

6.1 The Data Set

One hundred individuals were randomly sampled from Sioux Fall data for the years 1987 to 1988. In this market there are three major brands- Heinz, Hunts, Del Monte and a group of generic brands. Heinz is the market leader, followed by Hunts and Del Monte. The market share for the four brands is provided in the table below.

Insert Table 3 about here

In order to limit the dimensionality of the underlying heterogeneity structure different sizes of the same brand were not included in the analysis. Rather, price was converted to price per ounce. In a way, all sizes were included in the study but were not distinguished. The dependent variable is a brand choice variable. Consumers can choose from the four alternative brands, and can also choose not to purchase any brand on a given store visit.

The exogenous variables such as price, advertising, purchase timing and promotional variables were used for estimating the model. The purchase history (consumers' purchases over time) and the retail tracking files (weekly store sales) were used to obtain price data. To calculate prices of the competing brands, store specific prices were used as prices differed significantly by stores. Prices per ounce of ketchup differed significantly depending on bottle size. Therefore, prices for the competing brands were calculated for the same bottle size.

There are in total twenty five stores, four brands and four different bottle sizes of each brand in each store for a total of one hundred and fifty weeks. Unfortunately, when no purchase is made during a week in a particular store, price data for that store is not available. Whenever price data was missing, chain price for the brand was used. Without this information, average price for the last quarter for the store is calculated. If no purchases were made in the past quarter, then prices were averaged over the previous year. If price information was still not available, then the prices were averaged over a period of three years.

The price variable is the actual price paid by the consumer before the use of coupons. Coupon usage is not incorporated as one of the independent variables. The information about the coupon value for the brand that is being purchase is only available. The information about coupon value or availability cannot be obtained for the competing brands. This has an effect of biasing the results. In addition, the process of coupon selection is an endogenous variable beyond the scope of this study.

The advertising variable is a dummy variable indicating whether a particular brand was locally advertised (in the local newspaper) or not. The promotional variable is also a binary variable that takes a value of 0 if a brand is not promoted, and a value of 1 if it is promoted. One important aspect of the promotion that needs to be pointed out is that it is recorded only if a price promotion is accompanied by a display or an ad feature. Hence, the promotional variable is an interaction between a price promotion and an in-store promotion or an ad feature.

Table 3

Market Share

Brand	Market Share
Heinz	70.61%
Hunts	14.42%
Del Monte	7.65%
Generic	7.29%

Chapter Seven

Data Analysis

In this chapter I provide the empirical results for the models developed in chapters four and five. Three models were estimated (1) the "satisficing" model described in chapter four; (2) a heterogenous logit (HGL) model that is a variant of the model proposed by Guadagni and Little (1983); and (3) a reference price model (Kalwani et. al 1989).

An important advantage of the model described by equation 7.1 is that the HGL model (Guadagni and Little 1983) is its limiting case. For example, if Bs are assumed to be equal to1, and if α_R is equal to zero, then the model is a HGL model. Hence the HGL model can be used for evaluating the performance of the satisficing model. The satisficing model is as follows:

$$\begin{aligned} U_{sijt} &= L_j W_i + B_{spijt}\alpha_p p_{jt} + B_{spijt}\alpha_R(\overline{p^e_{ijt}} - p_{jt}) \\ &+ (1 - B_{spijt})\alpha_P \overline{p^e_{ijt}} + B_{saijt}\alpha_a a_{jt} + \alpha_{std} d_{ijt^*-1} + \varepsilon_{ijt} \end{aligned} \tag{7.1}$$

where,

$$W_i \sim N(\overline{W},1).$$

The HGL model can be viewed as a variant of the Guadagni and Little (1983) model. In particular, the HGL tries to disentangle the effects of state dependence and heterogeneity - state dependence is captured by the inclusion of lagged choice and heterogeneity is captured by the assumed functional form for the household specific brand intercepts. As a result in this variant of the Guadagni and Little (1983) model an exponentially decaying function for brand loyalty is not estimated. In addition, the HGL model can also be considered as an extension of the internal market structure model proposed by Elrod (1988), as it incorporates the impact of exogenous variables on choice. The equation for the HGL model is given as follows:

$$U_{ijt} = L_j W_i + \alpha_{Pi} p_{jt} + \alpha_{ai} a_{jt} + \alpha_{std} d_{ijt^*-1} + \varepsilon_{ijt} \tag{7.2}$$

where,

L_j is the factor loading for the "latent" attribute being used by consumer i to make decisions,

W_i is the weight attached by consumer i to this attribute,

α_{Pi} is the price coefficient of consumer i,

α_{std} is the coefficient for state dependence common to all consumers, and

α_{ai} is the advertising coeffcient of consumer i. More specifically, I make the following assumption:

$$W_i \sim N(\overline{W},1) \tag{7.3}$$

$$\alpha_{Pi} \sim N(\overline{\alpha_p},\sigma_p^2) \tag{7.4}$$

$$\alpha_{ai} \sim N(\overline{\alpha}_a,\sigma_a^2) \tag{7.5}$$

Equations 7.3, 7.4 and 7.5 assume that consumer specific weights for the "latent" unobserved variable and that for the coefficient of price and advertising are normally distributed across the population.

The other model used for comparison to the satisficing model is a variant of the reference-price model proposed by Winer (1986) and Kalwani et al. (1990). By constraining the Bs to be equal to 1, $\alpha_p = 0$ and factoring the impact of the difference between price expectation and price $(p^e_{ijt} - p_{jt})$, into gains and losses, one obtains the reference price model specified below:

$$U_{ijt} = L_j W_i + \alpha_{Pi} * (Gain_{ijt} + \lambda_p * Loss_{ijt}) + \alpha_{ai} a_{jt} + \alpha_{std} d_{ijt^*-1} + \varepsilon_{ijt} \qquad 7.6)$$

where,

λ_P is the loss aversion coefficient with respect to price, and,

$Gain_{ijt}$ is defined as the positive difference between reference price and observed price,

$Loss_{ijt}$ is defined as the positive difference between observed price and reference price, and,

α_{Pi}, α_{ai} and W_i have been defined in equations 7.3, 7.4 and 7.5.

The operationalization of Gains and Losses are defined as follows:

When the price of brand j is equal to or below that of the reference price for the j-th brand, then the consumer i faces a positive price difference given by:

$Gain_{ijt} = p^e_{ijt} - p_{jt}$, $\qquad Loss_{ijt} = 0$.

When the price of brand j is above that of the reference price, consumer i faces a negative or loss price difference given by:

$Gain_{ijt} = 0$, $\qquad Loss_{ijt} = p_{jt} - p^e_{ijt}$.

Also, the reference price is defined as follows:

$$p^{e}_{ijt} = \lambda * p_{jt} + (1-\lambda) * p^{e}_{ijt-1} \tag{7.7}$$

where λ is the exponential smoothing function.

It is important to note that the utility function specified in equation 7.2 is an extension of models specified by Winer (1986), and Hardie, Johnson and Fader (1993) as it incorporates heterogeneity in the price coefficient. A few researchers have argued (Kalyanaraman and Winer 1996) that the significance of the loss aversion parameter may in fact be due to the non-inclusion of a heterogeneity measure of the price coefficient. This model provides the opportunity to study the effect of incorporating heterogeneity on the loss aversion coefficient.

In all the three models, the utility for no-purchase has been defined as below:

$$U_{i\bar{j}t} = \alpha_{np} + \alpha_{Pt}(Date_{it} - Date_{it^{*}-1}) + \varepsilon_{i\bar{j}t} \tag{7.8}$$

where,

α_{np} is the intercept term for "no-choice", and,

α_{Pt} is the impact of purchase timing on the overall value for no-purchase option. Hence, the purchase timing model is nested within the brand choice models.

Both the HGL and reference price models are representation of the normative unbounded model in which consumers are represented as utility maximizers who use all available information to make decisions. It is important to note here that HGL and reference price models are limiting and not nested cases of the satisficing model described in equation 7.1. Hence nested tests like the log-likelihood ratio test cannot be used for identifying the best model. Indeed I use AIC, BIC, CAIC and HQ fit criteria for identifying the better model. In the next few sections of this chapter I provide empirical results and model comparisons.

7.1 HGL Model

The parameter estimates and the standard errors of the estimates for the HGL model are provided in Table 4.

