NORTHWESTERN UNIVERSITY

Empirical Models of Consumer Behavior in Retailing

A DISSERTATION

SUBMITTED TO THE GRADUATE SCHOOL
IN PARTIAL FULFILLMENT OF THE REQUIREMENTS

for the degree

DOCTOR OF PHILOSOPHY

Field of Marketing

By

Lei Karen Wang

EVANSTON, ILLINOIS

December 2008

ABSTRACT

Empirical Models of Consumer Behavior in Retailing

Lei Karen Wang

The main objective of this research is to enhance our understanding of consumer behavior in retailing. This objective is accomplished through the analysis of retailers' customer database. This research provides methodologies for retailers to process the large amount of readily available customer data and make more effective marketing decisions.

This dissertation consists of two essays. The first essay provides an empirical analysis of consumers' learning process about a multi-product brand and its implication on managing brand equity. We propose a structural model to describe how consumers learn and form brand equity based on information from product usage experiences and mailing catalogs across multiple product categories of a retailer. The model is applied to a direct mail retailer that sells products in five categories. The results show significant learning within and across categories and also considerable heterogeneity across consumers in their learning processes. The model provides us a tool to track the evolution of brand equity and to identify the key product categories that have the most significant impact on this brand equity formation process for each individual consumer.

The second essay provides an empirical analysis of consumers' product return behavior. To control product returns is as important as to increase sales for retailers to improve their profitability. In this paper we investigate how price influences product returns. We theoretically and empirically test a widely accepted assumption in the operation literature that return rate is

constant. We identify two effects that may influence return rate when an item is discounted: the *perceived value effect* and the *incremental customer effect*. Empirically, we measure these two effects on two different datasets. We find that both effects have substantial impact on return rates and the effect size and direction vary by product categories.

ACKNOWLEDGEMENTS

This dissertation would never have been possible without the continuous encouragement, guidance and help from my advisors: Eric Anderson and Karsten Hansen. They are the best advisors that I could ever ask for. My gratitude to them is beyond words. I am also grateful to the others who served on my committee, Professors Vincent Nijs and Igal Hendel, for their time and valuable feedback.

I would like to thank my parents Min Zhu and Chengan Wang. Although they are thousands miles away, their love, trust and support is always by my side. Last but not the least, I would like to thank my wonderful husband Wangyang Ni, for being a great role model in work and life, for being the first listener to all my research ideas, for lots of fun in my life.

Table of Contents

Lists of Tables

List of Figures

Chapter 1

Introduction

Retail is the second-largest industry in the U.S. by number of businesses and number of employees. Retail sales in the U.S. were up about 3.8% to $4.49 trillion in 2007 (Plunkett Research). The growth means not only opportunities but also challenges for retailers. In this increasingly competitive environment, retailers are pressed to provide more high quality products and services to meet consumers' constantly changing and heterogeneous preferences. A thorough understanding of what consumers like is critical for retailers to design effective marketing strategies. The good news is that retailers have the advantage of interacting with their consumers directly and they can record every interaction with their customers using the latest database technology. However, how to turn the readily available data into managerial insights is one of the key questions to be answered.

This dissertation consists of two essays analyzing consumer behavior in retailing. The first essay provides an empirical analysis of consumers' learning process about a multi-product brand and its implication on managing brand equity. Retailers' brand equity is a critical differentiator in today's competitive marketplace. However, retail managers have few metrics that they can rely on to manage their brand equity. We propose a structural model to describe how consumers learn and form brand equity based on information from product usage experiences and mailing catalogs across multiple product categories of a retailer. Based on the conceptual framework of customer-based brand equity, we represent a consumer's overall preferences for a brand and its specific product categories as nodes that are linked in an associative network in consumer's memory. This associative network enables consumers to generalize what they learn from one product category to the other categories as well as to the brand. We integrate this associative network structure into a Bayesian learning model, which can be estimated empirically from consumers' purchase behavior over time across categories.

The model is applied to a direct mail retailer that sells products in five categories. We analyze the complete transaction history of a sample of customers over eight years. The results show the existence of learning within and across product categories as well as considerable heterogeneity in consumers' learning process. The model provides a tool to track the evolution of brand equity at the individual consumer level. It also identifies the key product category by measures the impact of each category in this brand equity formation process.

The second essay provides an empirical analysis of consumers' return behavior. Product returns has become a serious problem for many catalog and internet retailers. A high return rate not only means less net sales but also higher inventory cost and higher cost in managing the return flow. Therefore, to predict returns is as important as to predict demand. However, there are few empirical papers on estimating customer returns. In this paper we investigate how price influences product returns. We theoretically and empirically test a widely accepted assumption in the operation literature that a constant fraction of items purchased by consumers are eventually returned. This suggests that price has impact on the number of returns only through the number of items sold, not through the return rate. In this paper, we identify two effects that may influence return rate when an item is discounted. First, when customers pay a lower price they receive more surplus and are less likely to return the item. We label this effect as the *perceived value effect*. Second, customers who buy at discounted prices may have different return propensities from those who buy at regular prices. We label this effect as the *incremental* *customer effect*. Empirically, we test and measure these two effects on two different datasets. The framework and analysis illustrate the importance and value of integrating operations and marketing decisions.

In both essays, we analyze consumer behavior at individual level across multiple product categories. Such customer and product specific analyses provide retailers valuable insight into customizing marketing strategies across their customers and products.

Chapter 2

Consumer Learning and Brand Equity Formation

2.1 Introduction

In today's competitive marketplace, brands become one of the most important assets of a firm. A strong brand name can generate not only superior profitability today but also sustainable growth in the future. The incremental value due to the brand name is referred to as brand equity. To build this intangible asset, firms are investing heavily in products, services, or marketing communications. They also realize the importance of measuring and monitoring brand equity to evaluate the effectiveness of their brand building activities.

Various brand equity metrics have been developed by academic researchers and industry practitioners to measure the value of a brand. These metrics measure brand equity from different perspectives and serve different purposes. One type of metrics measures the market value of a brand from firm's perspective. For example, according to Interbrand's 2007 brand value report, apparel retailer Gap's brand value is worth 6.4 billion in 2006 and 5.5 billion in 2007. Knowing the market value of a brand is useful for situations, such as merge, acquisition and brand licensing. However, if brand managers of Gap want to know why the brand equity decreases and how to improve it, this aggregate level brand value estimate is not very informative. The other type of brand equity measures brand equity from consumers' perspective. Based on the premise that the power of a brand lies in consumers' minds, the consumer-perspective metrics measure consumers' brand knowledge by asking consumers directly or inferring it from consumers' purchase patterns (Green and Srinivasan 1978, Green and Srinivasan 1990, Park and Srinivasan

1994, Kamakura and Russell 1993, Rangaswamy et. al 1993). This type of metrics provides more insight into the source of brand equity and how to manage it.

In this paper, we take the consumers' perspective and look at how brand equity forms in consumer minds. Brand equity, which is based on brand knowledge, is dynamic. It may change as consumers learn from their product experiences, advertising, and other interactions with the brand. Since consumers' learning may have impact on their subsequent purchases, it is critical for firms to understand this process in order to manage it and influence it strategically. Furthermore, the brand equity may evolve differently for different consumers because of heterogeneous consumer characteristics or experiences. We incorporate both dynamics and heterogeneity in our model.

We study the dynamics of brand equity from consumer learning's perspective. Consider a new customer acquired by a multi-product brand. Initially, she is uncertain about her valuation of the products and the overall brand. Over time, she learns from various sources and forms preference for the specific products and the overall brand. One important feature of the learning captured in our model is information spillover. When the consumer learns about one product, she may generalize to the other products with the same brand name or to the overall brand. For example, a consumer who likes Eddie Bauer's apparel for the high quality fabric may also like to try Eddie Bauer's bedding products because these products share similar fabric. Another consumer may like Eddie Bauer's apparel so much that he thinks highly of the brand Eddie Bauer and is willing to purchase the other Eddie Bauer's products (such as bedding, furniture, or outdoor gears), no matter how similar these products are to the apparel products. Such information spillover can be explained by an associative network model from the behavioral theory (Anderson 1983, Keller 1993 and Kardes et. al. 2004), which states that consumers' memory can be represented as an associative network with connected nodes. In our model, we assume that a consumer's memory about a multi-product brand is an associative network that consists of a brand node and product nodes. The valuations stored in these nodes are updated based on new information about the products from usage experiences and mailing catalogs.

We apply the model to a direct mail retailer that sells products in five categories under their store brand. In this application, consumers are assumed to form valuations about the average quality of the product categories and about the brand based on their direct product experiences and mailing catalogs. We analyze the complete transaction history of a sample of customers over eight years. From the observed purchase patterns within and across categories, we infer each consumer's underlying learning process. The results show significant learning as is reflected in the changes in the purchase patterns across categories. Between the two types of information, the cumulative impact of mail catalogs is found to be bigger than that of direct product experiences. Since this is a catalog retailer selling durable goods, it is reasonable that the high frequency of catalogs received outweighs the low frequency of direct experiences to influence brand equity. Furthermore, the catalogs' impact on brand equity is found to decay much slower than the direct experiences. In other words, consumers learn faster from direct experiences than mail catalogs.

Our model offers two important managerial contributions. First, it allows us to track the brand equity over time at individual consumer level. Following the revenue premium definition of brand equity (Ailawadi, et. al. 2003), we compute the brand equity at each time period as the difference between the simulated revenues with and without the brand for each consumer at each time period. Our brand equity metric is superior to the firm-based brand equity metrics in terms of providing more diagnostic insight at individual consumer level. It is superior to the other customer-based metrics by accounting for dynamics using the readily available transaction data. Second, our model uncovers the key categories that drive the changes in brand equity. Analytically, we show that the impact of each category on the overall brand equity depends not only on how much each consumer's average actual experiences exceeds her prior expectation (disconfirmation) in that category, but also the weight associated with that category (disconfirmation weight). Categories that are perceived to be more related to the brand have larger weights. These properties derived from our learning model add to the consumer satisfaction/dissatisfaction literature (Boulding, et. al. 1993).

This paper also offers methodological contribution. We develop a structural model to describe how consumers form brand equity based on direct experiences and mailing catalogs. A behavioral associative network model is integrated with an empirical Bayesian learning model to describe the learning process. In our model, consumers learn from direct product experience and product information in catalogs in the same way as in a regular Bayesian learning model. In addition, we also allow consumers' perceived category-brand relationships to be influenced by the category frequencies in the retailer's catalogs. Since mailing catalog is a major way to advertise for a catalog retailer, it is reasonable to assume that how often a consumer receives information about a specific category in the catalogs influences his/her category-brand association. For example, if a consumer receives catalogs with women's clothing much more often than men's, she probably associates the brand with women's clothing more than men's.

The remainder of the paper is organized as follows. In section 2.2, we review the relevant literatures. In section 2.3, we develop a model of consumer learning. In section 2.4, we describe the data. We then discuss the estimation and identification issues in section 2.5 and 2.6. In section 2.7, we present model results. In section 2.8, we discuss the managerial implications. The paper concludes with a brief discussion.

2.2 Related Research

2.2.1 Conceptual framework of brand equity

Our conceptual framework is motivated by the research on customer-based brand equity. Brand equity is the value of a brand (Farquhar 1989). It has been defined from the perspective of a firm and the perspective of a consumer. From the firm's perspective, brand equity is the additional value that accrues to a firm with a brand name compared to a firm without a brand name. It is measured as the aggregate level market outcome (such as revenue, profit or price premium) that is due to a brand name (Ailawadi, et. al. 2003). This measure aggregates over all products and customers and doesn't provide much insight into the source of brand equity. In this paper, we adopt a more relevant perspective for managing brand equity: the customer-based perspective.

The customer-based perspective defines brand equity as the differential effect of the brand knowledge on consumer response to the marketing of that brand (Keller 1993). The basic premise of the customer-based brand equity is that the power of a brand lies in what customers have learnt about the brand as a result of their interactions with the brand over time. Notice that this definition consists of three elements. The first is the "differential effect". Brand equity is a relative concept with respect to a benchmark (a generic product or competing brands). Secondly, this differential effect arises because of consumers' "brand knowledge". The brand equity may change as brand knowledge accumulates in consumers' minds. The third element in the definition is "consumer response to marketing". The differential effect of brand knowledge is reflected in consumers' response to marketing, including their choice of a brand, response to sales promotions and advertising, or evaluation of a product. The metric that we develop reflects these three elements.

Various metrics have been proposed to measure customer-based brand equity. One way is to measure multiple dimensions of brand knowledge, i.e. brand awareness and brand associations, using surveys or lab experiments (Park and Srinivasan 1994, Srinivasan et. al. 2005). The advantage of this method is that researchers can learn what consumers know and how they feel about a brand in an accurate and detailed manner. The disadvantage is that it relies on consumers' self report, which can be potentially biased and may not be consistent with consumers' actual responses to marketing. Moreover, to collect such data for a large sample on a continuous base is costly. Another way to measure brand equity is to infer brand knowledge from what consumers do. Kamakura and Russell (1993) use scanner data to estimate brand equity, which is operationalized as the residual utility after accounting for the utility of the physical products. This method utilizes the readily available transaction data and relies on statistical model to make inferences about brand equity. It is less costly for long-run and large sample tracking purposes. For our research purposes, we adopt the choice modeling framework to estimate brand equity from consumers' purchase patterns.

In contrast to previous works which measure brand equity at a certain point of time, our study models the dynamics of brand equity. The model specification is motivated by the conceptualization of customer-based brand equity and related theories on brand equity formation. Consistent with an associative network memory model (Anderson 1983), Keller (1993) conceptualizes brand knowledge (the basis of brand equity) as a network of associations. These associations differ in their levels of abstraction. While associations about attributes and benefits are specific to product categories, attitudes toward the brand are more general and can be applied to all products that share the same brand name. Ultimately, both the beliefs about the objective reality of the products and the belief about the overall brand may be reflected in brand choice and brand loyalty (Park 1991). In this paper, we decompose consumer's overall utility of consuming a product into two components: brand component and product component. Initially, consumers have uncertainty about these components and hold beliefs about them. According to the associative network model, the beliefs about the brand component and the product components are stored as connected nodes in consumer's memory. The network has the property that activation of one node may activate the other connected nodes depending on the strength of the link between the nodes.

As summarized by Keller (1993), there are three ways brand beliefs are created. The first way is by experiencing the products. The second way is by learning from information about the products communicated by the company, other commercial sources or word of mouth. The third way is on the basis of inferences from other brand associations, such as beliefs about other related products. In our specification, consumers not only learn from their direct product experiences and the brand's mailing catalogs, but also make inferences across products. Over time, the brand belief arises as consumers learn about the brand's products from various sources.

In summary, our paper contributes to the customer-based brand equity research in two ways. First, it quantifies the dynamic process of brand equity evolution as a result of consumer learning. Secondly, this paper proposes a methodology to estimate the associative network for a multi-product brand from consumer's actual purchase behavior. A better understanding of the brand equity formation process and consumers' perceived brand structure helps managers better manage products in order to maximize brand equity.