Insert Table 4 about here

For identification purpose, the utility of generics has been set to zero. Therefore, statistical tests for the common factors of the three brands hold no meaning. Hence, internal market structure studies do not report the t-values. All parameter estimates have the right sign and are significant, except for the parameter of heterogeneity for the advertising coefficient. It is important to note here that the effects of state dependence and heterogeneity have been incorporated separately in this model. A significant positive parameter for state dependence shows that that the past purchases associated with a brand (brand loyalty) increases the probability of brand choice, even though the utility function is being controlled for heterogeneity (consumer specific idiosyncratic intrinsic brand preference). This result is consistent with Keane (1995), who found significant impact of past purchases on current choice despite inclusion of complex functional forms of heterogeneity.

The negative parameter estimate for the purchase timing variable shows that as time increases since the last brand was purchased, the probability of a consumer making a purchase also increases. This is captured by the decreasing utility of no choice over time.

7.2 Reference Price Model

The results for this model have been provided in Table 5.

Insert Table 5 about here

All parameters have the right sign and are significant, except for the heterogeneity parameter for the advertising coefficient. As hypothesized, the loss aversion parameter has a negative sign, though it is not significantly different from one. This result contradicts the findings of earlier studies (Hardie, Johnson and Fader 1993; Bell and Lattin 1993). Though a proper understanding of this anomaly will require further investigation, one can conjecture that the significance of the loss aversion parameter may in part be due to the non-inclusion of the heterogeneity structure for the price coefficient. Incorporating heterogeneity may render the loss-aversion parameter insignificant.

7.3 Structural Satisficing Model of Brand Choice and Purchase Incidence

The parameters for the satisficing model are the same as the ones for the HGL model, except for the parameters of the activation function and the specification of the impact of price on brand utilities. The satisficing model assumes that it is not only the absolute values of attributes that are important, but also the relative values of the attributes evaluated about a reference point. Values that are less than this reference point are viewed as gains, while those that are more than the reference point are viewed as losses. Hence, the expected sign of the parameter α_R or the parameter for gain is positive.

γ_I captures the effect of intrinsic brand preferences on activation, γ_T incorporates the impact of time since the last brand was purchased on the activation function and γ_D captures the impact of promotion on attention. Though the model in equation 7.1 specifies the attribute specific activation function for the two marketing mix variables included in the model - (price and advertising), in order to decrease the dimensionality of the estimation problem these two are constrained to be equal. Not only does the constraining of the parameters ease the burden of computing high dimensional integrals, it also increases the interpretability of results. The intercept term, γ_o for the activation mechanism, is the threshold level for attention. Large positive values would mean that there is a probability of consumers using all available information to make decisions and conversely, large negative values would mean a small probability of using that information. Hence, the deterministic component of the constrained model estimated here is as follows:

$$V_{sijt} = L_j W_i + B_{sijt}\alpha_p p_{jt} + B_{sijt}\alpha_R(\overline{p^e_{ijt}} - p_{jt}) + (1 - B_{sijt})\alpha_P \overline{p^e_{ijt}} + B_{sijt}\alpha_a a_{jt} + \alpha_{sd} d_{ijt^*-1} \quad (7.9)$$

where,

$B_{sijt} = 1$, if consumer i is attending to the attributes of the j-th brand at t-th time period in the s-th state,

= 0, otherwise.

Insert Table 6 about here

The parameter estimates and the standard errors are provided in table six. A few results in this table warrant an explanation. For identification purposes, the utility for generics has been fixed to an arbitrary value of zero. Hence latent common factors reported in table six are relative to generics and do not have any absolute meaning. As this model differentiates the positions of brands in the attribute space, the result suggests that Heinz has the maximum and Hunts has the lowest unobserved attribute in question.

The results show that all parameters are significant. The coefficients for the activation mechanism, such as that for promotion, are positive. This indicates that the promotional variable makes the attributes of the brands more salient, which leads to an increase in utility and, consequently, an increase in the probability of choice. The intercept term γ_0 is a measure of the threshold value. A positive significant threshold value signifies that consumers are predisposed to use information about attributes to make decisions.

The significant negative parameter for γ_l can be interpreted in the following way. A consumer can simplify a decision problem either (1) by considering a subset of brands available in the market (Andrew and Srinivasan 1995), or (2) by selectively processing pertinent information at any given time (Payne, Bettman and Johnson 1988). Most studies in marketing that attempt to capture the limitations of consumer information processing assume one of the these two positions. γ_l captures both these elements of consumer decision making. A positive parameter signifies that there is a higher likelihood of using information about the attributes of the brands that one has bought in the past (brand loyalty). Conversely, a negative parameter signifies that consumers, while making decisions, use only latent attributes, for example, taste and freshness, for the brands that they are loyal to, and, both unobserved and observed attributes, such as taste and price respectively, of brands that they are not loyal to.

The negative parameter for γ_T shows that with the increase in time since the last brand was purchased, the impact of marketing mix variables on brand choice decreases. A rationale for this result is as follows. As inter-purchase timing is related to inventory, one can say that as time since the last brand was purchased increases, the stock of inventory available to a consumers for use goes down. With a decrease in inventory, consumers wish to replenish the stock. Hence, with the increase in inter-purchase time the inventory effects will dominate the decision, thereby reducing the impact of marketing mix variable on choice. On the other hand, if consumers are making a decision to buy a brand in the market at a time when inventory levels are high, there is a higher probability of them using the information, such as gains in the price, to make decisions. As the need to replenish the stock is not that urgent, consumers can choose to buy the product at a later time. Their processing of information can be ascribed to the concept of regret. A negative coefficient would mean that utility gained by making a choice earlier than expected (by virtue of their mean purchase timing) is a result of not regretting to pass up a current attractive offer in anticipation of an attractive offer in the future. Regret may in turn be due to exogenous variables, such as promotional activity, advertising, or price cuts.

An interpretation of the values and signs of the parameters of the activation function is that both brand familiarity and external cues in the environment decrease the perceived risk involved in making decisions. The satisficing model seems to suggest that both experience associated with brands and promotional signals tend to decrease perceived risk and, hence, increase the probability of purchase. Prior experience with a brand makes the exogenous variables, such as price, not an important component of consumer decision making. Brand loyal consumers use unobserved latent attributes to make choice among brands available in the market. As marketing mix variables are less important to loyal consumers, positive past experiences will have a substantial impact on choice.

A positive parameter for promotion signifies that consumers will take notice of a price cut only when it is accompanied by some form of promotion. This result provides an explanation for the empirical anomaly of promotional senstitivity. The activation function can be thought of as bringing pertinent information to short term memory. Promotional activity brings information about price gains to the working memory, thereby making the gain more salient.

The mechanism of making certain bits of information more active than others can also be thought of as an interaction between long term and short term memory. The price expectation, which can be thought of as an arousal level, is stored in the long term memory as a reference criterion. On the other hand, the activation function can be thought of as bringing the difference between price and the reference price into the active working memory. When a brand is not promoted, the likelihood that this difference will be noticed will be lower than when a brand's price promotional activity is accompanied with a promotional signal.

The satisficing model also tends to suggest that different consumer histories can lead to consumers making radically different decisions. As the price variables that a consumer is exposed to on a store visit can have an impact on the reference point, a gain for one consumer could be a loss for another. Consistent with the findings of Kalwani et al. (1990), the parameter λ_o is not significantly different from one which means that only the last time period's price (store visit) is important in the formation of price expectations (brand specific reference prices).

The parameter estimates for the activation mechanism also seem to suggest that consumers tend to consider only the "latent" attributes of the brands (such as taste, freshness) that they have had some positive experience with, in the past. There is a lower probability of consumer using information about marketing mix variables of these brands, such as price and advertising, to make decisions. Contrary to the findings of Hauser and Wernerfelt (1990), and Roberts and Lattin (1991)[8] that suggest the existence of an optimal size of the consideration set, the results of this research suggest its non-existence. Rather, consumers tend to selectively process information about all brands available in the market. External cues in the form of promotional activities increase the likelihood that observable attributes (price gain due to promotional activity) are brought to the notice of the consumers.

A positive parameter for absolute price can be inferred to mean that consumers use absolute price to discriminate between high quality and low quality brands. The relative price of the brand is used to asses a gain or a loss.