2.2.2 Consumer learning

As consumers collect information from various sources, they update their beliefs about the products. The empirical literature on Bayesian learning provides a tool to estimate this learning process from consumer purchase patterns over time. In these papers, consumers learn about the uncertain quality of some experience goods in a Bayesian manner by combining their prior beliefs and the newly received information (Erdem and Keane 1996; Crawford and Shum 2000; Ching 2002; Ackerberg 2003; Narayanan and Manchanda 2006). The single-product learning model has been extended to multi-product learning (Erdem 1998, Ackerberg 2003, Coscelli and Shum 2004). In these models, information about one product may be informative about the quality beliefs of the other products. The information spillover can be captured by covariances in the prior beliefs (Erdem 1998, Ackerberg 2003) or in the information signals (Coscelli and Shum 2004). A property of the learning process specified in these Bayesian learning papers is that the covariances are monotonically non-increasing. In other words, the perceived covariance between products never increases as consumers learn. This may be too restrictive in reality. For example, as a consumer receives more catalogs with men's apparel from a brand, he may associate the brand more with men's apparel. In our model, we allow the perceived brand-product relationship to be influenced by the brand's marketing communications and on top of this we model the consumer learning process.

In our study, we consider a multi-product brand with an umbrella structure, i.e. all products share the same brand name. This setup is similar to Erdem (1998). In her model, information about one product can spill over to the other product due to consumers' perceived correlation between the two products with the same brand name. However, the mechanism underlying the correlation is not explicitly modeled. The correlation can be driven by the common brand associations at the abstract level or the similarities at the concrete product level. Without separating the brand preference from the product preference, we don't know what the value of a brand is and which product is more important in influencing the value of the brand. In this work, we model consumers learning about the brand and the products separately based on the associative network model from the behavioral literature. The model allows us to estimate the retail brand equity accrued during the learning process and to identify the key product that drives the evolution of the overall brand equity. In addition, our model adds to Erdem's model in two ways. First, we account for learning from both usage experiences and marketing communications, whereas Erdem's model only accounts for the former. Furthermore, as mentioned earlier, we allow the perceived brand-product relationship to be influenced by marketing communications. Second, our model allows consumers to differ in more aspects of the learning process, including priors, prior covariances, signal variances, and true valuations, whereas Erdem's model only accounts for heterogeneity in the true valuations. Better controlling for heterogeneity ensures us to get better estimates of the learning process.

2.2.3 Cross-selling and category management

Consumers often purchase multiple products from the same brand. This provides cross-selling opportunities to the firm. A number of papers have studied consumers' cross-category purchase behavior. They focus on modeling consumers' purchase sequences across categories and predicting which product consumers might purchase next (Kamakura et. al. 1991, Knott et. al. 2002, Li et. al. 2005). When there is a natural sequence in which consumers purchase multiple products (e.g., a computer before a printer, a checking account before a brokerage account), the firm does not have much choice to influence the order. However, when there is no natural sequence of purchase, firms may have strategic reasons to influence consumers' purchase sequences. In our study, we provide such a motivation: to build overall brand equity. Since the order in which consumers make purchases may influence the brand equity formed in their minds and alter their subsequent purchases, firms need to design optimal product sequence to promote to consumers. This requires mangers to understand the role of each product (category) in influencing the overall brand equity.

2.3 Model

Consider a multi-product brand selling experience goods. All products, denoted as $k=1, \ldots, K$, share the same brand name. This brand structure is prevalent in reality and is often called a branded house or an umbrella brand. For example, Eddie Bauer, J. Crew, Apple, and

Colgate are all multi-product brands with such a brand structure. A customer who is newly acquired by the brand is uncertain about her valuation of the products. She holds beliefs about them. During the learning process, the beliefs are revised based on information from her direct product experiences and the brand's mailing catalogs. In what follows, we first discuss the consumer's purchase decision and then the learning process.

2.3.1 Purchase decisions

In our model, consumer *i* (i=1, …, I) makes a purchase decision for product *k* (k=1, …, K) at time *t* (t=1, …, T). Since consumer *i* does not know the utility of the product perfectly, she decides whether or not to buy the product based on her expected utility at time *t*. To simplify the model, let's assume that the consumer is risk neutral and myopic. At time *t*, her expected utility of product *k* is specified as:

$$U_{kit} = \bar{Q}_{kit-1} + \beta_{ki} X_{kit} + \varepsilon_{kit} \tag{2.1}$$

where the intercept $\bar{Q}_{kit-1}$ is consumer *i*'s expected utility of product *k* given information up to time *t-1*. X_{kit} is the covariates and β_{ki} is the corresponding heterogeneous coefficients. The error terms ε_{kit} denotes the idiosyncratic utility shock, which is independently distributed as $N(0,1)$. These utility shocks are observable to the consumer, but not to us researchers.

Suppose the utility of consuming the outside option is normalized to 0. The consumer decides to purchase, if and only if U_{kit} is positive. Let $Y_{it} = \{Y_{1it}, Y_{2it}, ..., Y_{Kit}\}$ denote the purchase decisions of all products, where $Y_{kit} = 1$ if product *k* is purchased at time *t*, and $Y_{kit} = 0$ otherwise.

2.3.2 Consumer learning

In this section, we first describe how consumer beliefs are specified. We then describe how these beliefs are updated upon newly arrived information.

Suppose the overall utility of a product is decomposed into two components: the utility derived from the brand (b_i) and the utility derived from the specific product (c_{ki}). Since these products are experience goods, consumer i does not know her b_i and c_{ki} perfectly prior to purchases. She holds some beliefs about them, which can be characterized by favorability, strength, perceived relationships among products and brand. Favorability captures how the consumer likes the brand or the products in general, strength captures how certain the consumer is about her beliefs, and perceived relationships captures how closely related the products and the brand are perceived to be in the consumer's memory. Mathematically, the beliefs about b_i and $\{c_{ki}\}$ can be represented as $(c_{1i}, c_{2i}, ..., b_i)'_t$, which is assumed to follow a multivariate normal distribution with mean and variance-covariance matrix specified as in (2.2):

$$\begin{bmatrix} c_{1i} \\ c_{2i} \\ \cdots \\ c_{Ki} \\ b_i \end{bmatrix}_t \sim MVN(\bar{Q}_{it} = \begin{bmatrix} \bar{C}_{1it} \\ \bar{C}_{2it} \\ \cdots \\ \bar{C}_{Kit} \\ \bar{B}_{it} \end{bmatrix}, \Sigma_{it} = \begin{pmatrix} \delta^2_{1it} & r_{12it} & \cdots & r_{1Kit} & \varpi_{1it} \\ r_{12it} & \delta^2_{2it} & \cdots & r_{2Kit} & \varpi_{2it} \\ \cdots & \cdots & \cdots & \cdots & \cdots \\ r_{1Kit} & r_{2Kit} & \cdots & \delta^2_{Kit} & \varpi_{Kit} \\ \varpi_{1it} & \varpi_{2it} & \cdots & \varpi_{Kit} & \delta^2_{bit} \end{pmatrix}) \qquad (2.2)$$

where $\{\bar{C}_{kit}\}$ and $\bar{B}_{it}$ are the means, which represent the favorability of product k and the brand. With this specification, $\bar{Q}_{kit}$ in equation (2.1) can be written as $\bar{Q}_{kit} = \bar{B}_{it} + \bar{C}_{kit}$. δ^2_{bit} and $\{\delta^2_{kit}\}$ are the variances, which capture the strength of the beliefs; $\{r_{kjit}, j \neq k\}$ is the covariance between products, and { ϖ_{kit} } is the covariance between the products and the brand. The product-product covariance r_{kjit} arises because of the similarities in specific and concrete product attributes or usage contexts, whereas the product-brand covariance ϖ_{kit} arises because of more general and abstract brand associations. The covariance matrix is used to represent the consumer's associative network in the memory. We will show later how these covariance terms drive the information spillover within the brand structure.

The learning process is characterized by the evolution of consumer's beliefs over time. It is assumed that the consumer updates her product and brand beliefs based on two types of information: direct product usage experiences and mailing catalogs (advertising). Although information from usage experiences is generated by the consumer herself and information from mailing catalogs is generated by the company, both types are assumed to contain information about the physical products (c_{ki}). Note that we assume the information received is directly about the products and not about the brand. The brand belief is assumed to be derived from the product information. This assumption is reasonable in the mailing catalog industry where the brand essence is mostly conveyed through products. It is also consistent with the behavioral research on abstract and concrete attributes which shows that concrete attributes are directly associated with the objects, whereas abstract attributes are computed or inferred from more concrete attributes (Bettman and Sujan 1987).

In addition, we assume that the frequency of each product in the mailing catalogs influences the consumer's learning process by changing the perceived product and brand relationship. This assumption is based on the rationale that a mailing catalog is not only a carrier of product quality information but also a messenger to communicate the position of each product in the brand family. It is expected that the more frequent a consumer receives information about a specific product, the more she would relate the product to the brand.

The learning process consists of two steps during each round of updating. In the first step, the frequency of each product in the mailing catalogs influences the consumer's perceived product-brand relationship. Suppose during time *t* the consumer *i* receives information about product *k*, ns_{kit} times. This frequency changes the consumer's variance-covariance in her belief from Σ_{it} to $\tilde{\Sigma}_{it}$, where $\tilde{\Sigma}_{it}$ is:

$$\tilde{\Sigma}_{it} = \begin{pmatrix} \delta_{1it}^2 & r_{12it} & \dots & r_{1Kit} & \tilde{\varpi}_{1it} \\ r_{12it} & \delta_{2it}^2 & \dots & r_{2Kit} & \tilde{\varpi}_{2it} \\ \dots & \dots & \dots & \dots & \dots \\ r_{1Kit} & r_{2Kit} & \dots & \delta_{Kit}^2 & \tilde{\varpi}_{Kit} \\ \tilde{\varpi}_{1it} & \tilde{\varpi}_{2it} & \dots & \tilde{\varpi}_{2it} & \delta_{bit}^2 \end{pmatrix} \text{ and } \tilde{\varpi}_{kit} = \varpi_{kit} + \alpha_i \delta_{kit}^2 \cdot ns_{kit}$$

Notice that the effect of product frequency ns_{kit} on $\tilde{\varpi}_{kit}$ is assumed to depend on the consumer's uncertainty about product *k* (i.e. δ_{kit}^2). The more uncertain the customer is, the more she would relate the product with the brand when α_i is positive.

In the second step, the customer updates her beliefs about b_i and $\{ c_{ki} \}$ in a Bayesian manner based on the product information received from her direct product experiences and the mailing catalogs. We specify each piece of information received as a continuous random variable which is only observable to the consumer after she receives the product and is not observable to us researchers. Each realization of an experience or an observation in the mail catalog, called a signal, is the consumer's holistic valuation of the corresponding product. For example, after a purchase of product k is made by consumer i at time t, a realization of the experience signal is observed by the consumer. It is assumed that each experience signal, denoted as E_{kit}, is drawn from the same normal distribution:

$$E_{kit} \overset{i.i.d.}{\sim} N(c_{ki}, \pi_i^2)$$

where the mean c_{ki} is the true utility of product k and variance π_i^2 reflects the consumer's perceived variation in the signals. Note that c_{ki} is unknown to the consumer initially and can be learned after the consumer observes repeatedly the experience signal $\{E_{kit}\}$ for a sufficiently large number of times. One of the factors that determines the number of signals needed for the consumer to learn c_{ki} (speed of learning) is the signal variance π_i^2. If the consumer believes that the signals have a small variance, then each signal is very informative about c_{ki} and the consumer needs only a few signals to learn c_{ki}. In an extreme case, if $\pi_i^2 = 0$, the consumer learns about c_{ki} perfectly after receiving only one experience signal.

In addition to product experiences, the consumer can also learn about products from the catalog content. Similar to the experience signals, the catalog signals are informative about the product quality c_{ki}. Suppose the consumer receives product k's information in the catalogs ns_{kit} times during time t. The average of these ns_{kit} signals is denoted as S_{kit}, which is assumed to be normally distributed around c_{ki}:

$$S_{kit} \overset{i.i.d.}{\sim} N(c_{ki}, \sigma_i^2 / ns_{kit})$$

where σ_i^2 is the variance of the each catalog signal. Since previous studies show using lab experiments that direct experience is more informative about experience attributes than advertising (Wright and Lynch 1995), it is expected that σ_i^2 is larger than π_i^2. Therefore, our model estimates allow us to test empirically whether or not usage experience is more informative than advertising.

Based on both experience signals and catalog signals, the consumer updates her beliefs in a Bayesian manner. To express the beliefs in a matrix form, we need the following notations.

Let $E_{it} = (E_{1it}, E_{2it}, ..., E_{Kit})'$, $S_{it} = (S_{1it}, S_{2it}, ..., S_{Kit})'$, $\underset{K*K}{T_{it}^{E}} = \begin{pmatrix} Y_{1it} & 0 & ... & 0 \\ 0 & Y_{2it} & ... & 0 \\ 0 & 0 & ... & ... \\ 0 & 0 & ... & Y_{Kit} \end{pmatrix}$ and

$\underset{K*K}{T_{it}^{S}} = \begin{pmatrix} ns_{1it} & 0 & ... & 0 \\ 0 & ns_{2it} & ... & 0 \\ 0 & 0 & ... & ... \\ 0 & 0 & ... & ns_{Kit} \end{pmatrix}$. Recall that after the first step of updating at time *t*, the

consumer's beliefs follow a multivariate normal distribution with mean $\bar{Q}_{it}$ and variance-covariance $\tilde{\Sigma}_{it}$. This is the prior belief which is to be combined with the signals received during time *t* to form posterior belief ($\bar{Q}_{it+1}, \Sigma_{it+1}$). Given the normality assumption of the prior beliefs and the signals, the posterior beliefs follow a multivariate normal distribution (DeGrout 1970):

$$\begin{bmatrix} c_{1i} \\ c_{2i} \\ \cdots \\ c_{Ki} \\ b_i \end{bmatrix}_{t+1} \sim MVN\left(\bar{Q}_{it+1} = \begin{bmatrix} \bar{C}_{1it} \\ \bar{C}_{2it} \\ \cdots \\ \bar{C}_{Kit} \\ \bar{B}_{it} \end{bmatrix}, \Sigma_{it+1} \right)$$

where the posterior mean and the posterior variance are:

$$\Sigma_{it+1} = (\tilde{\Sigma}_{it}^{-1} + \begin{bmatrix} T_{it}^{E}\pi_i^{-2} + T_{it}^{S}\sigma_i^{-2} & \mathbf{0} \\ \mathbf{0}' & 0 \end{bmatrix})^{-1}$$

$$\bar{Q}_{it+1} = \Sigma_{it+1}(\tilde{\Sigma}_{it}^{-1}\bar{Q}_{it} + \begin{bmatrix} T_{it}^{E}\pi_i^{-2}E_{it} + T_{it}^{S}\sigma_i^{-2}S_{it} \\ 0 \end{bmatrix})$$

Here $\mathbf{0}$ is a K dimensional vector of 0's. Notice that posterior mean vector $\bar{Q}_{it+1}$ includes the posterior mean belief of each product (i.e. $\bar{C}_{1it}$, $\bar{C}_{2it}$, …, $\bar{C}_{Kit}$) and the brand (i.e. $\bar{B}_{it}$). All *K+1* elements in the belief vector is updated simultaneously based on information signals about K products.