7.4 Goodness of Fit

In order to identify the best of the three models, HGL, reference price and the satisficing brand choice, four different measures were employed. It is important to realize that these models belong to three different class of models – substantive,

[8]The idea that consumers, while making a purchase decision consider an optimal number of brands, has its roots in information economic theory (Stigler 1961). This argument can be viewed as being equivalent to the concepts of sampling theory which calculates the number of observations required to make inferences about the population with a certain amount of confidence. Simon (1978) points out that though this might explain the reason for the existence of the latent construct of consideration sets, it violates the assumptions of bounded rationality. The number of steps in the decision making are increased and not decreased by viewing the decision process as a two stage process. Not only does a consumer have to make decisions about which brand to choose, but also how many brands to consider at any given time. Thus, although the explanation of information economics might have explanatory power, it lacks procedural rationality. The end result may make sense, but the process outlined by the theory cannot even approximately describe the process being used by consumers to make decisions.

processual, and procedural rational models. Erdem (1993) specifies the model proposed by Guadagni and Little (1983) (HGL is a variant of the original G & L model) as an "approximation to the reduced form" model. As G& L model assumes unbounded rationality, I refer to this as a substantive rational model. The reference price model is a variant of the original model proposed by Winer (1986) and tries to understand the process by which consumers use price to make decisions. Although it attempts to understand the process of making choice, thus making it a processual model, it still assumes unbounded rationality in the form of consumer ability to use all available information to make decisions. The satisficing model goes beyond the processual models as it incorporates computational limitations of consumers in the form of activation functions. Probabilistically, at any given time only certain bits of information are active and are, therefore, being used.

Table seven provides a comprehensive summary of the evaluation of the three models on these four criterion - AIC, HQ, BIC and CAIC (Elrod and Keane 1995). The log-likelihoods for the HGL, reference price and satisficing models are provided in the table.

AIC is defined as :

$$AIC=-2*LL+2*R, \tag{7.10}$$

where R is the number of parameters estimated in the model. A major limitation of this criterion is that it suffers from the phenomenon termed dimension-inconsistency (Elrod and Keane 1995) and, thus, is not able to identify the most parsimonious model.

To remove this limitation of the AIC criterion Schwarz (1978) proposed the BIC that takes into account the number of observations in the data set. BIC is defined as follows:

$$BIC=-2*LL+R*\ln(N), \tag{7.11}$$

where R is the number of parameters and N the number of observations in the data set.

Perhaps the criteria that imposes maximum penalty for additional parameters is the Consistent Akaike Information Criterion (CAIC) developed by Bozdogan (1987) and is defined as follows:

$$CAIC=-2*LL+R*\ln(N)+R \tag{7.12}$$

Elrod and Keane (1995) contend that both BIC and CAIC over penalize the log-likelihood function and argue that the HQ measure derived by Hannan and Quinn (1979) is the ideal measure for model selection. Their argument is based on the fact that HQ penalizes the log-likelihood for an extra penalty by a minimum value required to preserve the property of dimensional consistency. Dimensional consistency is a necessary requirement of the criteria used for model selection. HQ is defined as follows:

$$HQ=-2*LL+2*R*\ln(\ln(N)) \tag{7.13}$$

Using different criteria for model selection often leads to accepting different models. Though this is true for a number of studies conducted in marketing (Erdem

1993; Elrod and Keane 1995), table seven shows that on all four selection criteria the satisficing model does better than the HGL and the reference price model.

Insert Table 7 about here

Table Four

Heterogeneous Logit Model

Parameters	Estimates (Std-Error)
L1 (Heinz)	2.24 (.054)
L2 (Del Monte)	1.25 (.083)
L3 (Hunts)	1.65 (.065)
No Choice	4.58 (.055)
$\overline{\alpha_p}$ (Price)	-1.45 (.026)
α_{std} (State Dependence)	1.02 (.096)
$\overline{\alpha_a}$ (Advertising)	1.80 (.163)
α_{PI} (Purchase Timing)	-1.32 (.028)
σ_p	1.041 (.028)
σ_a	.0032 (.220)*
$\overline{W}$	.00018 (.017)*

* Not significant at .05 level of significance

Table Five

Reference Price Model

Parameters	Estimates (Std-Error)
L1 (Heinz)	2.40 (.010)
L2 (Del Monte)	-.68 (.200)
L3 (Hunts)	1.41 (.140)
No Choice	4.59 (.130)
$\overline{\alpha_P}$ (Gain)	2.05 (.914)
λ_P	-1.31 (.370)***
α_{std} (State Dependence)	1.03 (.017)
$\overline{\alpha_a}$ (Advertising)	1.87 (.176)
α_{Pt} (Purchase Timing)	-1.17 (.112)
σ_p	3.83 (.347)
σ_a	.0064 (.017)*
$\overline{W}$	.00018 (.017)*
λ	.942 (1.302)**

*not significant at .05 level of significance

** not significantly different from 1

*** not significantly different from -1

Table Six

Satisficing Model

Parameters	Estimates (Std-Error)
L1(Heinz)	2.42 (.149)
L2 (Del Monte)	-1.33 (.123)
L3 (Hunts)	-1.65 (.118)
No Choice	4.00 (.146)
α_R (Reference Price)	5.92 (.491)
α_P (Price)	1.28 (.490)
α_{std} (State Dependence)	1.40 (.056)
α_a (Advertising)	4.22 (.489)
α_{Pt} (Purchase Timing)	-1.03 (.200)
$\overline{W}$	.00009 (.035)*
γ_o	.87 (.227)
γ_I	-1.15 (.649)
γ_T	-0.505 (.193)
γ_D	3.75 (1.741)
λ	.99 (.019)**

*Not Significant at .05 level.

**Not Significantly different from 1.

Table Seven

Goodness of Fit Criteria

	Satisficing Model	Reference Price	HGL
Log-Likelihood	-2780.68	-3370.4	-3328.8
No. of Parameters	15	13	11
AIC	5591.36	6766.8	6657.6
BIC	5699.81	6860.5	6758.91
CAIC	5714.81	6873.5	6769.91
HQ	5628.03	6798.26	6705.02

Chapter 8

Discussions and Conclusions

8.1 General Discussions and Implications

Though it is heartening to see that the satisficing model has a better explanatory power than the rational HGL model, it is important to remember that the true purpose of structural models is to understand and describe choice processes. For example, the satisficing model assumes an inability of consumers to use all available information to make decisions. This assumption can be captured in a number of different ways such. These different ways are seen in the model proposed by Payne, Bettman and Johnson (1988), and Roberts and Lattin (1991). In this thesis, neither one of these two methods is used to capture limitations on consumer decision making. These methods attempt to capture the scarcity of human thinking as a technological and not a psychological construct. This assumption of scarcity of mind being a technological limitation is at odds with the theory of bounded rationality.

In this thesis a more plausible process of consumer thinking process is proposed. Borrowing ideas from Rumelhart and McCleland (1986), this work implements the psychological limitations of the human mind by the use of activation mechanisms. It argues that certain bits of information can be activated more than others, hence will be used by consumers to make decisions. Although determining the optimal consideration set size (Hauser and Wernerfelt 1990) and symbolic manipulations required for heuristic decision making can be costly human thinking processes, it can be argued that the spread of activation is not expensive and can be achieved in parallel (Anderson 1993). In addition, the activation mechanism mitigates the computational complexity of the process based heuristic models. Of the several heuristics that are available to a consumer it is not certain which will be used by a consumer to make decisions. Furthermore, in an uncertain environment it is not certain how consumers decide which heuristics to use as the costs and benefits of using different rules are not known (Meyer and Kahn 1991).

Activation models have the advantage of not requiring the information on cost and benefits of using different heuristic rules. Given the psychological limitations, knowledge of the environment in which the decision is being made is enough to predict the bits of information that will become salient for any decision. The variables that impact activation can often be determined from theoretical development in the area of information processing (Newell and Simon 1972; Simon 1978; Simon 1976), connectionism (Anderson 1976, 1983, 1993) and memory (Johnson-Laird 1983, Anderson 1976).[9] Infact, the satisficing model is the only model in marketing that attempts to understand and model consumer choice behavior under uncertainty.[10] The activation mechanism can be thought of as capturing either the

---

[9] This argument comes from the views proposed by Rumelhart and McCleland (1986). They argue that symbolic manipulations leading to models of heuristic decision rules are macroscopic accounts, analogous to Newtonian mechanics, whereas activation models offer more microscopic accounts, analogous to quantum theory. In a broad range, just as the two theories of physics -Newtonian mechanics and quantam theories, both the symbolic and activation theories are able to predict behavior of individuals. However, similar to Newtonian mechanics, heuristic theory breaks down. One such example is the choice of decision rule used by consumers to make decisions in an uncertain environment where the cost and benefit of using different decision rules are not known. It is for an understanding of these situations that a more microscopic theory is required and is provided by the activation theory.

[10] Economic theory frequently uses terms "risk" and "uncertainty" interchangeably (Erdem 1993). It is important to distinguish between these two concepts. When the moments of distributions are known with certainty, then the decision is being made under risk. On the other hand, uncertainty can stem from a number of different sources. Some of them may be as follows: (1) Consumers may not be aware of the different alternatives available to them at any given time; (2) Consumers may not have the ability to use all the available information to make decisions; (3) Consumers making decisions in an absolutely new environment have no idea of the moments of distribution of the variables that might affect decisions. All these conditions render economic models of utility maximization under risk highly inappropriate to capture consumer decision making. The satisficing model captures the uncertainty of the

uncertainty or the risk involved in making a decision. The lower the activation, the more risky the choice alternative and, consequently, the less the derived utility from the alternative. A small positive or a large negative brand intercept would mean that (1) a brand (alternative) is unfamiliar and may not be a part of the consideration set, or (2) that the consumer does not have the ability to use all the available information and is not able to consider an alternative. Hence, brand loyalty or brand familiarity not only increases the probability of choice, as measured by purchase feedback, but also decreases the perceived risk by increasing activation with the result that information pertaining to a brand is brought into the working memory. As bits of information are competing with each other for attention, activation not only brings information into working memory but also inhibits information about other brands becoming available at the time of making decisions. This mechanism of inhibition can substantially increase the probability of choice as the probability of choosing a brand is largely dependent upon the number of brands in the consideration set. Thus, it is in the interest of brand managers to make certain that the number of brands a consumer is considering is kept to as few as possible, and that their brand is part of this consideration set.

There are a number of different ways of accomplishing this. One way is to increase brand name awareness. Another way is to decrease perceived risk related to brand choice by positive brand associations and increased perceived quality. A third alternative is to increase brand loyalty by offering high perceived quality. All these variables in effect lead to higher brand equity. In another study it has been suggested by Erdem (1993) that perceived risk (in this model perceived risk is being captured

---

form described by (1) and (2) and has the ability to capture the uncertainty of the form described in (3).

by the activation function) is related to brand equity, and that decrease in perceived risk leads to an increase in brand equity. Aaker (1991) writes that brand equity can be increased by: brand name awareness, brand loyalty, perceived quality and brand associations. The activation function that captures different components of brand equity[11] shows that a brand that has low brand equity will have to use price promotions to attract consumers.

8.2 Future Research

A number of issues have been left for future research. Certain aspects of cognitive modeling, such as declarative and procedural memory, have not been distinguished in the current model. Declarative memory is the knowledge of facts about the world. On the other hand, procedural knowledge is the knowledge about how to do something. In the models proposed in this thesis the activation functions can be described as production systems (Newell and Simon 1972; Anderson 1976) that attempt to capture procedural knowledge. As the parameters of the activation function are constant over time, an implicit assumption of the model is that of no learning, that is, experience does not change the interaction of memory with behavior. This assumption is a major limitation of the models proposed in this dissertation. An advantage of using activation models lies in their ability to learn and unlearn concepts from the environment. As the focus of research in marketing today[12] is in establishing the long term effects of marketing mix variables, consumer brand choice models that attempt to understand the process of choice can play an important part in doing so.

[11]Positive brand associations and brand awareness are measured by the intercept term of the activation function. The value of the activation function itself measures the perceived risk of the attribute.

[12]A special session on long term impact of marketing mix variables on market share was held at the 1996 Marketing Science Conference, Gainesville, March 7-10.

For example, Erdem (1993) showed that by using structural models a researcher could understand the reason for insignificant advertising effect, that is, not only can one study effect but also the reason for that effect. Activation models go a step further in that not only do they capture a more realistic process of choice but they also describe it. In addition, relationships among different variables and the activation function represent learning effects over time. The parameters for the activation function measure the association of different variables with the underlying construct. Learning about brands would be captured by the updating of parameters. Bayesian techniques should be useful in capturing this learning process.

Another major limitation of the models proposed in this thesis is their computational complexity that allows only a few brands and attributes to be included in the study. Even for fast machines (workstations and supercomputers) the estimation problem is often an infeasible exercise. An avenue for further research could be in the development of new algorithms that would substantially reduce the required run time. Taylor series expansion of the likelihood function provides a promising avenue for such future research.

Structural modeling can easily be extended to incorporate the process of forming price expectations. The uncertainty of the price expectation is modeled as an individual specific normal error time that is constant over time. One can easily relax this assumption. In addition, these models can be used to study the effects of frequency and intensity of price promotions on price expectations.

8.3 Conclusion

This thesis accomplished three purposes. First, borrowing concepts from bounded rationality and cognitive psychology it introduced a framework that incorporates the impact of risk and uncertainty, consumer price expectation, and consideration sets on consumer brand choice. In addition, it provides a plausible explanation for phenomenon such as promotional signal sensitivity that have often been observed in the market environment (Inman and Mcalister 1993). Secondly, it shows a relationship among brand equity, perceived risk, brand loyalty, brand awareness and promotional signal. Finally, it has developed brand choice models that are based on cognitive theory and thus, has enabled researchers to understand the process of choice. To the best of my knowledge, this is the first procedural rational structural dynamic brand choice model that has several applications and implications. The model developed in this thesis attempts to break new ground by providing a framework for capturing the adaptive character of human decision making without resorting to a cost-benefit analysis.

References

Aaker, David A. (1991), *Managing Brand Equity: Capitalizing on the Value of a Brand*, New York: Macmillan.

Allenby, Greg and Peter Lenk (1994), "Reassessing Brand Loyalty, Price Sensitivity and Merchandising Effects on Consumer Brand Choice," *Working Paper*, Ohio State University.

Anderson, John R (1976), *Language, Memory, and Thought*, Hillsdale : Lawerence Erlbaum Associates.

______________ (1983), *The Architecture of Cognition*, Cambridge: Harvard University Press.

______________ (1986), *Cognitive Psychology and its Applications*, New York: W. H. Freeman and Company.

______________ (1993), *The Adaptive Character of Thought*, Hillsdale: Lawerence Erlbaum Associates.

______________ and G. H. Bower (1973), *Human Associative Memory*, Washington, DC: Winston and Sons.

Andrew, Rick L. and T. C. Srinivasan (1995), "Studying Consideration Effects in Empirical Choice Models Using Scanner Panel Data," *Journal of Marketing Research*, 32, 30-41.

Assuncao, Joao and Robert J Meyer (1990), "The Rational Effects of Price Expectations on Sales-Price Relationship," *Working Paper*, John E. Anderson Graduate School of Management, University of California, Los Angeles.

Bass, Frank M. (1974), "The Theory of Stochastic Preference and Brand Switching," *Journal of Marketing Research,* 11, 1-20.

__________, Abel P. Jeuland and Gordon P. Wright (1976), "Equilibrium Stochastic Choice and Market Penetration Theories: Derivations and Comparisons," *Management Science*, 22, 1051-1063.

__________ and Thomas L. Pilon (1980), "A Stochastic Brand Choice Framework for Econometric Modeling of Time Series Market Share Behavior," *Journal of Marketing Research*, 17, 486-497.

Bell, D. E. (1982), "Regret in Decision Making Under Uncertainty," *Operations Research,* 30, 961-981.

Bell, David R. and James M. Lattin (1993), "Loss Aversion and Heterogeneity in Price Sensitivity," *Working Paper*, Standford University, Stanford, CA.

Bettman, James R. (1971), "Information Processing Models of Consumer Behavior ," *Journal of Marketing Research*, 7, 370-376.

__________ (1979), *An Information Processing Theory of Consumer Choice*, Reading, MA: Addison Wesley.

__________, Eric J. Johnson and John W. Payne (1990), "A Componential Analysis of Cognitive Effort in Choice, " *Organizational Behavior and Human Decision Prcocesses*, 45, 111-140.

__________, Eric J. Johnson and John W. Payne (1991), "Consumer Decision Making," *in Handbook of Consumer Behavior*, editors Thomas. S. Robertson and Harold H. Kassarjan, 50-84, New Jersey: Prentice-Hall.