2.3.3 Model comparison with a regular multivariate Bayesian learning model

The consumer's belief is characterized by mean vector $\bar{Q}_{it}$ and variance-covariance matrix Σ_{it} and they evolve over time as the consumer learns. It is important to summarize the key differences between our model and a regular multivariate Bayesian learning model. In a regular Bayesian learning model, the consumers update their beliefs about the product quality in a Bayesian manner (the second step in our model). In this process, the mean of the quality beliefs is assumed to evolve randomly depending on the level of the received signals, but the variance-covariance of the quality beliefs is assumed to evolve deterministically and is decreasing upon each update. In our model, we relax this non-increasing assumption about the covariance of the beliefs. We assume that before each Bayesian updating the consumer revises her perceived covariance between the products and the brand based on the frequency of the product shown up in the catalogs.

In what follows, we derive the recursive updating equations for the posterior mean and variance-covariance in our model and compare them to a regular Bayesian learning model.

First, let's look at the posterior variance-covariance. Recall that $\tilde{\Sigma}_{it} = \begin{pmatrix} \delta_{1it}^2 & r_{12it} & \ldots & r_{1Kit} & \tilde{\varpi}_{1it} \\ r_{12it} & \delta_{2it}^2 & \ldots & r_{2Kit} & \tilde{\varpi}_{2it} \\ \ldots & \ldots & \ldots & \ldots & \ldots \\ r_{1Kit} & r_{2Kit} & \ldots & \delta_{Kit}^2 & \tilde{\varpi}_{Kit} \\ \tilde{\varpi}_{1it} & \tilde{\varpi}_{2it} & \ldots & \tilde{\varpi}_{2it} & \delta_{bit}^2 \end{pmatrix}$. We can partition $\tilde{\Sigma}_{it}$ into four parts: $\tilde{\Sigma}_{it} = \begin{pmatrix} \Sigma_{it}^{K*K} & \tilde{W}_{it} \\ \tilde{W}_{it}' & \delta_{bit}^2 \end{pmatrix}$

where Σ_{it}^{K*K} is the upper K*K submatrix of Σ_{it} and it includes variance and covariance of the product beliefs. $\tilde{W}_{it}$ is the vector of the revised covariances between the product beliefs and the brand belief, i.e. $\tilde{W}_{it}' = (\tilde{\varpi}_{1it}, \ldots, \tilde{\varpi}_{Kit})'$. δ_{bit}^2 is the variance of the brand belief. Then we can show that Σ_{it+1} can be expressed as a function of the petitioned elements: Σ_{it}^{K*K}, $\tilde{W}_{it}$, and δ_{bit}^2 (see Appendix 2.1 for details):

$$\Sigma_{it+1} = (\tilde{\Sigma}_{it}^{-1} + \begin{bmatrix} T_{it}^{E}\pi_i^{-2} + T_{it}^{S}\sigma_i^{-2} & \mathbf{0} \\ \mathbf{0}' & 0 \end{bmatrix})^{-1}$$

$$= (I_{(K+1)*(K+1)} + \begin{pmatrix} \Sigma_{it}^{K*K} & \tilde{W}_{it} \\ \tilde{W}_{it}' & \delta_{bit}^2 \end{pmatrix} \begin{bmatrix} T_{it}^{E}\pi_i^{-2} + T_{it}^{S}\sigma_i^{-2} & \mathbf{0} \\ \mathbf{0}' & 0 \end{bmatrix})^{-1} \begin{pmatrix} \Sigma_{it}^{K*K} & \tilde{W}_{it} \\ \tilde{W}_{it}' & \delta_{bit}^2 \end{pmatrix} \quad (2.3)$$

$$= \begin{pmatrix} \Sigma_{it+1}^{K*K} & W_{it+1} \\ W_{it+1}' & \delta_{bit+1}^2 \end{pmatrix}$$

where

$$\Sigma_{it+1}^{K*K} = (I_{K*K} + \Sigma_{it}^{K*K}(T_{it}^{E}\pi_i^{-2} + T_{it}^{S}\sigma_i^{-2}))^{-1}\Sigma_{it}^{K*K} = ((\Sigma_{it}^{K*K})^{-1} + T_{it}^{E}\pi_i^{-2} + T_{it}^{S}\sigma_i^{-2})^{-1}$$

$$W_{it+1} = (\Sigma_{it}^{K*K}(T_{it}^{E}\pi_i^{-2} + T_{it}^{S}\sigma_i^{-2}) + I_{K*K})^{-1}\tilde{W}_{it}$$

$$\delta_{bit+1}^2 = -\tilde{W}_{it}'(T_{it}^{E}\pi_i^{-2} + T_{it}^{S}\sigma_i^{-2})(\Sigma_{it}^{K*K}(T_{it}^{E}\pi_i^{-2} + T_{it}^{S}\sigma_i^{-2}) + I_{K*K})^{-1}\tilde{W}_{it} + \delta_{bit}^2$$

Notice that the updating of the variance-covariance of the product beliefs Σ_{it+1}^{K*K} doesn't depend on the product-brand covariance W_{it}. Σ_{it+1}^{K*K} evolves in exact the same way as in a regular Bayesian updating process. On the other hand, W_{it+1} and δ_{bit+1}^2 depend on the revised covariance between the product beliefs and the brand belief (i.e. $\tilde{W}_{it}$). Therefore, their evolutions differ from a regular Bayesian learning model.

Then, let's look at the posterior mean. We can partition $\bar{Q}_{it+1}$ into two parts $\bar{C}_{it+1}$ and $\bar{B}_{it+1}$, where $\bar{C}_{it} = (\bar{C}_{1it}, ..., \bar{C}_{Kit})'$. We can rewrite $\bar{Q}_{it+1}$ as:

$$\bar{Q}_{it+1} = \begin{pmatrix} \bar{C}_{it+1} \\ \bar{B}_{it+1} \end{pmatrix} = (I_{(K+1)*(K+1)} + \tilde{\Sigma}_{it} \begin{bmatrix} T_{it}^{E}\pi_i^{-2} + T_{it}^{S}\sigma_i^{-2} & \mathbf{0} \\ \mathbf{0}' & 0 \end{bmatrix})^{-1} \begin{pmatrix} \bar{C}_{it} \\ \bar{B}_{it} \end{pmatrix}$$

$$+ (I_{(K+1)*(K+1)} + \tilde{\Sigma}_{it} \begin{bmatrix} T_{it}^{E}\pi_i^{-2} + T_{it}^{S}\sigma_i^{-2} & \mathbf{0} \\ \mathbf{0}' & 0 \end{bmatrix})^{-1} \tilde{\Sigma}_{it} \begin{bmatrix} T_{it}^{E}\pi_i^{-2}E_{it} + T_{it}^{S}\sigma_i^{-2}S_{it} \\ 0 \end{bmatrix}$$

Based on the partition results of Σ_{it} and $\tilde{\Sigma}_{it}$ in equation (2.3), we can expand the matrix multiplication to get the following expressions (see Appendix 2.2 for details):

$$\bar{C}_{it+1} = \Sigma_{it+1}^{K*K}((\Sigma_{it}^{K*K})^{-1}\bar{C}_{it} + T_{it}^{E}\pi_i^{-2}E_{it} + T_{it}^{S}\sigma_i^{-2}S_{it})$$

$$\begin{aligned}\bar{B}_{it+1} &= \bar{B}_{it} - \tilde{W}_{it}'(T_{it}^{E}\pi_i^{-2} + T_{it}^{S}\sigma_i^{-2})\Sigma_{it+1}^{K*K}(\Sigma_{it}^{K*K})^{-1}\bar{C}_{it} \\ &\quad + \tilde{W}_{it}'(I_{K*K} - (T_{it}^{E}\pi_i^{-2} + T_{it}^{S}\sigma_i^{-2})\Sigma_{it+1}^{K*K})(T_{it}^{E}\pi_i^{-2}E_{it} + T_{it}^{S}\sigma_i^{-2}S_{it})\end{aligned} \tag{2.4}$$

Notice that $\bar{C}_{it+1}$ depends on not only $\bar{C}_{it}$ but also signals E_{it} and S_{it}. This recursive expression for $\bar{C}_{it+1}$ is no different from a regular Bayesian learning model. The revised covariances $\tilde{W}_{it}'$ doesn't influence the updating of $\bar{C}_{it+1}$ and only influences the updating of $\bar{B}_{it+1}$. Therefore, the posterior mean brand belief $\bar{B}_{it+1}$ evolves differently from a regular Bayesian model.

2.3.4 Decomposing product's contribution to the evolution of brand valuation

Since the main interest of the paper is to understand how $\bar{B}_{it}$ evolves over time, we focus on the evolution process of $\bar{B}_{it}$. Recall that the evolution of $\bar{B}_{it}$ is driven by product information from direct product experiences and mailing catalogs, i.e. E_{it} and S_{it}. To see more explicitly how $\bar{B}_{it}$ changes with these information signals, we can rearrange the terms in equation (2.4) and get the following expression:

$$\begin{aligned}\bar{B}_{it+1} &= \bar{B}_{it} + \tilde{W}_{it}'(I_{K*K} - (T_{it}^{E}\pi_i^{-2} + T_{it}^{S}\sigma_i^{-2})\Sigma_{it+1}^{K*K})[T_{it}^{E}\pi_i^{-2}(E_{it} - \bar{C}_{it}) + T_{it}^{S}\sigma_i^{-2}(S_{it} - \bar{C}_{it})] \\ &= \bar{B}_{it} + \sum_{k=1}^{K}[\tilde{\varpi}_{kit}(1-\delta_{kit+1}^{2}(Y_{kit}\pi_i^{-2} + ns_{kit}\sigma_i^{-2})) - \sum_{j\neq k}^{K}\tilde{\varpi}_{jit}r_{jkit+1}(Y_{jit}\pi_i^{-2} + ns_{jit}\sigma_i^{-2})] \\ &\quad *[Y_{kit}\pi_i^{-2}(E_{kit} - \bar{C}_{kit}) + ns_{kit}\sigma_i^{-2}(S_{kit} - \bar{C}_{kit})]\end{aligned} \tag{2.5}$$

The change in $\bar{B}_{it}$ upon each updating can be decomposed by products and information type. For each type of information, each product's contribution to $\bar{B}_{it}$ is a multiplication of two factors. The first factor is $E_{kit} - \bar{C}_{kit}$ or $S_{kit} - \bar{C}_{kit}$, which is the difference between the actual product signals received and the prior expectation before receiving the signals. Following the consumer satisfaction literature (Woodruff et. al. 1983), we call this difference *disconfirmation*. A positive disconfirmation means the actual information received is better than expected. The second fact is the product specific *weight* of the disconfirmation, i.e. $[\tilde{\varpi}_{kit}(1-\delta^2_{kit+1}(Y_{kit}\pi_i^{-2} + ns_{kit}\sigma_i^{-2})) - \sum_{j\neq k}^{K} \tilde{\varpi}_{jit} r_{jkit+1}(Y_{jit}\pi_i^{-2} + ns_{jit}\sigma_i^{-2})]Y_{kit}\pi_i^{-2}$ for direct experience and $[\tilde{\varpi}_{kit}(1-\delta^2_{kit+1}(Y_{kit}\pi_i^{-2} + ns_{kit}\sigma_i^{-2})) - \sum_{j\neq k}^{K} \tilde{\varpi}_{jit} r_{jkit+1}(Y_{jit}\pi_i^{-2} + ns_{jit}\sigma_i^{-2})]ns_{kit}\sigma_i^{-2}$ for catalog information. We can compare the disconfirmation weights across products to identify the product that has the largest disconfirmation weight in influencing $\bar{B}_{it}$. We define such a product as a *key product*. Among all products, the product with smaller δ^2_{kit+1} (the consumer has less uncertainty), higher $\tilde{\varpi}_{kit}$, lower $\tilde{\varpi}_{jit} r_{kjit+1}, j \neq k$ (perceived to be more related to the brand than the other products), and higher Y_{kit} or ns_{kit} (the more frequently purchased or advertised) gets more weight. We can also compare the disconfirmation weights across information types. Between two types of information, the type of with larger $Y_{kit}\pi_i^{-2}$ and $ns_{kit}\sigma_i^{-2}$ (more frequently received and more accurate) gets more weight.

In summary, the learning process is characterized by the evolution of $\bar{Q}_{it}$ and Σ_{it}. The evolution of $\bar{B}_{it}$ is found to depend not only on each product's *disconfirmation* but also the product's *disconfirmation weight* for each information type. In the recursive updating process, the parameters that characterize the learning process are $(\bar{C}_{ki0}, \bar{B}_{i0})$ in the initial mean beliefs, $(\varpi_{ki0}, \delta_{ki0}^2, r_{kji0})$ in the initial variance-covariance matrix, α_i in $\tilde{\Sigma}_{it}$, signal variances (σ_i^2, π_i^2) and signals E_{it} and S_{it}. Not all of these parameters are identifiable. In the empirical application, we will discuss the necessary normalizations for the identification.

2.4 Data

In the empirical application, we study a retail brand which offers products in multiple categories under a common store brand name. We first describe the data and then motivate our learning model by providing a preliminary analysis of consumers' purchase patterns within the brand.

This retailer offers thousands of products under their store brand and these are durable and experience goods. To keep our model parsimonious and tractable, we merge all products into five categories, following the categorization by the retailer. As summarized in Table 2.1, three of the five categories are apparel. The other two categories are non-apparel categories and for confidentiality reasons they are denoted as Non-apparel 1 and Non-Apparel 2. The summary statistics of prices in these five categories are shown in Table 2.2.

Table 2.1: Descriptions of Five Product Categories

Category Name	Items
Tops and Bottoms	Tops, Bottoms, Dresses, Sleepwear, Underwear
Footwear	Footwear
Outerwear	Outerwear Tops, Outerwear Bottoms, Headwear, Gloves
Non-Apparel 1	NA
Non-Apparel 2	NA

Table 2.2: Mean and Standard Deviation for Price in Five Categories

Category Name	Average Full Price (Standard Deviation)	Average Sell Price (Standard Deviation)
Tops and Bottoms	32.36	30.02
	(14.67)	(14.17)
Footwear	45.62	43.12
	(25.16)	(25.95)
Outerwear	63.94	59.18
	(40.12)	(38.89)
Non-Apparel 1	46.37	44.84
	(53.21)	(48.38)
Non-Apparel 2	35.56	34.92
	(78.49)	(20.91)

The retailer's customer base consists of thousands of customers. For our research purposes, we take a random sample of 500 customers who meet the following two criteria. First, we include customers who were acquired in July 1999. The acquisition occurs when the company first obtains the customer information due to first purchase, gift card registration, and catalog request, etc. Secondly, we only include customers who don't have access to the retailer's local stores. These customers make purchases remotely (by telephone, mail or internet) without physically inspecting the products. Since customers who have access to local stores may engage in a different learning process, we save them for future research.

We observe these customers' transactions with the retailer in all five categories between July 1999 and June 2007. Let's look at some summary statistics for their purchase patterns.

First, Table 2.3 provides an overview of the number of customers in each categories and the number of items purchased by these customers. In terms of the number of customers, Non-Apparel 2 is the largest category with 370 customers, followed by Tops & Bottoms with 223 customers. However, among the customers who have made purchases in the categories, Tops & Bottoms is the category that the customers buy mostly frequently. An average customer who has made purchases in Tops & Bottoms has bought 9.73 items over the eight years. Therefore, Tops & Bottoms is the largest category in terms of total items sold, followed by Non-Apparel 2.

Table 2.3: Summary Statistics for the Number of Items Purchased among Customers Who Have Made Purchases in These Category

	Number of Customers	Number of Items Purchased by Each Customer				Total Items Purchased
		Mean	Std. Dev.	Min	Max	
Tops & Bottoms	223	9.73	16.64	1	119	2170
Footwear	153	2.97	3.02	1	19	454
Outerwear	175	2.94	3.08	1	23	515
Non-Apparel 1	77	3.71	3.85	1	19	286
Non-Apparel 2	370	3.99	5.69	1	81	1476

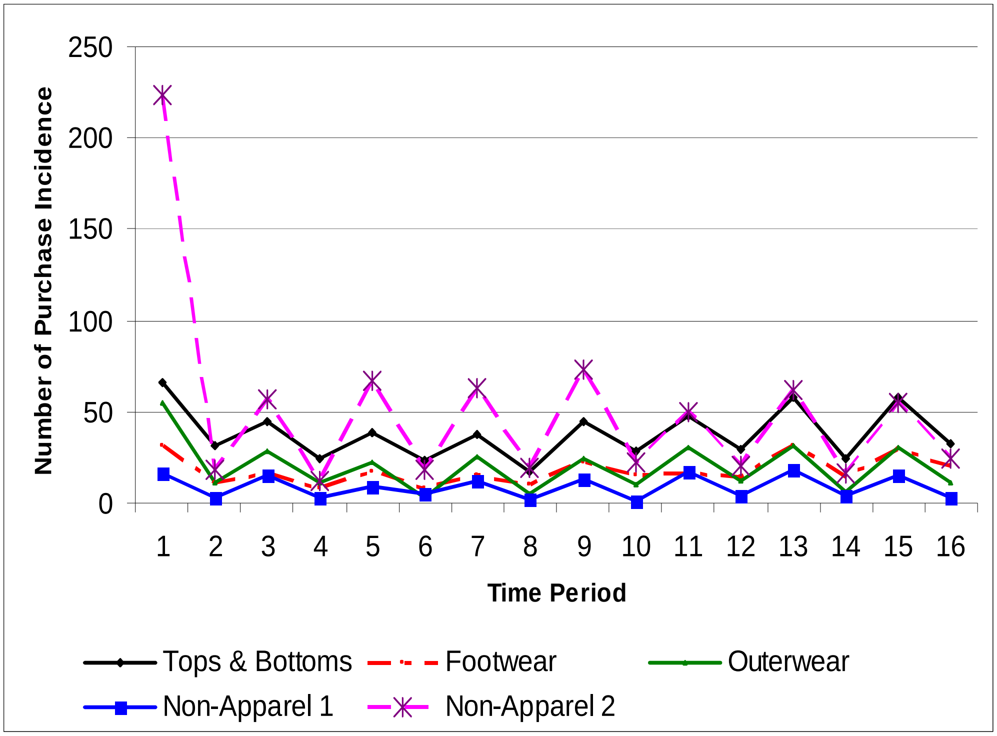
Second, how many consumers are multi-category buyers? Over the eight years, 48.6% customers have purchased only 1 category from the retailer, 23.2% have purchased 2 categories, 13.4% have purchased 3 categories, 9.6% have purchased 4 categories and 5.2% have purchased all five categories. Table 2.4 shows the 3 most popular purchase sequence across these five categories. For customers who have only bought 1 category, most of them (71.6%) have bought Non-Apparel 2. For those who have bought only 2 categories, most of them (22.41%) first bought Non-Apparel 2 and then Tops & Bottoms. It seems that Non-Apparel 2 and Tops & Bottoms are typically purchased before the other categories. The purchase sequence could be due to various reasons, such as consumer's different prior expectations, different purchase cycles or correlations in the preference across categories. We account for these elements in our model.

Table 2.4: Top 3 Most Frequent Purchase Sequences

Number of Category Purchased	1st	2nd	3rd	4th	5th	Percent
1	Non-Apparel 2	-	-	-	-	71.60%
	Tops & Bottoms	-	-	-	-	11.11%
	Outerwear	-	-	-	-	8.23%
2	Non-Apparel 2	Tops & Bottoms	-	-	-	22.41%
	Non-Apparel 2	Outerwear				12.93%
	Tops & Bottoms	Non-Apparel 2	-	-	-	10.34%
3	Non-Apparel 2	Tops & Bottoms	Outerwear	-	-	10.45%
	Tops & Bottoms	Footwear	Non-Apparel 2	-	-	10.45%
	Tops & Bottoms	Footwear	Outerwear	-	-	10.45%
4	Non-Apparel 2	Tops & Bottoms	Outerwear	Footwear	-	10.42%
	Tops & Bottoms	Outerwear	Non-Apparel 2	Footwear	-	6.25%
	Tops & Bottoms	Outerwear	Footwear	Non-Apparel2	-	6.25%
5	Non-Apparel 2	Outerwear	Tops & Bottoms	Footwear	Non-Apparel 1	7.69%
	Non-Apparel 2	Tops & Bottoms	Outerwear	Footwear	Non-Apparel 1	7.69%
	Outerwear	Tops & Bottoms	Footwear	Non-Apparel 2	Non-Apparel 1	7.69%

The aggregate number of purchase incidence over time is shown in Figure 2.1. Consistent with the purchase sequence, the number of purchase incidence in Non-Apparel 2 starts significantly higher than the number of purchase incidence in the other categories, but it drops sharply over time.

Figure 2.1: Category Purchase Incidences over Time



Third, how is the number of categories bought related to the number of items bought per category? As shown in Table 2.5, the consumers who have made purchases in more categories tend to buy more products per category. It seems that multi-category buyers are better customers in terms of quantity purchased. In our model, we are going to attribute the purchases within and across categories to consumers' brand and category preference.

Table 2.5: Number of Item Purchased per Category by Number of Categories Purchased

Number of Categories Purchased	Number of Customers	Average Number of Items Purchased per Category	
		Mean	**Std. Dev.**
1	243	2.91	3.90
2	116	3.59	4.11
3	67	3.40	2.28
4	48	5.36	4.57
5	26	12.66	9.03

We observe when and what mailing catalogs are received by each consumer. On average, the consumers receive 8.62 catalogs every year. To characterize the product combination in each catalog, we calculate the percentages of each category in each catalog. As shown in Table 2.6, on average, Tops & Bottoms is the category most frequently shown in the catalogs among all categories.

Table 2.6: Summary Statistics for Category Percentage among All Catalogs

	Category Percentages in Catalogs	
	Mean	**Std. Dev.**
Tops & Bottoms	45.56%	18.78%
Footwear	14.98%	6.70%
Outerwear	10.88%	7.67%
Non-Apparel 1	12.99%	21.37%
Non-Apparel 2	15.59%	19.02%

2.5 Estimation

2.5.1 Empirical specification

In the empirical application, we model the purchase incidences in each category over time. It is assumed that a consumer's decision to make a purchase depends on the expected utility specified in equation 2.1. In addition to consumer's changing preferences for the brand and the categories, there could be other factors that influence consumer's purchase decisions. To control for these factors, we include the following covariates in the utility specification. First, we include price, time since last purchase, seasonality, i.e. $X_{kit} = [price_{kt}, timelast_{kit}, season_t]'$. The price variable $price_{kt}$ is a category level price index calculated as the ratio of average selling price to the average full price in category k during time t. Using the same dataset, Anderson et al. (2006) found that using the price index provides more sensible results than using the actual selling prices. The variable $timelast_{kit}$ is the number of periods since last purchase in that category. We estimate a category specific coefficient to control for the difference in the purchase frequencies across categories. The seasonality variable $season_t$ is a binary variable. It takes value 1 for months 1~6 and value 0 for the other months. We also include year dummies to control for the changes in the environment, such as competitive landscape or economic environment.

2.5.2 Identification

It is important to discuss the identification of the model parameters. All parameters can be grouped into the five groups: 1) prior means: { $\bar{C}_{ki0}, k=1,\dots,K$ } and $\bar{B}_{i0}$; 2) prior variances{ $\delta^2_{ki0}, k=1,\dots,K$ }, prior covariances between categories{ $r_{j,k,i,0}$, $j \neq k, k=1,\dots,K$ }, and prior covariances between brand and categories { $\varpi_{b,i,0}, k=1,\dots,K$ } and effect of product frequency in catalogs on brand-category covariance α_i ; 3) category quality to be learned { $c_{ki}, k=1,\dots,K$ }; 4) signal variance of product experience σ^2_i and mailing catalogs π^2_i; and 5) non-learning parameters { $\beta_{ki}, k=1,\dots,K$ }. However, not all these parameters are identified from the consumers purchase behaviors. In what follows, we discuss what parameters are identified and how they are identified.

The identification mainly comes from consumers' purchase patterns over time. As a consumer learns about the products and the brand, her preferences changes as well as her future purchase patterns. The more she likes the brands, the more frequently she would buy within and across the product categories. If she doesn't have much strong preference for the brand, but she likes a particular category, she would keep buying in that category. Therefore, the brand preference and the category preferences are identified from the purchase frequencies within and across categories.

To see the identification for each group of parameters, we can divide the time horizon into three periods. The prior learning period is the period between acquisition and first purchase. How long each consumer waits to make a first purchase is assumed to be associated with the consumer's prior expectation $\bar{Q}_{ki0}$ after controlling for other covariates. Consumers with higher prior expectations are expected to make a first purchase sooner after acquisition than consumers with lower prior expectations. However, since $\bar{Q}_{ki0} = \bar{B}_{i0} + \bar{C}_{ki0}$, we cannot separately identify $\bar{B}_{i0}$ and $\bar{C}_{ki0}$. To resolve this problem, we assume that a newly acquired consumer holds a neutral preference for the brand, i.e. $\bar{B}_{i0}$ = 0. With this assumption, the consumer's initial purchases are only influenced by her preference for the products. This assumption is consistent with the findings in the behavioral literature that consumers tend to use product attribute-based evaluations in their earlier phase of choice and brand-based evaluations in the later phase of choice as they become more familiar with the brand (Bettman and Park 1980).

The post-learning period is the period when the consumer's preferences and the purchase rates have converged to steady states. During this period, the intercepts in the utility $\bar{B}_{it} + \bar{C}_{kit}$ are not changed by any newly received information. Therefore, the intercept $\bar{B}_{it} + \bar{C}_{kit}$ and the non-learning parameters β_{ki} are identified just as in the regular Probit model. Furthermore, $\bar{B}_{it}$ and $\bar{C}_{kit}$ can be identified separately because of the learning process assumed restricts $\bar{B}_{it}$ and $\bar{C}_{kit}$ to be dependent, i.e., both of them are functions of the previous product information { E_{kit} }

and { S_{kit} }. From a consumer's purchase patterns over time, we can identify her average experiences $\bar{E}_{ki}$ and $\bar{S}_{ki}$, which allows us to back out the true category valuation c_{ki}.

The other parameters are identified from the consumer's speed of learning within and across categories. The sooner the consumer's preferences converge to steady states, the faster she learns. This can be explained either by accurate signals (small σ_i^2 and π_i^2) or large initial uncertainty (large δ_{ki0}^2). Since we can't identify both the signal variance and the prior variance (δ_{ki0}^2), we choose to normalize $\delta_{ki0}^2 = 1$, $k=1,\ldots, K$.

The information spillover effect driven by the prior covariance terms are identified from the cross-category purchase patterns over time. Since the information about one category may spill over to the other categories due to the perceived relationships between categories (r_{jki0}) or the perceived relationships between category and brand (ϖ_{ki0}) and r_{jki0} and ϖ_{ki0} cannot be separately identified, we choose to normalize ϖ_{ki0} =0, $k=1,\ldots, K$. This normalization implies that initially the categories and the brand are perceived to be uncorrelated. This is a reasonable assumption because consumers' brand preference is yet to be formed at that point. Therefore, the initial information spillover is only driven by perceived covariance at the product level. In this way, we can identify r_{jki0}. Over time, the cross-category spillover that cannot be explained by r_{jki0} can be attributed to the category-brand associations, which allows us to identify α_i.

2.5.3 Estimation

Since we are interested in estimating the learning process for each individual consumer, we specify the parameters at the individual level. Let the vector θ_i denote the vector of parameters belong to consumer i, i.e. θ_i =({ $\bar{C}_{ki0}, k=1,...K$ } , { r_{jki0} , $j \neq k, k=1,...K$ }, { $c_{ki}, k=1,...K$ }, σ_i^2 , π_i^2 , α_i and { $\beta_{ki}, k=1,...K$ })'. We use a hierarchical Bayesian approach to induce data shrinkage (Rossi, Allenby, and McCulloch 1996), since we have a large set of parameters to estimate but relatively few observations for some of the customers. By assuming a population distribution to restrict the individual level parameters, hierarchical Bayesian model can fit the data well while avoiding the problem of overfitting. The individual level parameters are specified as a function of customer characteristics, denoted as Z_i. Then,

$$\theta_i = \Pi Z_i + \eta_i$$

where η_i captures unobservable heterogeneity and it is assumed to be distributed as:

$$\eta_i \sim \text{MVN}\ (\mathbf{0},\ \Omega)$$

Variable Z_i includes customer demographics, such as age, household income, number of kids, marriage status, gender of the head of household and whether or not the customer is acquired by catalogs. These variables are collected in 2007 by the retailer. To complete the model, we specify the prior of Π and Ω as:

$$\pi = vec(\Pi) \sim N(\hat{\pi}, \Phi) \text{ and } \Omega \sim Wishart(g, G)$$

Given that we don't observe the signals, we treat them as the augmented latent variables (Tanner and Wong 1989). To reduce the number of signals to be estimated, we define

$Y_{kit}\pi_i^{-2}E_{kit} + ns_{kit}\sigma_i^{-2}S_{kit}$ as one latent variable. Since E_{it} and S_{it} are both normally distributed around c_{ki}, $Y_{kit}\pi_i^{-2}E_{kit} + ns_{kit}\sigma_i^{-2}S_{kit}$ is also normally distributed:

$$Y_{kit}\pi_i^{-2}E_{kit} + ns_{kit}\sigma_i^{-2}S_{kit} \sim N(m_{kit}, v_{kit})$$

where $m_{kit} = (Y_{kit}\pi_i^{-2} + ns_{kit}\sigma_i^{-2})c_{ki}$ and $v_{kit} = Y_{kit}\pi_i^{-2} + ns_{it}\sigma_i^{-2}$.