Biswas, Abhijit and Edward A. Blair (1991), "Context Effects of Reference Price in Retail Advertisements," *Journal of Marketing*, 55, 1-12.

____________, Elizabeth J. Wilson and Jane W. Licata (1993), "Reference Pricing Studies in Marketing: A Synthesis of Research Results," *Journal of Business Research*, 27, 239-256.

Bruner, J. S. (1957), "On Perceptual Readiness," *Psychological Review*, 64, 123-152.

Bucklin, Randolph E. and James M. Lattin (1991), "A Two State Model of Purchase Incidence and Brand Choice," *Marketing Science*, 10, 24-39.

Burrel, Q. L. (1980), "A Simple Stochastic Model for Library Loans," *Journal of Documentation*, 41, 115-132.

__________ (1985), "A Note on Aging on a Library Circulation Model," *Journal of Documentation*, 41, 100-115.

Chiang, J. (1991), "A Simultaneous Approach to the Whether, What and How Much to Buy Questions," *Marketing Science*, 10, 297-315.

Chintagunta, Pradeep (1993), "Investigating Purchase Incidence, Brand Choice and Purchase Quantity Decisions of Households," *Marketing Science*,12, 184-212.

Chomsky, N. (1957), *Syntactic Structure*, The Hague: Norton.

__________ (1968), *Language and Mind*, New York: Harcourt Brace and World.

Corstjen, Marcel L. and David A. Gatuschi (1983), "Formal Choice Models in Marketing," *Marketing Science*, 2, 19-56.

Currim, Imran S., Robert J. Meyer and Nhan Le (1988), "Disaggregate Tree Structured Modeling of Consumer Choice Data," *Journal of Marketing Research*, 25, 253-265.

Dickson, Peter and A. G. Sawyer (1990), "The Price Knowledge and Search of Supermarket Shoppers, " *Journal of Marketing Research*, 54, 42-53.

Eckstein, Z., Dan Horsky and Yeol Raban (1988), "An Empirical Dynamic Model of Optimal Brand Choice," *Mimeo. Tel Aviv*, Tel Aviv University.

Elrod, Terry (1988), "Choice Map: Inferring a Product Market Map from Observed Choice Behavior, " *Marketing Science*, 7, 21-40.

__________ (1991), "Internal Analysis of Market Structure: Recent Developments and Future Prospects," *Marketing Letters*, 2, 253-266.

__________ and Michael P . Keane (1995), "A Factor Analytic Model for Representing the Market Structure in Panel Data," *Journal of Marketing Research*, 32, 1-16.

Emory, Fred (1970), "Psychological Aspects of Price," in *Pricing Strategy*, editors Bernard Taylor and Gordon Wills, Princeton, NJ: Brandon Systems, 98-111.

Erdem, Tulin (1993), "Dynamic Brand Chòice and Product Market Signalling under Imperfect Information," *Unpublished Manuscript*, University of Alberta.

Gijsbrechts, Els (1993), "Prices and Pricing Research in Consmer Marketing: Some Recent Developments, " *Internation Journal of Research in Marketing*, 10, 115-151.

Gonul, Fusun and K. Srinivasan (1993), "Modeling Unobserved Heterogeneity in Multinomial Logit Models: Methodological and Managerial Implications," *Working Paper*, Carnegie Mellon University.

Gould, Stephen J. (1980), *The Panda's Thumb*, Chicago: Norton Press.

Grether and Louis Wilde (1984), "An analysis of Conjunctive Choice: Theory and Experiments," *Journal of Consumer Research*, 10, 373-385.

Grossberg, Stephen and William E. Gutowski (1987), "Neural Dynamics of Decision Making Under Risk: Affective Balance and Cognitive-Emotional Interactions," *Psychological Review*, 94, 300-318.

Guadagni, Peter and John D. C. Little (1983), "A Logit Model of Brand Choice Calibrated on Scanner Panel Data," *Marketing Science*, 2, 203-238.

Gupta, Sunil (1988), "The Impact of Sales Promotion on When, What and How Much to Buy," *Journal of Marketing Research*, 25, 342-355.

__________ (1991), " Stochastic Models of Inter-Purchase Time and Time Dependent Covariates," *Journal of Marketing Research*, 28, 1-15.

Hardie, Bruce G. S., Eric J Johnson and Peter S. Fader (1993), "Modeling Loss Aversion and Reference Dependence in Brand Choice," *Marketing Science*, Forthcoming.

Hauser, John and Birger Wernerfelt (1990), "An Evaluation Cost Model of Consideration Sets," *Journal of Consumer Research*, 16, 393-408.

Helsen, Kristiaan and D. C. Schmittlein (1993), "Analyzing Duration Times in Marketing Research," *Working Paper*, Wharton School, Philadelphia.

Hogarth, Robin (1980), *Judgement and Choice, New York* : John Wiley Publication.

Huber, Joel and Noreen Klein (1988), "Choosing Which to Choose: A Model of Cutoff Selection in a Two-Stage Decision Process," *Working Paper*, Fuqua School of Business, Duke University, Durham, NC.

Inman, J. Jeffery and Leigh Mcalister (1993), "A Retailer Promotion Policy Model Considering Signal Sensitivity," *Marketing Science*, 12, 339-357.

____________ and ____________(1994), "Do Coupon Expiration Dates Affect Consumer Behavior," *Journal of Marketing Research*, 31, 423-428.

Jacobson, Robert and Carl Obermiller (1990), "The Formation of Expected Future Price: A Reference Price for Forward Looking Consumers," *Journal of Consumer Research*, 16, 420-432.

Johnson, Eric J. and John W. Payne (1985), "Effort and Accuracy in Choice," *Management Science*, 31, 395-414.

Johnson-Laird, P. N. (1983), *Mental Models*, Cambridge, MA: Harvard University Press.

Jones, J Morgan and Jane T. Landwehr (1988), "Removing Heterogeneity Bias from Logit Model Estimation," *Marketing Science*, 7, 41-59.

Kahneman, Daniel and Amos Tversky (1979), "Prospect Theory: An Analysis of Decision Making Under Risk," *Econometrica*, 47, 263-291.

Kahn, Barbara E., William L. Moore and Rashi Glazer (1987), "Experiments in Constrained Choice," *Journal of Consumer Research* , 14, 96-113.

Kalwani, Manohar U., Chi Kin Yin, Heikki J. Rinne and Yoshi Sugita (1990), "A Price Expectations Model of Customer Brand Choice," *Journal of Marketing Research*, 27, 251-262.

______________ and ________ (1992), " Consumer Price Promotion Expectations: An Experimental Study," *Journal of Marketing Research*, 29, 90-100.

Kalyanaraman, G. and J. D. C. Little (1989), "A Price Response Model Developed from Perceptual Theories," *Working Paper*, MIT.

______________ and Russell S. Winer (1995), "Empirical Generalizations from Reference Price Research," *Marketing Science*, 14, G161-169.

Kamakura, Wagner and G. J. Russell (1989), "A Probabilistic Choice Model for Market Segmentation and Elasticity Structure," *Journal of Marketing Research*, 26, 379-390.

Katona, George (1960), *The Powerful Consumer*, New York: McGraw-Hill.

Keane, Michael (1995), "Modeling Heterogeneity and State Dependence in Consumer Choice Behavior," *working paper*, University of Minnesota.

Klein, Noren M. and Janet E. Oglethorpe (1987), "Cognitive Reference Points in Consumer Decision Making," in *Advances in Consumer Research*, editors Melanie Wallendrof and Paul Anderson, 14, 107-128.

Koslow, Scott (1990), "A Theory of Deterministic Consumer Choice Behavior: Applying Generalizability Theory of Measurement to Consumer Panel Data, " *Unpublished Doctoral Dissertation*, University of Southern California.

Krishna, Aradhna (1992), "The Normative Impact of Consumer Price Expectations for Multiple Brands on Consumer Purchase Behavior," *Marketing Science*, 11, 266-287.

Krishnamurthi, Lakshman and S. P. Raj (1988), "A Model of Brand Choice and Purchase Quantity Price Sensitivities," *Marketing Science*, 7, 1-20.

Kunreuther, Howard (1978), *Protecting Against High Risk Hazards: Public Policy Lessons*, New York: James Wiley Publication.