To obtain the posterior distributions of all the parameters, we take draws from the joint posterior distribution. Based on these model specifications, the joint posterior can be written as:

$$\begin{aligned}
&Posterior \propto \\
&\prod_i \left[\begin{array}{l} \prod_k \left[\begin{array}{l} \prod_{t=1}^{T} \left[\exp(-.5*(U_{kit} - \bar{B}_{it} - \bar{C}_{kit} - \beta_{ki} X_{kit})^2 I(U_{kit} > 0)^{Y_{kit}} I(U_{kit} < 0)^{1-Y_{kit}} \right] \\ \prod_{t=1}^{T-1} [v_{kit}^{-1} \exp(-.5*((T_{it}^{E}\pi_i^{-2}E_{it} + T_{it}^{S}\sigma_i^{-2}S_{it} - m_{kit})/v_{kit})^2)] \end{array} \right] \\ |\Omega|^{-1/2} \exp(-\frac{1}{2}(\theta_i - \Pi Z_i)'\Omega^{-1}(\theta_i - \Pi Z_i)) \end{array} \right] \\
&|\Phi|^{-1/2} \exp(-\frac{1}{2}(\pi - \hat{\pi})'\Phi^{-1}(\pi - \hat{\pi})') \frac{\Omega^{(g-K-1)/2}}{|G|^{g/2}} \exp(-\frac{1}{2} tr(G^{-1}\Omega))
\end{aligned}$$

We draw parameters from the joint posterior distribution using Gibbs Sampling, i.e. draw a set of parameters conditional on the others sequentially. When the conditional distribution is an unknown distribution, we use the Metropolis Hastings Algorithm (Chib and Greenberg 1995) to obtain these draws. The detailed algorithm is described in Appendix 2.3. Simulations are conducted to assure the recoveries of the parameters.

2.5.4 Simulation

The purpose of this simulation study is to make sure our proposed model and estimation strategy can recover the model parameters using simulated data. Without loss of generality, we simulate a case in which the retailer offers products in two categories. The simulated sample consists of 300 customers and 60 time periods. The individual level parameter θ_i is simulated from distribution: θ_i ~MVN (Π, Ω) where Π is a 10-element vector and Ω is assumed to be a 10*10 diagonal matrix. The true values of the aggregate level parameters Π and Ω are listed in column 3 of Table 2.7 and Table 2.8. In the estimation, we ran the MCMC chain for a total of 500,000 iterations to obtain draws from the full conditional distributions. Among all the draws, we discarded the first 100,000 as "burn-in" and kept the latter 400,000 to make inference. The posterior mean and the posterior standard deviation are calculated based on these draws.

The posterior means and standard deviation of the aggregate level parameters Π and Ω are reported in Table 2.7 and Table 2.8. The parameters are recovered with a reasonable precision, i.e., most of the estimates are within one standard deviation of the truth.

Table 2.7: True Π and Posterior Mean and Standard Deviation of Estimated Π

		True Parameter (Π)	**Posterior Mean**	**Posterior Standard Deviation**
Prior Mean	category 1	-.3	-0.297	.054
	category 2	-.3	-0.309	.074
True Mean	category 1	.5	0.512	.051
	category 2	.5	0.520	.057
Log (Signal Variance)	category 1	-4	-3.955	.067
	category 2	-4	-3.822	.126
Prior Correlation	category 1 and category 2	.5	0.433	.075
Effect of Catalog Intensity on Category-Brand Links		.1	.089	.016
Price Index Coefficients	category 1	-.08	-.081	.015
	category 2	-.08	-.082	.015

Table 2.8: True Σ and Posterior Mean and Standard Deviation of Estimated Σ

		True Parameter (diagonals of Σ)	Posterior Mean	Posterior Standard Deviation
Prior Mean	category 1	.05	.062	.014
	category 2	.05	.060	.014
True Mean	category 1	.05	.061	.014
	category 2	.05	.046	.015
Log (Signal Variance)	category 1	.05	.066	.015
	category 2	.05	.064	.015
Prior Correlation	category 1 and category 2	.05	.059	.015
Effect of Catalog Intensity on Category-Brand Links		.05	.06	.006
Price Index Coefficients	category 1	.05	.048	.006
	category 2	.05	.049	.006

Note: To save space, we only show the diagonals of Σ.

2.6 Results

2.6.1 Do consumers learn?

To assess how much consumers learn in each category, we can compare $\bar{C}_{ki0}$ and c_{ki}. In Table 2.9, we report the mean and standard deviation of $\bar{C}_{ki0}$ and c_{ki} across all consumers. First, we notice that among all categories the initial expectation about Non-apparel 2 (0.202) is the highest, followed by Top & Bottoms (-0.39). The initial expectation about Footwear and Non-apparel 1 are the lowest (-0.74). The rank of the initial beliefs is consistent with the typical order in which the consumers make purchases in these categories as shown in Table 2.4. Second, we find that the difference between c_{ki} and $\bar{C}_{ki0}$ as an estimate of the accumulative disconfirmation are all positive except for Non-Apparel 2. Since Non-Apparel 2 is the category where most customers make the first purchases, its negative disconfirmation could be a concern.

Table 2.9: Summary Statistics for Posterior Mean of Each Consumer's $\bar{C}_{\mathrm{ki0}}$ and c_{ki}

	Across Consumer Mean (std dev.) of $\bar{C}_{\mathrm{ki0}}$	Across Consumer Mean (std dev.) of c_{ki}	Disconfirmation c_{ki} - $\bar{C}_{\mathrm{ki0}}$
Tops & Bottoms	-0.390 (0.203)	-0.142 (0.295)	0.247*
Footwear	-0.743 (0.169)	-0.183 (0.232)	0.560*
Outerwear	-0.662 (0.132)	-0.121 (0.192)	0.541*
Non-Apparel 1	-0.746 (0.127)	-0.210 (0.204)	0.536*
Non-Apparel 2	0.202 (0.325)	-0.043 (0.082)	-0.245*

* The difference is significant at 5% level.

Notice that the differences across categories in c_{ki} are much smaller than in $\bar{C}_{ki0}$. This suggests after learning consumers' expectations across these categories become more consistent.

2.6.2 How do consumers learn?

There are three sets of parameters describing how consumers learn: the perceived category covariance in the initial period ({ r_{jki0} , $j \neq k, k = 1,...K$ }), the effect of category frequency in catalogs on category-brand associations (α_i), and the signal variances of usage experiences and mail catalogs (σ_i^2 and π_i^2). First, the mean and the standard deviation (across consumers) of r_{jki0} are reported Table 2.10. These parameters describe consumers' perceived covariance (similarity) prior to any purchases.

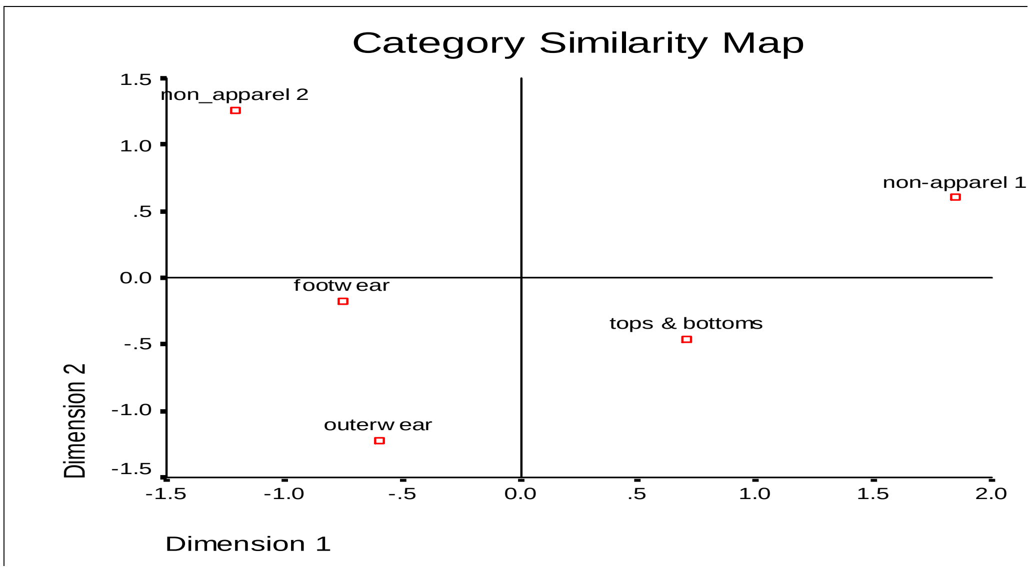
Table 2.10: Summary Statistics for Posterior Mean of Each Consumer's r_{jki0}

Prior Covariance	*Tops & Bottoms*	*Footwear*	*Outerwear*	*Non-Apparel 1*
Tops & Bottoms	---			
Footwear	0.372 (0.221)	---		
Outerwear	0.240 (0.259)	0.166 (0.205)	---	
Non-Apparel 1	0.272 (0.196)	0.008 (0.150)	-0.018 (0.141)	---
Non-Apparel 2	0.027 (0.215)	0.194 (0.193)	0.094 (0.184)	-0.016 (0.234)

Note: The table reports the means and standard deviations (in parentheses) across consumers.

Based on these estimates, we can use Multidimensional Scaling method (MDS) to place these five categories on a two dimensional similarity map as shown in Figure 2.2. The dimension 1 in the map can be interpreted as outdoor / non-outdoor and the dimension 2 can be interpreted as apparel / non-apparel. The distance between two categories measures the dissimilarity. For example, tops & bottoms, outerwear and footwear are perceived to be more similar to each other than to the other products. Since we have estimates for each individual customer, we can produce this similarity map for each of them.

Figure 2.2: Category Similarity Map Using MDS Based on Estimated Initial Prior Covariance between Categories



While such category similarities based on physical attributes in consumers' minds are not under the influence of the retailer, how these categories are related to the overall brand is assumed to be influenced by the retailer's catalogs. Recall that we assume the perceived category-brand association is a function of category frequency. The coefficient capturing this effect is α_i. In Table 2.11, we report the mean and the standard deviation of α_i across consumers. On average, α_i is positive (0.394) which means the more catalogs received about the category, the stronger association the consumer held between the category and the brand.

Table 2.11: Mean and Standard Deviation of Category-Brand Association α_i

	Across-Consumer Mean	Across-Consumer Std. Deviation
α_i	0.394	0.117

How fast consumers learn not only depends on consumers perceived brand structure, but also the signal variances. As shown in Table 2.12, the signal variance of direct experiences is on average 0.663, which is lower than the signal variance of mail catalogs (0.916). This suggests that information from direct usage experiences is perceived to be more accurate than information from advertising. This is consistent the findings in Wright and Alice (1995).

Table 2.12: Experience Signal Variance and Catalogs Signal Variance

Signal Variance	Variables	Across-Consumer Mean (Std. Dev.)
Experience	σ_i^2	0.663 (0.181)
Catalog	π_i^2	0.916 (0.188)

2.6.3 Non-learning parameters

In Table 2.13, we report the mean and the standard deviation of price, time since last purchase, and seasonality across consumers. At the aggregate level, consumers are significantly less price sensitive in Tops & Bottoms and Non-Apparel 2 than in the other categories (significant at 5% level). This suggests a possibility that the category price sensitivity might be related to the category preference. The category specific time-since-last-purchase coefficient is to control the different purchase cycles across categories. A large coefficient indicates a short purchase cycle. Tops & Bottoms and Non-Apparel 2 have significantly shorter purchase cycles than the other categories (significant at 5% level). In terms of seasonality, non-apparel 2 is the category has the most significant seasonality.

Table 2.13: Model Estimates for Non-learning Parameters

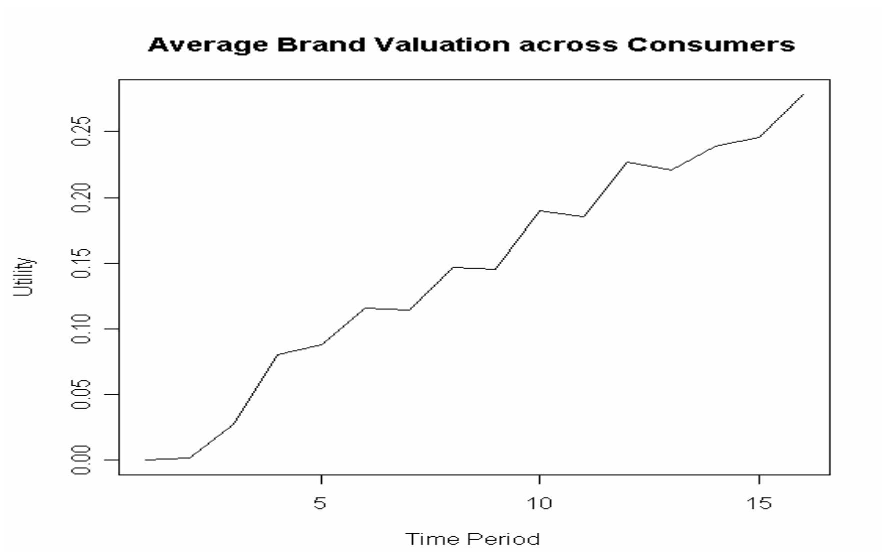
Prior means	**Across-Consumer Mean (Std. Dev.)**
Price_Tops & Bottoms	-0.468 (0.323)
Price_ Footwear	-.0.522 (0.183)
Price_ Outerwear	-0.507 (0.247)
Price_Non-Apparel 1	-0.561 (0.159)
Price_ Non-Apparel 2	-0.403 (0.357)
Time since last purchase_ Tops & Bottoms	-0.043 (0.112)
Time since last purchase_ Footwear	-0.094 (0.165)
Time since last purchase_ Outerwear	-0.088 (0.159)
Time since last purchase_ Non-Apparel 1	-0.273 (0.203)
Time since last purchase_ Non-Apparel 2	0.042 (0.133)
Season_Tops & Bottoms	-0.226 (0.209)
Season_ Footwear	-0.177 (0.153)
Season_ Outerwear	-0.297 (0.171)
Season_Non-Apparel 1	-0.099 (0.260)
Season_ Non-Apparel 2	-0.538 (0.257)