Lattin, James M. and Randolph E. Bucklin (1989), "Reference Effects of Price and Promotion on Brand Choice Behavior," *Marketing Science*, 11, 266-287.

Lehmann, Donald and William L. Moore (1989), "A Combined Simply Scalable and Tree Based Preference Model," *Working Paper*, Columbia University.

Luce, R. Duncan (1959), *Individual Choice Behavior*, New York: Wiley.

Margolis, Howard (1987), *Patterns, Thinking and Cognition*, Chicago: University of Chicago Press.

Mcfadden, Daniel (1986), "The Choice Theory Approach to Marketing Research," *Marketing Science*, 5, 275-297.

_____________ (1989), "A Method of Simulated Moments for Estimation of Discrete Response Model without Numerical Intergration," *Econometrica*, 57, 995-1026.

_____________ (1991), "Advances in Computation, Statistical Methods, and Testing of Discrete Choice Models, " *Marketing Letters*, 2, 215-229.

McCleland, D. C. (1955), "The Psychology of Mental Content Reconsidered," *Psychological Review*, 62, 297-320.

Meyer, Robert J. and Arvind Sathi (1985), "A Multi-Attribute Model of Consumer Choice During Product Learning," *Marketing Science*, 4, 41-61.

____________ and J. Assuncao (1990), "The Optimality of Consumer Stock Piling Strategies," *Marketing Science*, 4, 41-61.

____________ and Barbara E. Kahn (1991), "Probabilistic Models of Consumer Choice," in *Handbook of Consumer Behavior*, editors Thomas. S. Robertson and Harold H. Kassarjan, 85-123, New Jersey: Prentice-Hall.

Miller, G. A. (1956), "The Magical Number Seven, Plus or Minus two: Some Limits on Our Capacity for Processing Information," *Psychological Review*, 63, 81-97.

Monroe, Kent B. (1990), *Pricing: Making Profitable Decisions*, New York: Harper and Row.

Moore, William L., Donald R. Lehmann and Edgar A. Pessemier (1989), "A Paired Comparison Nested Logit Model of Individual Preference Structures," *Journal of Marketing Research*, 26, 420-428.

Neslin, S. A., C. Henderson and J. Quelch (1985), "Consumer Promotions and the Accelaration of Product Purchases," *Marketing Science*, 4, 147-165.

Newell, Alan (1991), *Unified Theories of Cognition*, Cambridge: Harvard University Press.

__________ and Herbert A. Simon (1972), *Human Problem Solving*, New Jersey: Prentice Hall.

Oliver, Richard L. and Russell S. Winer (1986), "Formation of Consumer Expectations," *Working Paper*, Vanderbuilt University.

Pakes, Ariel and David Pollard (1989), "Simulation and the Asymptotics of Optimization Estimates," *Econometrica*, 57, 1027-1058.

Payne, John W. , James R. Bettman and Eric J. Johnson (1988), "Adaptive Strategy Selection in Decision Making," *Journal of Experimental Psychology: Human Learning, Memory and Cognition*, 14, 534-552.

Popkowski Leszczyc, Peter (1994), "The Effect of Competitive Promotions on Brand Choice and Purchase Timing," *paper presented at Marketing Science Conference*, Tucson, Arizona.

Quillian, M. R. (1969), "The Teachable Language Comprehender," *Communications of the Associations for Computing Machinery*, 12, 459-476.

Restle, F. (1961), *Psychology of Judgement and Choice: A Theoretical Essay*, New York: James Wiley.

Roberts, John H. and Glen Urban (1988), "Modeling Multiattribute Utility, Risk and Belief Dynamics for New Consumer Durable Brand Choice, " *Management Science*, 34, 167-185.

____________ and James M. Lattin (1991), "Development and Testing of a Model of Consideration Set Composition," *Journal of Marketing Research*, 28, 429-440.

Rumelhart, David E. and James L. McCleland (1986), *Parallel Distributed Processing*, Cambridge, MA: MIT Press.

Russo, J. E. and B. A. Dosher (1983), "Strategies for Multiattribute Binary Choice," *Journal of Experimental Psychology, Learning, Memory and Cognition*, 9, 676-696.

Rust, John (1987), "Optimal Replacement of GMC Bus Engines: An Empirical of Harold Zurcher," *Econometrica*, 55, 999-1035.

Shocker, Allan D., Moshe Ben-Akiva, Bruno Boccara and Prakash Nedungadi (1991), "Consideration Set Influences on Consumer Decision Making and Choice: Issues, Models and Suggestions," *Marketing Letters*, 2, 181-197.

Shugan, Steve M (1980), "The Cost of Thinking," *Journal of Consumer Research*, 7, 99-111.

Simon, Herbert A. (1955), "A Behavioral Model of Rational Choice," *Quaterly Journal of Economics*, 69, 99-118.

______________ (1976), *Administrative Behavior*, New York: Macmillan.

______________ (1977), *Models of Discovery*, Boston: D. Riedel Publishing Company

______________ (1978), "Rationality as Process and as Product of Thought," *American Economic Review Proceedings*, 68, 1-16.

______________ (1982), *Models of Bounded Rationality*, Cambridge, MA: MIT Press.

Smolensky, P (1986), "Information Processing in Dynamical Systems: Foundations of Harmony Theory," in *Parrallel Distributed Processing*, editors David A. Rumelhart and James L. McCleland, 194-282, Cambridge, MA : MIT press.

Stigler, George (1961), "The Economics of Information," *Journal of Political Economy*, 69, 213-225.

Stritter, E. P. (1977), "File Migration," *Unpublished Doctoral Dissertation*, Standford University, Stanford, CA.

Supan, Axel-Bosch and Vassilis A. Hajivassiliou (1993), "Smooth Unbiased Multivariate Probability Simulators for Maximum Likelihood Estimation of Limited Dependent Variable Models," *Journal of Econometrics*, 58, 347-369.

Tellis, G. J. (1988), "Advertising Exposure, Loyalty and Brand Exposure: A Two Stage Model of Choice," *Marketing Science*, 4, 147-165.

Thaler, Richard (1985), "Mental Accounting and Consumer Choice," *Marketing Science*, 4, 199-214.

Theil, H (1954), "Econometric Models and Welfare Maximization," *Weltwirtschaftliches Archiv.*, 72, 60-83.

Tversky, Amos (1972), "Elimination by Aspects: A Theory of Choice," *Psychological Review*, 79, 281-299.

_____________ and Shumel Sattah (1979), "Preference Trees," *Psychological Review*, 86, 542-573.

Urbany, Joel E., William O. Bearden and Dan C. Weilbaker (1988), "The Effect of Plausible and Exaggerated Reference Prices on Consumer Perceptions and Price Search," *Journal of Consumer Research*, 15,95-110.

Van Raaij, W. F. (1977), *Consumer Choice Behavior: An Information Processing Approach*, Voorschoten, Netherlands: VAM.

Winer, Russell S. (1985), "A Reference Price Model of Demand for Consumer Durables: Preliminary Development," *Marketing Science*, 4, 74-90.

_____________ (1986), "A Refernce Price Model of Demand for Frequently Purchased Products," *Journal of Consumer Research*, 13, 250-256.