2.7. Managerial Implications

2.7.1 Measuring brand equity

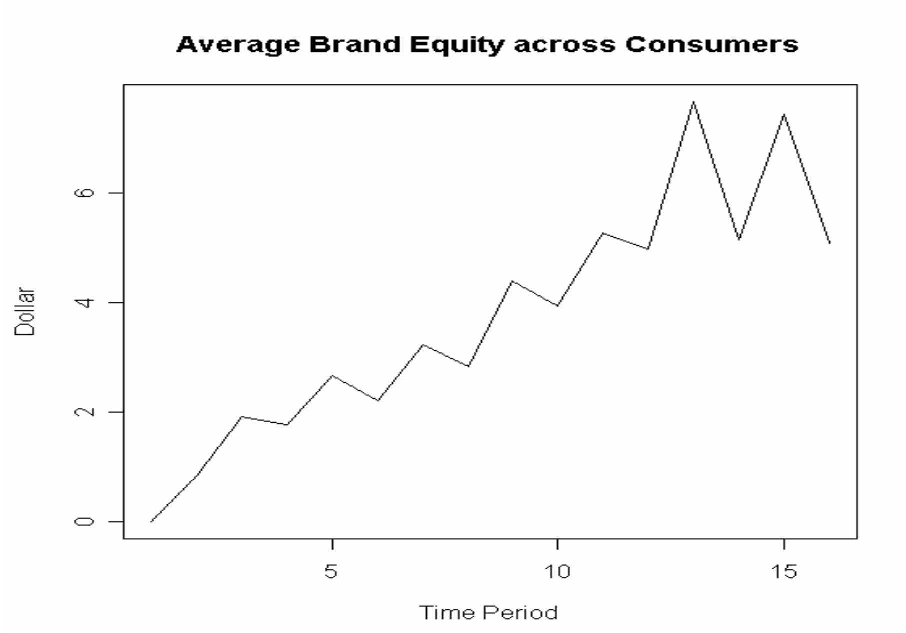
As a result of the learning process, consumers form preferences for the overall brand. An important contribution of this model is that it provides estimates to track the change in $\overline{B}_{it}$. Recall that the initial mean brand valuation $\overline{B}_{i0}$ is normalized to 0. How each consumer's mean brand valuation changes from 0 to $\bar{B}_{iT}$ depends on the signals they have received during the learning process. Since the signals are not observable to us researchers, the path of $\bar{B}_{it}$ is random and only the average is identified. So, we simulate the learning process for each consumer 100 times and take the average as an estimate of the path. In Figure 2.3, we plot the average of all consumers' estimated $\bar{B}_{it}$. We can see that the average $\bar{B}_{it}$ increases over time from 0 to 0.26.

Figure 2.3: Evolution of Average Brand Valuation over Time (in Utility)



This increase in brand valuation can be translated into dollars, which is brand equity. Following the revenue premium definition of brand equity (Ailawadi et. al. 2003), we define brand equity as the incremental revenue that is due to the brand. To calculate the incremental revenue, we conduct a counterfactual experiment using our model and the estimated parameters to simulate revenues with and without the brand name. This counterfactual experiment describes two scenarios. In these two scenarios, the retailers sell exactly the same products. The only difference is that in the without brand case, all products are not branded and therefore as consumers learn, there is no preference accumulating to the brand. The brand equity is calculated as the difference between the revenues in these two scenarios. As shown in Figure 2.4, the average brand equity across consumers increase from \$0 per period (6 months) to about \$6. This means the consumers are on average willing to pay \$6 dollar more for only the brand name every 6 months.

Figure 2.4: Evolution of Average Brand Equity over Time (in Dollar)

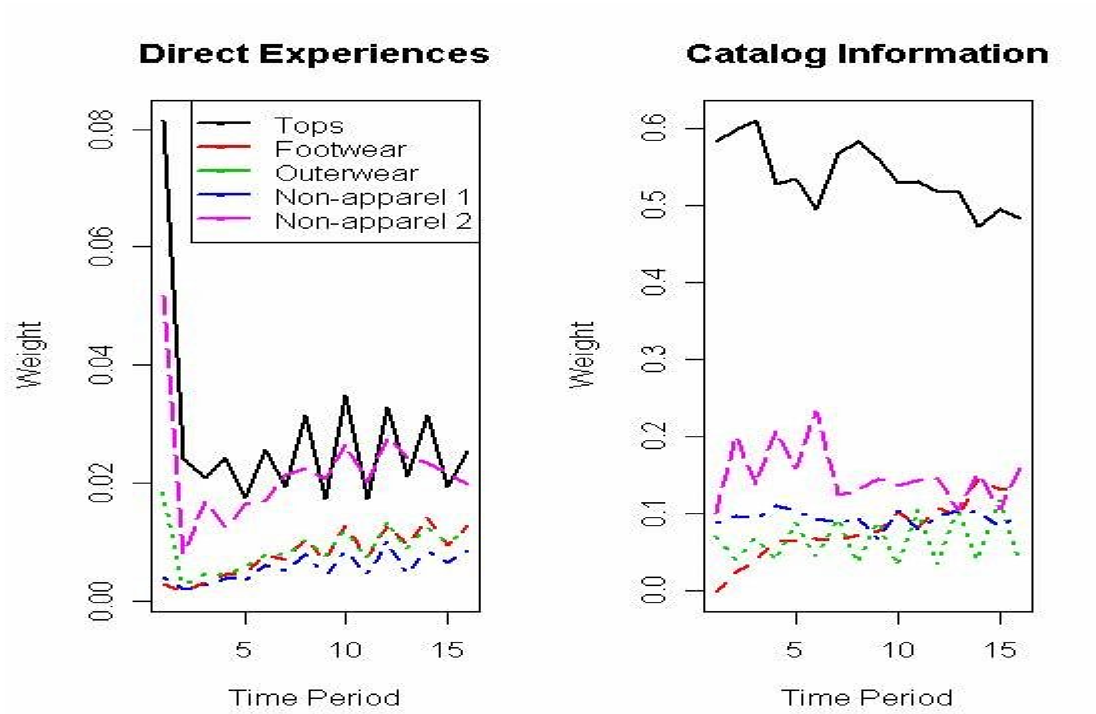


2.7.2 Key category

In section 2.6.1, we present the results for disconfirmation in each category. How much the disconfirmation influences the overall brand equity also depends on the category's disconfirmation weight. Based on the equation 2.5 in section 2.3.4, we calculate the category disconfirmation weight for each consumer and each information type at each time period. Averaging across all consumers, we obtain the category disconfirmation weight over time.

Suppose the total weight over all categories and both information types is one. We can decompose the total weight by categories and information types. In Figure 2.5, we plot the category disconfirmation weight by categories and information types. The first thing to notice is that across two information types catalogs information has larger weight in influencing the brand value than direct experiences. Since this is a catalog retailer selling durable goods, it is reasonable that the high frequency of catalogs received outweighs the low frequency of direct experiences. Also we notice that over time the weight of direct experiences decay much faster than the catalog information. This means the learning from direct product experiences is faster, whereas the learning from catalogs is more gradual.

Figure 2.5: Evolution of Disconfirmation Weight by Category and by Information Type



For direct experiences, Tops & Bottoms and Non-Apparel 2 have higher weights than the other three categories. The differences become smaller over time. While the weights of Tops &

Bottoms and Non-Apparel 2 decrease significantly, the weights of the other categories remain at a constant level or even increase over time. This suggests that the initial experiences in Tops & Bottoms and Non-Apparel 2 are very important in forming brand equity in consumers' minds. For catalog information, Tops & Bottoms has significantly higher weight than the other categories, followed by Non-apparel 2.

Overall, Tops & Bottoms and Non-apparel 2 are the key categories that have higher weights in influencing the brand value. However, these categories have lower disconfirmations than the other categories. These results suggest to the retailer that in order to improve the overall brand value, they need to either improve the disconfirmations in these key categories or improve the weights of the other categories with higher disconfirmations. To improve the weights of certain categories, the retailer can enhance the association between the categories and the brand by sending more catalogs in these categories and inducing consumers to buy more often in these categories.

2.7.3 Heterogeneity

To understand consumer heterogeneity is important, for direct marketers, in order to sell the right products, to the right consumers and at the right time.

First of all, heterogeneity is found to be significant since the across-consumer standard deviations are large relative to the across-consumer means as shown in Table 2.10, Table 2.11, Table 2.12, and Table 2.13. In our model, the observable heterogeneity can be captured by Π. In Table 2.14, we report the posterior mean and standard deviation of Π.

Table 2.14: Posterior Mean and Standard Deviation of Π

Variables	Intercept	Income	Age	Female	Acquired by a Purchase	Acquisition Channel (Phone)	Acquisition Channel (Internet)
Initial Prior Mean: Tops & Bottoms	-0.117* (0.061)	0.065* (0.020)	-0.453* (0.181)	-0.174* (0.035)	0.317* (0.048)	-0.195* (0.028)	-0.184* (0.050)
Initial Prior Mean: Footwear	-0.509* (0.059)	-0.042* (0.022)	-0.132 (0.138)	-0.108* (0.031)	0.139* (0.033)	-0.207* (0.050)	-0.062 (0.045)
Initial Prior Mean: Outerwear	-0.481* (0.046)	-0.111* (0.030)	-0.095 (0.112)	-0.194* (0.038)	-0.061 (0.062)	0.141* (0.032)	0.175* (0.047)
Initial Prior Mean: Non-Apparel 1	-0.840* (0.036)	-0.019 (0.013)	0.126* (0.063)	-0.093* (0.029)	-0.010 (0.029)	0.107* (0.028)	0.326* (0.035)
Initial Prior Mean: Non-Apparel 2	-0.167* (0.061)	-0.073* (0.023)	0.255* (0.089)	-0.061* (0.025)	0.475* (0.025)	0.179* (0.030)	0.421* (0.041)
True Quality: Tops & Bottoms	-0.579* (0.063)	0.188* (0.024)	-0.011 (0.130)	0.343* (0.035)	-0.006 (0.039)	-0.004 (0.046)	0.209* (0.066)
True Quality: Footwear	0.141* (0.060)	-0.182* (0.034)	-0.005 (0.079)	0.065* (0.018)	0.271* (0.041)	-0.441* (0.042)	-0.428* (0.035)
True Quality: Outerwear	0.002 (0.062)	-0.152* (0.041)	-0.076 (0.103)	-0.025 (0.031)	-0.155* (0.051)	0.079 (0.040)	0.313* (0.037)
True Quality: Non-Apparel 1	-0.196* (0.050)	-0.130* (0.039)	0.739* (0.123)	-0.198* (0.032)	0.084 (0.080)	-0.046 (0.047)	-0.312* (0.059)
True Quality: Non-Apparel 2	-0.174* (0.053)	-0.028 (0.025)	0.121 (0.101)	0.054 (0.046)	0.029 (0.056)	0.101* (0.039)	0.014 (0.048)
Prior Covariance: Tops vs. Footwear	0.704* (0.063)	-0.112* (0.024)	-0.416* (0.132)	0.171* (0.022)	-0.214* (0.019)	-0.053 (0.059)	-0.389* (0.055)
Prior Covariance: Tops vs. Outerwear	-0.230* (0.038)	0.141* (0.023)	0.589* (0.071)	0.194* (0.021)	-0.439* (0.024)	0.239* (0.033)	0.104* (0.028)
Prior Covariance: Tops vs. Non-Apparel 1	0.518* (0.029)	0.103* (0.018)	-0.318* (0.083)	-0.085* (0.022)	-0.317* (0.032)	-0.035 (0.024)	-0.017 (0.047)
Prior Covariance: Tops vs. Non-Apparel 2	-0.264* (0.069)	0.092* (0.028)	-0.167* (0.073)	-0.005 (0.023)	0.073* (0.020)	0.253* (0.080)	0.555* (0.104)
Prior Covariance: Footwear vs. Outerwear	-0.295* (0.153)	0.079* (0.034)	0.370 (0.272)	0.248* (0.035)	0.000 (0.062)	-0.011 (0.063)	0.340* (0.052)
Prior Covariance: Footwear vs. Non-Apparel 1	0.022 (0.043)	0.071 (0.038)	-0.172* (0.079)	0.006 (0.052)	0.147* (0.032)	-0.062 (0.044)	-0.143* (0.041)
Prior Covariance: Footwear vs. Non-Apparel 2	0.019 (0.033)	-0.142* (0.025)	-0.030 (0.077)	0.059* (0.024)	0.076* (0.032)	0.320* (0.044)	0.292* (0.041)
Prior Covariance: Outerwear vs. Non-Apparel 1	-0.220* (0.097)	0.133* (0.037)	-0.035 (0.209)	-0.007 (0.044)	0.220* (0.054)	0.016 (0.048)	0.085 (0.066)
Prior Covariance: Outerwear vs. Non-Apparel 2	0.209 (0.129)	0.050 (0.032)	-0.106 (0.161)	-0.323* (0.040)	-0.147* (0.046)	0.184* (0.042)	0.108* (0.047)
Prior Covariance: Non-Apparel 1 vs. Non-Apparel 2	0.480* (0.049)	0.205* (0.013)	-0.799* (0.101)	-0.125* (0.025)	-0.017 (0.034)	-0.303* (0.051)	-0.413* (0.031)

Table 2.14 (Continued)