Appendix 1

Gauss Source Code of the Program Used for the Estimation of the Satisficing Model

```
closeall;
library maxlik;
#include maxlik.ext;
maxset;
closeall;

/*
Data Set y5.out Contains Random Generates From a Standard Normal Distribution
*/

load y51=y5.out;
y5=y51[.,1];
j12=1;


_max_MaxIters=80;
_max_CovPar=1;
_max_Algorithm=6;
_max_Active={1,1,1,1,1,1,1,1,1,1,1,1,1,0,1,0,1};

/*
 Parameters for Maxlik
*/

x0={2.61,-1.34,-1.73,3.86,6.70,1.38,4.98,-.02,1.05,-.25,-.21,.49,2.12,
0.00,-1.49,0,-.51
};

/*
Call to Maxlik
*/

{x,f,g,cov,rec}=maxlik("logi",0,&likfun,x0);
output file=hrf81.out reset;
call maxprt(x,f,g,cov,rec);
output off;
```

closeall;

/*
Global Parameters Used in the Program
*/

```
clearg t1,t2,t3,t4,t5,t6,t7,y6,x1,g,cov11,
e,j,d,i1,i2,i3,i4,i5,numhh,y11,lik,lik1,prob1,l1,l2,l3,l4,tot2,
t31,pr1,pr2,pr3,lk1,lk2,lk3,t11,y7,y71,e11,e21,e31,t51,t41,t63,t61,t62;
```

/*
Likelihood Function
*/

```
PROC likfun(beta,d22);
local r,denov,logv,logiv,hid,i,y1,y2,x,
        ii,step,area,are,ha,hb,xx,tot,yy,f1,
        gam,jk,y8,y10,t8,y9,x2,x3,t20,x33
        ,price,adv,loyal,temp,jt,cov22,cov33,cov44,cov,
        g1,jt,d22,hess1,logiv1,tot1,cov55,b11,b1g,b2l,b2g,b3l,b3g,rf,b21,b22,b23,
        e1,e2,e3,t10,t32,b1,b2,b3,kappa,va,ve,jk1,t19,cov2,v1,v2,v3,i6,y111,lk,l11
        ,l21,l31,l41,i7,y119,y200,del;
closeall;
        price=-beta[17];
        y6=y5+(beta[8])^2;
        y71=y6*(beta[1]~beta[2]~beta[3]);
        y119=beta[12];
it=1;
j=1;hid=1;lik=0;numhh=1;
closeall;
```

/*
Data Set x.dat Has the Combinations of the Possible Bits of Information That Could Have Been Used by Individuals to Make Decisions. There are 8 (2^3) Possibilities. The Data Set has 3 Rows and 8 Columns. TheTransposed Data Set is Provided Below:

State	Heinz	Del Monte	Hunts
1	0	0	0
2	1	0	0
3	0	1	0

4	1	1	0
5	0	0	1
6	1	0	1
7	0	1	1
8	1	1	1

0 ---- Represents That Attributes of a Brand is Not Being Used by an Individual to Make Decisions

1----- Represents That Attributes of a Particular Brand is Being Used by an Individual to Make Decisions

```
*/
        load x[3,8]=x.dat;

/*
The Data Set. There are 10200 Observations and 100 Households.
*/

        open f1=ash59.dat;
j12=1;
lik1=0;
do while numhh <= 100;
        xx=readr(f1,1);

/*
tot1 -- Total Number of Observations Per Household
*/

        tot1=xx[1,34];
        yy=seekr(f1,j);
        d4=readr(f1,tot1);
        area=0; logv=0;i=1;ii=1;temp=0;
tot=tot1-1;


if tot1>1;
        d=d4[2:tot1,.];
        j13=j12+39;
        y7=y71';
        y200=y119[.,j12:j13];
        l1=zeros(tot,1);
```

```
        l2=zeros(tot,1);
        l3=zeros(tot,1);
        l4=zeros(tot,1);
i=1;

/*
Setting up Variables for Measuring State Dependence
*/

do while i<=tot;
        if d[i,28]==1;
        l1[i,1]=1;
elseif d[i,28]==2;
        l2[i,1]=1;
elseif d[i,28]==3;
        l3[i,1]=1;
else;
        l4[i,1]=1;
endif;
i=i+1;
endo;

/*
These Are The Bernoullis
*/

b1=exp(y200+beta[9]*d[.,4]+beta[10]*d[.,32]+beta[11]*y7[1,.])./
        (1+exp(y200+beta[9]*d[.,4]+beta[10]*d[.,32]+beta[11]*y7[1,.]));
b2=exp(y200+beta[9]*d[.,5]+beta[10]*d[.,32]+beta[11]*y7[2,.])./
        (1+exp(y200+beta[9]*d[.,5]+beta[10]*d[.,32]+beta[11]*y7[2,.]));
b3=exp(y200+beta[9]*d[.,6]+beta[10]*d[.,32]+beta[11]*y7[3,.])./
        (1+exp(y200+beta[9]*d[.,6]+beta[10]*d[.,32]+beta[11]*y7[3,.]));


/*
Capturing Uncertainty Due to Decision/ No Decision
*/

del=.5;

/*
 Exponential Smoothing Function For Reference Price
*/
```

```
lk=exp(beta[13])/(1+exp(beta[13]));

/*
 Capturing the Effect of Attention on Updating of the Reference Price
*/

lk1=lk.*b1;
lk2=lk.*b2;
lk3=lk.*b3;

/*
Reference Price
*/

pr1=zeros(tot,40);pr2=zeros(tot,40);pr3=zeros(tot,40);
do while i<=tot;
if i ==1;

/*
 Initial Values for Reference Price
*/
        pr1[i,.]=pr1[i,.]+d4[i,12];
        pr2[i,.]=pr2[i,.]+d4[i,13];
        pr3[i,.]=pr3[i,.]+d4[i,14];
else;
        pr1[i,.]=(d[i,12].*lk1+pr1[i-1,.].*(1-lk1));
        pr2[i,.]=(d[i,13].*lk2+pr2[i-1,.].*(1-lk2));
        pr3[i,.]=(d[i,14].*lk3+pr3[i-1,.].*(1-lk3));
endif;
        i=i+1;
endo;
tot2=tot1-1;

        i1=zeros(tot2,1);
        i2=zeros(tot2,1);
        i3=zeros(tot2,1);
        i4=zeros(tot2,1);
        i5=zeros(tot2,1);
        i6=zeros(tot2,1);

i=1;
```

```
/*
Dependent Varaibales
*/

do while i <= tot2;
        if d[i,30]==1;
i1[i,1]=1;
        elseif d[i,30]==2;
i2[i,1]=1;
        elseif d[i,30]==3;
i3[i,1]=1;
        elseif d[i,30]==4;
i4[i,1]=1;
        elseif d[i,30]==0;
i5[i,1]=1;
endif;
i=i+1;
endo;

/*
Loyalty Variable
*/

        l11=l1;
        l21=l2;
        l31=l3;
        l41=l4;

e11=zeros(tot2,1);e21=zeros(tot2,1);e31=zeros(tot2,1);i=1;w2=zeros(1,1);

/*
 Evaluating The Likelihood Function. This Loop Incorporates Heterogeneity Due to
Individual Tastes. It Also Accounts for Researcher's Lack of Knowledge of What
Information is Being Used By Consumers to Make Decision.

As There are Three Brands - Heinz, Hunts and Del Monte (The Utility For the Fourth
Brand Has Been Fixed to a Value of Zero) There Are In All Eight Possible
Combinations
*/

        do while i<=8;

/*
```

```
 w1 Measures the Probability that Individual i is in the s-th State at the t-th Time
Period, Given That a Decision is Being Made.
*/

        w1=del.*(x[1,i].*b1+(1-x[1,i]).*(1-b1)).*((x[2,i]).*b2+(1-x[2,i]).*(1-
              b2)).*(x[3,i].*b3+(1-x[3,i]).*(1-b3));

        w2=w2~w1[1,1];

/*
 Utility Function For Each Brand Given the State in Which the Decision is Being
Made */

        e1=y7[1,.]+(x[1,i]).*(d[.,12]*price+beta[5].*(pr1-
        d[.,12]))+beta[6]*l11+(x[1,i]).*beta[7].*d[.,17]+((1-x[1,i])).*pr1*price;

/*
Utility For Heinz
*/

        e2=y7[2,.]+(x[2,i]).*(d[.,13]*price+beta[5].*(pr2-
        d[.,13]))+beta[6]*l21+(x[2,i]).*beta[7].*d[.,18]+(1-x[2,i]).*pr2*price;

/*
Utility for Del Monte
*/

        e3=y7[3,.]+(x[3,i]).*(d[.,14]*price+beta[5].*(pr3-
        d[.,14]))+beta[6]*l31+(x[3,i]).*beta[7].*d[.,19]+((1-x[3,i])).*pr3*price;

/*
Utility for Hunts
*/

        t1=exp(e1);
        t2=exp(e2);
        t3=exp(e3);

/*
Numerator for  the Logit
*/

        t31=exp(beta[4]+beta[15]*d[.,32]);
```

```
/*
Numerator for No Choice Alternative
*/

        t4=(ones(tot2,40)+t1+t2+t3+t31);

/*
Denominator for the Logit
*/

        t5=(t1.*i1+t2.*i2+t3.*i3+i4+t31.*i5);
        t51=(t1.*i1+t2.*i2+t3.*i3+i4+t31.*i6);t41=(t1+t2+t3+ones(tot2,40)+t31);

        e21=e21~(w1.*(t51./t41));

/*
e21 Captures the Fact that It is Known with Certainty that Individual i is Making a
Decision, When a Brand is Bought in The Product Category.
e21 is the Probability of Individual i Purchasing the j-th Brand in the t-th Time
Period.
*/

        e11=e11~(w1.*(t5./t4));

/*
e11 is the Probability of Individual i Not Purchasing a Brand in the Market Given
That a a Consumer is Making a Decision in the Product Category
*/

i=i+1;
endo;

        t6=(sumc(e11')/40);
        t63=sumc(e21')/40;

/*
Incorporating Uncertainty Due to Decision/ No Decision on a Store Visit
*/

        t61=(i5.*(1-del)+(1-i5).*t63);

/*
```

```
Likelihood for the i-th Individual
*/

        t62=t6+t61;

/*
Log-Likelihood for the i-th Individual
*/

        prob1=ln(prodc(t62));
        lik1=lik1+prob1;
        j=j+tot1;
        j12=j12;
        numhh=numhh+1;
        else;
        j=j+tot1;
        numhh=numhh+1;
endif;
        endo;

/*
Return Value of Likelihood Function from the Procedure
*/

        retp(lik1);
endp;
```

Appendix 2

Simulation

To test the performance of the simulated maximum likelihood algorithm used for the estimation of the satisficing model, simulated data set with known true parameters were generated. A variant of the algorithm used in this thesis was then employed to recover the true parameters. The following specification for the four brand choice model was assumed:

$$U_{ijt} = L_j W_i + \pi_{ijt}\alpha_p p_{jt} + \pi_{ijt}\alpha_R(\overline{p^e_{ijt}} - p_{jt}) + (1-\pi_{ijt})\alpha_P \overline{p^e_{ijt}} + \pi_{ijt}\alpha_a a_{jt} + \varepsilon_{ijt} \qquad \text{(A.2.1)}$$

where,

ε_{ijt} is assumed to have a double extreme gumble distribution,

α's are the parameters with known values,

p is Price,

a is the Advertising,

i.....indexes for the consumer (i=1,......N),

j....indexes for the brand (j=1,.....J),

t.....indexes for the time period of store visits. (t=1,.....T),

π_{ijt} is the probability of consumer i attending to the attributes of the j-th brand at the t-th time period, and is given as follows:

$$\pi_{ijt} = \frac{\exp(\gamma_{0i} + \gamma_1(L_j W_i) + \gamma_2 * D_{jt} + \gamma_3 * Weeks_t)}{1+\exp(\gamma_{0i} + \gamma_1(L_j W_i) + \gamma_2 * D_{jt} + \gamma_3 * Weeks_t)} \qquad \text{(A.2.2)}$$

where,

D_{jt} = 1, if the j-th brand is on promotion in the t-th time period

= 0, otherwise, and,

p^e_{ijt} is the price expectations of the i-th consumer for the j-th brand at the t-th time period, which is given as follows:

$$P^e_{ijt} = (\lambda_o * (\pi_{ijt})^{\lambda_1}) * P_{jt}. \quad \text{(A.2.3)}$$

Also, the following specifications were assumed:

$$W_i \sim N(\overline{W},1) \quad \text{(A.2.4)}$$

and,

$$\gamma_{0i} \sim N(\overline{\gamma_o},\sigma_\gamma) \quad \text{(A.2.5)}$$

The utility for no purchase is defined as follows:

$$U_{i\bar{j}t} = \alpha_{np} + \alpha_{Pl} * Weeks_t + \varepsilon_{i\bar{j}t} \quad \text{(A.2.6)}$$

The data set has 1000 individuals over a period of 32 weeks, a total of 32000 observations. The parameters for the simulated data set has been provided in Table 8:

Insert Table 8 about there

With known parameters for the distributions specified in equations A.2.4 and A.2.5, W_i and γ_{oi} were generated from a multivariate normal distribution. Knowing the values of the exogenous variables given in equations A.2.1 and A.2.2., utility for each brand or that for no purchase was evaluated. The alternative with the highest utility was the individual choice for a given time period.

Table 9 reports the parameter estimates and the respective true values as well as the standard errors of the estimates. For identification of the model, the brand specific intercept is assumed to be equal to zero. The standard errors of the estimates show that only two of the sixteen parameters, $\overline{W}$ and α_{pl}, are biased at .05 level of significance. Also, the likelihood tests shows that the parameters estimated by the algorithms are not significantly different from the true parameters at .05 level of significance. Hence the algorithm does a reasonably well at recovering the true parameters of the data set.

Insert Table 9 about here

Table 8

True Values of the Paramaters

Parameters	**True Values**
L1	1.6
L2	2.0
L3	1.0
$\overline{W}$	1.0
α_p	-6.0
α_R	2.0
α_a	3.0
α_{np}	1.7
α_{PI}	-.7
$\overline{\gamma_o}$	1.0
σ_{γ_o}	1.6
γ_1	-3.0
γ_2	3.0
γ_3	1.0
λ_o	.85
λ_1	.5

Table 9

Estimates of the Simulated Model

Parameters	**Estimates**	**True Values**	**Standard Errors**
L1	1.5986	1.6	.0084
L2	1.9926	2.0	.0076
L3	1.0255	1.0	.0153
$\overline{W}$	1.0608*	1.0	.0124
α_p	-5.9950	-6.0	.0057
α_R	1.9991	2.0	.0052
α_a	3.0136	3.0	.0079
α_{np}	1.7123	1.7	.0076
α_{PI}	-.6753*	-.7	.0100
$\overline{\gamma_o}$	1.0004	1.0	.0053
σ_{γ_o}	1.6061	1.6	.0055
γ_1	-3.0075	-3.0	.0062
γ_2	2.9964	3.0	.0056
γ_3	.9933	1.0	.0057
λ_o	.8490	.85	.0053
λ_1	.5005	.5	.0052

Log-Likelihood at Estimated Values = -37009.2

Log-Likelihood at True Values = -37021.9

$\chi^2_{16} = -2 * (LL_{TrueValues} - LL_{EstimatedValues}) = 23.6$[13]

[13] Critical Value at .05 level of significance is 26.2

Appendix 3

Technical Appendix

This appendix provides the technical details of the method used to estimate the parameters of the satisficing model. As the likelihood function does not have a closed form solution, a simulated maximum likelihood (SML) procedure was used to estimate the parameters of the model. There are in all four discrete distributions and one continuous distribution that need to be simulated. In particular, the SML needs to integrate the informational uncertainty and heterogeneity out of the likelihood function. Details of the algorithm are provided below.

(1) The current estimate of the mean of the weights $\overline{W}$ for the latent attribute is used to generate 80 simulates from a normal distribution given by $N(\overline{W},1)$.

(2)There are in all 2^(J-1)+1 possible states, where J is the total number of brands available in the market. As a first step the probability of each state is evaluated. Using gauss estimates of $\gamma_o, \gamma_I, \gamma_T, \gamma_D$, equation 4.4 is used to calculate the probability of using a bit of information. The current estimate of this probability is then used to evaluate equation 4.2.2, which provides the probability of any given state.

(3) Given (i) the state in which the consumer is making decisions (that is the information being used by consumer) (ii) simulates from the normal distribution, and (iii) current estimates of the parameters, equations 5.1 and 5.2 are used to calculate the value of the utility function for choice of the brand and that of not choosing a brand in the market for each state is obtained.

(4) The utility value for each brand and that for no choice is used to estimate the probabilities given by equations 5.14 and 5.19, which are used to compute the likelihood function provided by equations 5.23 and 5.24. GAUSS optimization routines such as BFGS were used to obtain the maximum likelihood estimates.

It is important to note here that algorithm proposed in this dissertation is a hybrid of two different methods found in the literature: Elrod and Keane (1995) and Andrews and Srinivasan (1995). More specifically, it extends the Elrod and Keane (1995) algorithm to incorporate the effects of informational uncertainty using the method proposed by Andrews and Srinivasan (1995). A major limitation of this algorithm is that the run time increases exponentially with the number of alternatives and attributes of the alternatives. This makes estimation of large choice sets often an infeasible exercise. For example, even for as few as four alternatives and two attributes, estimation times on a Risc 6000 workstation were often close to 50 hours. Although the availability of fast computers will allow us to estimate these complex models, a major focus of future research should be to develop methods that considerably shorten the run time. Taylor series approximations are a promising avenue for future research.