Variables	Intercept	Income	Age	Female	Acquired by a Purchase	Acquisition Channel (Phone)	Acquisition Channel (Internet)
Exp(Signal Variance) Experiences	-0.037 (0.051)	0.140* (0.034)	-0.485* (0.127)	-0.072* (0.030)	-0.347* (0.106)	-0.182* (0.045)	-0.245* (0.032)
Exp(Signal Variance) Catalog	0.051 (0.056)	-0.040 (0.031)	-0.065 (0.082)	0.164* (0.029)	-0.228* (0.029)	-0.218* (0.038)	-0.076 (0.062)
Catalog Intensity on Brand-category Association	-0.106 (0.062)	-0.186* (0.024)	-0.637* (0.090)	-0.316* (0.033)	-0.211* (0.041)	-0.147* (0.034)	-0.329* (0.032)
Price: Tops & Bottoms	-0.842* (0.110)	-0.214* (0.024)	0.725* (0.250)	0.038 (0.059)	0.186* (0.041)	0.224* (0.051)	0.206* (0.065)
Price: Footwear	-0.625* (0.065)	-0.053* (0.025)	0.095 (0.150)	-0.149* (0.048)	0.194* (0.028)	0.218* (0.035)	0.131* (0.035)
Price: Outerwear	-0.179* (0.053)	0.109* (0.030)	-0.670* (0.126)	-0.079* (0.042)	0.405* (0.043)	-0.399* (0.044)	-0.185* (0.048)
Price: Non-Apparel 1	-0.698* (0.099)	0.134* (0.027)	0.444* (0.152)	-0.197* (0.029)	-0.121* (0.050)	0.075 (0.061)	-0.005 (0.059)
Price: Non-Apparel 2	-1.070* (0.083)	0.026 (0.030)	0.161 (0.152)	0.268* (0.026)	0.503* (0.034)	0.301* (0.030)	0.182* (0.035)
Time since last purchase: Tops & Bottoms	-0.060 (0.047)	0.014 (0.014)	0.159* (0.066)	-0.001 (0.020)	-0.117* (0.017)	0.000 (0.028)	-0.054 (0.033)
Time since last purchase: Footwear	-0.179* (0.052)	0.109* (0.022)	-0.670* (0.078)	-0.079* (0.022)	0.405* (0.022)	-0.399* (0.030)	-0.185* (0.035)
Time since last purchase: Outerwear	-0.122* (0.062)	0.033 (0.020)	0.098 (0.097)	0.023 (0.033)	-0.069* (0.028)	0.004 (0.035)	-0.104* (0.039)
Time since last purchase: Non-Apparel 1	-0.292* (0.073)	0.085* (0.022)	-0.193 (0.109)	0.131* (0.036)	-0.018 (0.032)	-0.063* (0.034)	-0.112* (0.039)
Time since last purchase: Non-Apparel 2	0.184* (0.041)	0.020 (0.020)	-0.381* (0.070)	0.022 (0.020)	-0.106* (0.022)	0.031 (0.025)	0.012 (0.035)
Seasonality: Tops & Bottoms	-0.174 (0.109)	-0.052 (0.031)	-0.071 (0.144)	-0.256* (0.114)	0.242* (0.037)	0.107* (0.034)	0.242* (0.035)
Seasonality: Footwear	-0.300* (0.080)	0.036 (0.033)	0.028 (0.170)	0.255* (0.041)	0.078* (0.021)	-0.163* (0.044)	-0.200* (0.059)
Seasonality: Outerwear	-0.275* (0.058)	0.008 (0.025)	0.155 (0.144)	0.029 (0.032)	-0.188* (0.018)	0.040* (0.023)	-0.257* (0.058)
Seasonality: Non-Apparel 1	0.190* (0.088)	-0.371* (0.024)	0.007 (0.096)	-0.120* (0.031)	0.010 (0.025)	0.087 (0.059)	0.217* (0.092)
Seasonality: Non-Apparel 2	-0.386* (0.075)	0.011 (0.031)	0.5378 (0.120)	-0.088* (0.043)	-0.225* (0.037)	-0.321* (0.042)	-0.262* (0.067)

***** “significant” means either at least 97.5% of the posterior mass is above 0 or 97.5% is below 0.

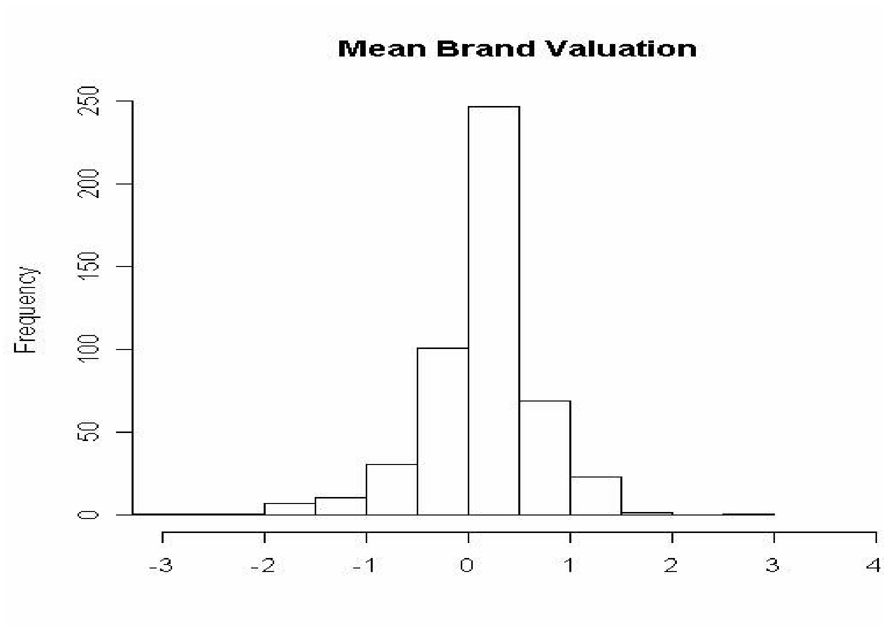
Overall, many of the demographic variables are significant in explaining the heterogeneity in the model parameters. For example, we find that females in general have lower initial prior means than males in all categories, which means that they wait longer to make their first purchase after being acquired. They also have smaller signal variances for direct experiences, but larger signal variances for catalog information. This means that they weigh direct experiences more than males do, but weigh less catalog information than males do in the learning process. Also, their brand-category associations are less influenced by the category frequency in the catalog.

Age is also significant in explaining some of the model parameters. For example, older customers tend to perceive Tops & Bottoms and Outerwear to be more similar (the only two clothing categories), but the other pairs less similar than the young customers. They also weigh the direct experiences more than younger customers in their learning process (smaller signal variance of direct experience). Their brand-category associations are less influenced by the category frequency in the catalog, compared to younger customers and female customers.

How consumers are acquired also matters to explain consumers' learning. Those who are acquired by making a first purchase have higher initial prior means than those who are acquired by other means (gift card registration, catalog request, etc). They also learn faster, since their signal variances of direct experiences and catalog information are smaller. In terms of acquisition channels (mail, phone or internet when first registered on the retailer's database), customers who are first acquired through phone or internet have smaller signal variance for direct experiences and lower category-brand associations than customers who are acquired through mails.

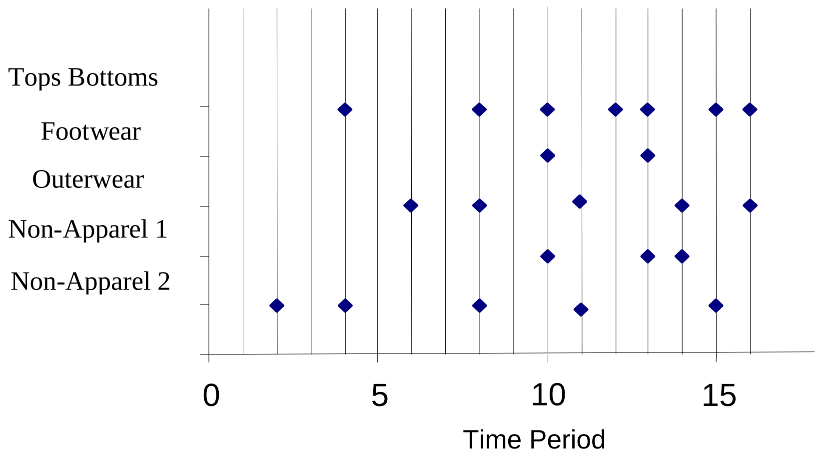
Retailers also want to understand the heterogeneity in brand equity. Consumers' brand valuations (brand equity) are found to be heterogeneous. In Figure 2.6, we plot the histogram of mean brand valuations across consumers. The mean is 0.16 and standard deviation is 0.88.

Figure 2.6: Histogram of Consumers' Mean Brand Valuations



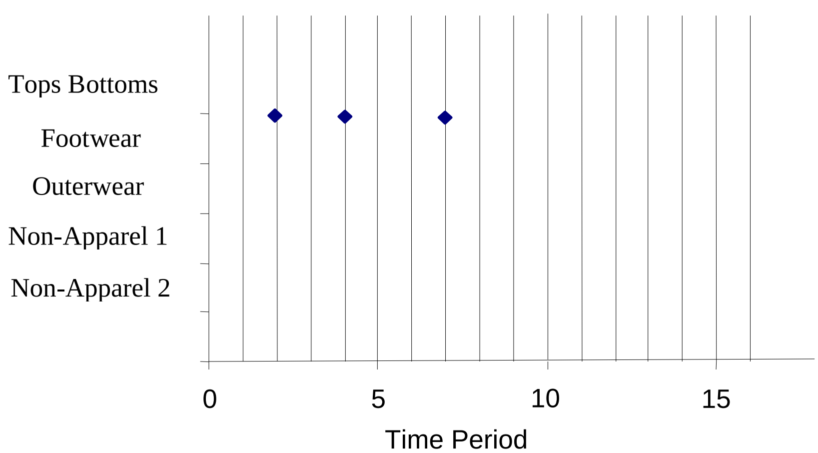
To see how the estimated brand equity relates to consumers purchase patterns, we show the purchase patterns of two consumers in our data as an example. One customer has high estimated average brand equity and the other has low estimated average brand equity. As shown in Figure 2.7, the high brand equity customer purchases in all categories and his/her purchase rates increase over time across most categories. In contrast, the low brand equity customer purchases only in Tops & Bottoms and his/her purchase rate decreases over time as shown in Figure 2.8.

Figure 2.7: Purchase Patterns of the Customer with High Brand Equity



* Note: Each dot represents a purchase incidence during a 6-month period

Figure 2.8: Purchase Patterns of the Customer with Low Brand Equity



* Note: Each dot represents a purchase incidence during a 6-month period

To identify consumers with higher brand valuation, we run a regression of customer's average brand valuation (i.e. $\sum_{t=1}^{T} \overline{B}_{it} / T$) on demographics. As shown in Table 2.15, customers who have higher income, who are older, female, acquired through internet have higher average brand valuation.

Table 2.15: Regression of Average Brand Valuation on Demographics

	Parameter Estimates (Std. Error)
Intercept	1.033 (0.151)
Income (10K)	0.0207*** (0.007)
Age	0.011*** (0.003)
Female	0.402*** (0.812)
Acquired at First Purchase	0.015 (0.809)
Acquired through Phone	0.123 (0.095)
Acquired through Internet	0.552*** (0.108)
Adjusted R-square	0.136

***: Significant at 1% level

The key category is also different in different consumers' eyes. For each consumer, we pick the category with the largest accumulative disconfirmation weight to be the key category. In Table 2.16, we show the frequency of a category being considered as a key category for direct experience and catalog information separately. In terms of direct experience, Tops & Bottoms is the key category for 221 customers and Non-apparel 2 is the key category for 244 customers. Considering that on average consumers have negative disconfirmation in Non-Apparel 2, the retailer needs to be concerned about the significantly negative impact of experiences in Non-

Apparel 2 on brand equity. In terms of catalog information, there is less heterogeneity. Tops & Bottoms is the key category for 495 customers. The good news is that not many consumers have Non-Apparel 2 as the key category in terms of catalog information.

Table 2.16: Frequency of Being a Key Category by Information Types

	Direct Experience	**Catalog Information**
Tops & Bottoms	221	495
Footwear	10	0
Outerwear	15	0
Non-Apparel 1	10	4
Non-Apparel 2	244	1

2.8 Conclusion

In this paper, we model the process of brand equity formation from the consumer learning's perspective. The model decomposes consumers' utility of consuming a product into the brand component and the product specific component. Since consumers are initially uncertain of these components, they learn about them over time from direct product experiences and firm's mailing catalogs. To model this learning process, we specify consumers' belief about the utility of the brand and its products as an associative network. This network of beliefs is then integrated into a Bayesian learning framework and estimated from consumers' purchase incidences across categories.

Our model contributes to the brand equity research in several aspects. First, it provides a tool to track individual consumer's retail brand equity based on his/her purchase behavior. Since the transaction data is usually readily available for retailers, our model is a more cost effective way to track individual customer brand equity than survey-based methods. Second, our theoretical derivation of disconfirmation and disconfirmation weight points out factors that influence brand equity formation. Then we empirically estimate the contribution of each product category in driving the evolution of overall brand equity. Third, we empirically estimate the associative network for all products. This perceived structure of the product portfolio has implications on the retailer's cross selling strategies. Lastly, since all estimates are at the individual level, we can identify brand equity, key category and associative network for each consumer. This is helpful for direct retailers to design more effective customized marketing strategies.

We conclude with a few comments on some of the limitations. First, although we have carefully controlled factors that might influence consumers' purchases other than brand and category preferences, such as seasonality, overall economic situation (using year and seasonality dummies), consumers' life cycles (using age), it is still likely that some factors are not included in the model, such as competition. This is more a data limitation, rather than a model limitation. If we have competitors' data, it is easy to fix this problem. Second, our model has not accounted for consumer experimentations. Consumers may leverage their knowledge in one category to make better purchase decisions in the other categories. This provides enough motivations for consumers to experiment to collect more information. Ideally, a learning model of forward looking consumers better describes consumer behavior in this context. Considering the complexity of the model, we leave it for future work.

Chapter 3

How Does Price Influence Product Returns?

3.1 Introduction

Sophisticated inventory models must account for not just new merchandise but also the flow of returned merchandise. While optimization of inventory is often sophisticated, the prediction of returns behavior is generally not as advanced. Most inventory models assume either that returns occur as a fixed proportion of sales or that returns are independent of sales. Surprisingly, a search of the literature reveals little empirical research describing the how marketing actions, such as pricing, affect customer return behavior. This absence of empirical work occurs despite recognition in the theoretical literature that poor estimates of returns behavior can significantly increase total inventory management costs (de Brito and van der Laan, 2002).

In this paper we examine the relationship between prices and returns. We provide evidence from a large-scale field experiment conducted with a women's clothing catalog to reject the straw-man hypothesis that a constant proportion of items are returned. Instead we will show that the rate of returns varies according to the price of an item. To understand why price may impact returns, we develop a model incorporating two effects. We label these effects the *perceived value* and *incremental customer* effects. The perceived value effect predicts that discounted items are less likely to be returned by a consumer as the lower price may compensate for disutility in product fit. The incremental customer effect recognizes that the mix of customer types may change when an item is discounted. A low price may attract customers with either high or low return propensities and therefore the impact of this effect on returns is ambiguous.

To test and measure these effects empirically, we build a joint model of consumer's purchase and return behavior and estimate it on a large scale dataset from a multi-category catalog retailer. The model improves our understanding of the price effect on returns in several aspects. First, we confirm our predictions that a discounted price reduces returns through the perceived value effect, but has an ambiguous effect on returns through the incremental demand effect. We find that in women's categories lower prices attract consumers with lower return propensities and therefore reduce aggregate return rates, whereas in kids' categories lower prices attract consumers with higher return propensities and increase aggregate return rates. In men's categories, we find no incremental customer effect. Second, the model allows us to measure the size of the perceived value effect and the incremental customer effect. Using simulations, we find that the perceived value effect and the incremental customer effect have substantial impact on returns. Ignoring these two effects and assuming a constant return rate would lead to overestimate or underestimate in the total number of returns (overestimate by 35%, 42% , 28% , 39%, and 9% for women's casual, outerwear, dress, men's casual and outerwear respectively, underestimate by 6% and 2% for kids casual and outerwear respectively). Furthermore, the incremental customer effect is found to be larger in size than the perceived value effect. Third, the model provides individual level estimates which are useful for retailers to design effective targeted price promotion strategies to prevent returns.

Our paper contributes to the existing literature in three ways. First, theoretically and empirically we reject the constant return rate assumption, a widely accepted assumption in the inventory management literature. These findings highlight the need for further research on this topic. Second, our model provides a framework to understand the effect of prices on returns. It is also a better tool to predict returns than the extant empirical models. More importantly, it provides new insight into managing returns: understanding customer heterogeneity and product heterogeneity is the key to designing targeted marketing strategies so that returns are effectively managed before sales take place.

It is important to clarify that this paper focuses on returns of unwanted merchandise by customers. Another common reason for returns is recycling of consumed merchandise for remanufacturing. For example, printer cartridges, disposable cameras, and automobile parts are often returned for remanufacturing (Rogers and Tibben-Lembke 2001). This is an important source of returns in some industries, but is not considered in this paper. It is also helpful to clarify our terminology. We use the term “rate of returns” (or “return rate”) to describe the proportion of items that a customer purchases and then subsequently returns. We distinguish this proportion from the “number of returns”, which represents a count of how many items are returned.

3.2 Previous Literature

The field of inventory management includes a wide range of models designed to support production planning and procurement processes. All of these models require a prediction of the relationship between sales and returns. Returns are typically assumed to be a constant proportion of sales, so that if a retailer sells more items the number of items returned will increase (Kiesmüller and van der Laan 2001, Savaskan et al. 2004). In remanufacturing contexts, researchers have assumed that returns are independent of sales (see for example Fleischmann et al. 2002).

In consumer product return literature, several theoretical papers have modeled how retail return policies affect customer purchase and return decision and consequently on firm profits (Davis et al. 1995, 1998 and Che 1996). In these models, a consumer decides to return a product if the residual consumption value is less than the consumer's value from returning the product: the refunded price minus the hassle cost involved. This assumption implies that for each individual consumer as price increases, the utility from returning the product increases and therefore the consumer is more likely to return the product.

While the literature on product return is extensive, it offers few empirical studies investigating how price affects product returns. One exception is Hess and Mayhew (1997), who analyze customer returns to an apparel catalog. The authors estimate a logit model to predict return rates and find that more expensive items are more likely to be returned. Note that this finding is at the aggregate level across consumers and products. In our paper, we further explore the mechanism of how price affects returns by accounting for consumer heterogeneity in the purchase and return decisions. We will present evidence that at aggregate level, the *proportion* of products that are returned may increase or decrease depending on consumer and product heterogeneity. Considerable variations across consumers and products have also been found in

Anderson, Hansen and Simester (2006) when they measure the option value of returns for consumers.

The paper also contributes to an emerging research stream that recognizes the need to coordinate marketing and operations decisions (Ho and Tang 2004). While research activity is growing, published research on the issue still remains somewhat limited. For example, Karmarkar (1996) points to “a lack of applied research that extends across marketing and manufacturing parameters and has consequences for practice” (p. 127). Our search of the literature revealed limited empirical research on either intra-firm coordination or inter-firm coordination between marketing and operations decisions. One exception is Kulp, Lee and Ofek (2004), who conduct a large-scale survey to investigate the value of inter-firm coordination between manufacturers and retailers. They find that there are limited gains from information sharing. They do report that collaborative initiatives in inventory management and new products and services increase performance, but caution that inter-firm coordination on reverse logistics programs can lead to the unexpected consequence of greater manufacturer stockouts.

A number of theoretical models have investigated inter-firm and intra-firm coordination. Eliashberg and Steinberg (1987) examine coordination of price and inventory policy in an industrial supply chain. Researchers have also examined the integration of marketing programs with operations decisions. This includes customer reward programs and capacity decisions (Kim, Shi and Srinivasan 2004) and customer advance booking programs with production policies (Tang, Rajaram and Alptekinoglu 2004). In related work, Hess and Lucas (2004) examine how a firm should allocate scarce resources between marketing and manufacturing.

The remainder of this paper is organized as follows. In section 3.3, we present a pilot study to investigate the effect of price on returns. In section 3.4, we develop hypotheses from a theoretical model of customer return behavior. We then test these hypotheses using our empirical model and discuss the findings in section 3.5. In section 3.6, we conduct simulation studies to explicitly measure the size of the perceived value effect and the incremental customer effect by product categories and discuss the managerial implications. The paper concludes in section 3.7.

3.3 Pilot Study

Before fully modeling the effect of price on returns, we conduct a pilot study to test the straw-man assumption that return is a constant proportion of demand, independent of price. In this study, we conduct a price experiment to establish the causal relationship between price and product returns.

The pilot study was conducted in a mail-order catalog that sells women’s fashion clothing, in the plus-size category, which is one of the fastest growing segments in the apparel industry. For confidentiality reasons we are unable to identify the name of the catalog. The items are all sold under the firm’s own private label brand and are only available through the company’s catalog. Although clothing with the same brand is not available in retail stores, other companies offer competing brands in both direct and traditional store channels.

The company offers a very liberal return policy: customers can return any item for any reason provided they pay for return shipping and handling.[1] A pre-paid mailing label allows customers to return the item via the US Postal service with no immediate out-of-pocket expense. After receipt of the item, the company refunds the item price less $4.00 for return shipping.

The pilot study focuses on a single catalog for which three catalog versions were produced. Each version was distributed to a random selection of 90,000 customers. The study was designed to investigate how varying the price, the price ending and the use of “Sale” cues impacted demand. The findings, which are reported in a previous paper (Anderson and Simester 2003), confirm that these were effective at increasing demand. A detailed description of the experimental design and summary statistics is provided in (Anderson and Simester 2003).

Our current analysis focuses on the price manipulations and its impact, if any, on product returns. For this reason, we focus on the 65 items in the test, which are sold at three price conditions across the three catalog versions and have positive sales at each price conditions. This allows us to calculate a “return rate” in each condition. We then compare the impact of prices on return rates by comparing between the three experimental conditions for all these 65 items. To ease our concern about the selection bias of these 65 items, we compare them with the other items in the test and find no significant difference in terms of average price ($58.22 vs. $56.50), total units sold (39.62 vs. 40.31) and return rate (24.66% vs. 23.98%). The random allocation of customers to three pricing conditions also overcomes other potential confounds. For example, we can rule out intervening events, such as competitive actions, because these

[1] Theoretical models have identified conditions under which such polices are optimal (see for example Hess et al. 1996; and Davis et al. 1998).

events are common to the three experimental conditions. Moreover, by exogenously varying prices between the three catalog versions, we overcome endogeneity concerns that potentially arise when using non-experimental data.

To evaluate how the experimental manipulation of prices across the three conditions affected the return rate we begin by presenting univariate findings. For ease of exposition we label the three price levels: “Low” Medium” and “High”. Aggregating across the 65 items yields measures of the return rate at each price level. The findings reveal a significant increase ($p < 0.05$) in the return rate in the High (28.2%) and Medium (28.5%) price conditions compared to the Low price condition (24.3%), but no significant difference between the Medium and High conditions.

In regression analysis we focus on the same 65 items and estimate a model with product fixed effects and price as independent variables. Note that the product fixed effects control for all item characteristics such as color, size and style. The dependent variable, $Y_{i,v}$, is either the *Number of Returns*$_{i,v}$ or the *Return Rate*$_{i,v}$ for item *j* in catalog version *v*. The findings are reported in Table 3.1.

Table 3.1: Within-Item Variation in Return

	Number of Return	Return Rate
Price	0.6137* (0.2092)	0.0130* (0.0061)
Adj. R-squared	0.704	0.289
Sample size	195	195

Standard errors are in parentheses. We omit coefficients for the binary variables identifying each item.
* Significantly different from zero ($p < 0.05$).

The results strongly reject the straw-man that return rates are independent of price paid. Consistent with the univariate results, we show that return rates are substantially larger when a product is sold at a higher price. We also show that the number of products returned is increased at a higher price. This latter result is surprising since higher prices yield fewer sales, resulting in an inverse relationship between sales and the number of returns.

Overall, the results strongly reject the straw-main hypothesis that return rates are constant. Clearly, the price paid has an impact on both return rates and the number of returns. The exogenous price experiment allows us to make such a causal statement. In order to further explore the underlying driving forces of the price effect on returns, we develop a theoretical model to generate hypotheses and then test them empirically.

3.4 A Model of Customer Return Behavior

In this section, we develop a theoretical model describing consumers' purchase and return decision. The model generates testable hypotheses about how prices affect returns.

Consider a consumer with utility $U = v - p$, where v is the valuation of the item and p is the price. Prior to purchasing an item, consumer h is uncertain about the item's valuation and has a prior cumulative distribution $F_h(V)$. For example, a consumer purchasing from a catalog may read an item description and see a photograph of an item prior to purchasing. After the item is received and inspected the true value, v, is revealed. At that point, the customer decides whether to keep or return the item. Due to variation in fit, styling, color and other item characteristics the true valuation may differ from the customer's expectations.

In contemplating the return decision, the customer considers the value of the outside option, $\tilde{U}$, and the return costs, c. The outside option represents the expected surplus when purchasing from a competing store. For ease of exposition we scale $\tilde{U} - c$ to zero ($\tilde{U} - c = 0$). Given these assumptions, a customer will return an item valued at v and purchase at price p iff:

$$v - p < \tilde{U} - c \equiv 0 \tag{3.1}$$

Therefore, the probability of returning the product is $F_h(p)$, which is positively correlated with price. We label this effect of price on return as the *perceived value* effect. It is the direct effect of price on each individual consumer's return decision.

When making a purchase decision, customers are forward looking and incorporate the return option into their purchasing decision. Let $\bar{V}_h = E_h\left(V \mid V \geq p\right)$ represents the expected value of an item that is not returned by customer h. Customer h will purchase an item if and only if:

$$\left[1 - F_h\left(p\right)\right]\left(\bar{V}_h - p\right) + F_h\left(p\right)\left(\tilde{U} - c\right) = \left[1 - F_h\left(p\right)\right]\left(\bar{V}_h - p\right) \geq \tilde{U}\ . \tag{3.2}$$

As we would expect, inequality (3.2) implies that customer demand for an item is negatively correlated with price paid. If customers with different return probability, $F_h(p)$, are attracted to purchase, the aggregate level returns would change with prices. We label this effect of price on return as the *incremental customer* effect. It is the effect of price on customer mix and indirectly on aggregate returns.

We can illustrate the perceived value effect and the incremental customer effect of price on returns using the following example. Consider a market with two segments of customers: a mass of n_H high type customers and a mass of n_L low type customers. There are two exogenous price levels, p_H and p_L, such that only the high type customers purchase at p_H and both types of customers purchase at p_L. Each customer segment has the same prior distribution of valuations, $F_h(V)$ where $h \in (H, L)$, and receives an independent draw from this distribution. If upon arrival of the item, inspection reveals that $v < p$ then the customer returns the item. Otherwise the customer keeps the item.

At the high price the return rate is $r(p_H) = F_H(p_H)$ and the total number of items returned is: $R(p_H) = n_H F_H(p_H)$. At the low price, the number of items returned is $R(p_L) = n_H F_H(p_L) + n_L F_L(p_L)$ and the return rate is $r(p_L) = R(p_L)/(n_H + n_L)$. If the price decreases from p_H to p_L then the change in total returns is:

$$\begin{aligned}\Delta R &= R(p_L) - R(p_H) \\ &= n_L F_H(p_H) + n_H\left[F_H(p_L) - F_H(p_H)\right] + n_L[F_L(p_L) - F_H(p_H)]\end{aligned} \tag{3.3}$$

In equation (3.3), the change in total return is decomposed into three terms. The first term is the expected change in the total return if the low type customers return at the same rate as the high type. This is the prediction by the straw-man model. The second term, which captures the perceived value effect, is the change in total returns due to the change in return probability of the high type customers as the price drops from p_H to p_L. The third term, which captures the incremental customer effect, is the change in returns due to the difference between the return probability of the incremental (the low type) customers at the new price p_L and that of the high type customers at the original price p_H.

Similar to the decomposition in equation (3.3), the change in return rate can be decomposed into two terms:

$$\begin{aligned}\Delta r &= r\left(p_L\right) - r\left(p_H\right) \\ &= \frac{n_H F_H(p_L) + n_L F_L(p_L)}{n_H + n_L} - F_H(p_H) \\ &= \frac{n_H\left[F_H(p_L) - F_H(p_H)\right]}{n_H + n_L} + \frac{n_L(F_L(p_L) - F_H(p_H))}{n_H + n_L}\end{aligned} \tag{3.4}$$

The first term captures the change in return rate due to the perceived value effect and the second term captures the incremental customer effect.

While the straw-man model predicts that the number of returns increases and the return rate remains the same as the price drops, we show that this may not be true. The perceived value effect implies that the total returns and the return rate decrease as price drops, since $F_H(p_L) < F_H(p_H)$. The incremental customer effect predicts that the change in total return and return rate depends on the comparison between the low type customers' return rate and the high type customers' return rate. If the incremental customers (low type) are bargain hunters and are less likely to return a discounted item than the high type customers, the incremental customer effect makes returns decrease with prices. On the other hand, if the incremental customers are poor fits for the product and are more likely to return, the incremental customer effect makes returns increase with prices. Therefore, the overall effect of price on return is ambiguous depending on both the perceived value and the incremental customer effect.

Given the ambiguity of the theoretical predictions, the actual outcome is an empirical question. In the next section, we present an empirical study to estimate how price affects returns.

3.5 Empirical Study

The theoretical model suggests that price affects returns in two ways. It influences individual consumer’s return decisions through the perceived value effect and the aggregate consumer mix through the incremental customer effect. In this section, we empirically test the existence of these two effects and measure their sizes. A better understanding of these effects work will contribute to retailers’ understanding of how to design more profitable promotions.

3.5.1 Data description

The data used in this study is from a second multi-product retailer that is different than the retailer involved in the pilot study. The products are sold under this (second) retailer’s store brand name and through mailing catalogs, internet and physical stores. In this study we focus on customers who live in areas without physical stores and therefore have to make purchases remotely (via the Internet, mail or phone). Such remote purchases increase quality uncertainty and therefore product returns. The sample used in this study is a random sample of 3000 active consumers who have made at least one purchase between 2000 and 2003 and at least one purchase between 2004 and 2006. We use the purchase and return transactions from the years 2000 through 2003 for estimation and transactions from 2004 to 2006 for holdout prediction. For each customer, we also have demographic information, including the number of children in the household, and the age, income, gender and marital status of the head of household.

The products under consideration are clothing. Due to the large number of SKUs, it is infeasible to estimate a model for each SKU. To keep our analysis tractable and parsimonious, we merge all products into seven categories: women's casual, women's outerwear, women's dress, men's casual, men's outerwear, kid's casual and kid's outerwear. Such categorization allows us to investigate how the price effect on returns varies by product type. The prices, units sold and return rates of these seven categories are summarized in Table 3.2. On average, casuals have lower regular prices, they are sold at lower prices, more units are sold, and return rates are lower than for other products. Women's dresses have the deepest discounts and the highest return rates among all categories.

Table 3.2: Average Price, Units Sold and Return Rate by Products

Categories	*Average Full Price*	*Average Selling Price*	*Average Discount*	*Units Sold*	*Return Rate*
women's casual	40.36	32.67	19.05%	15551	14.44%
women's outerwear	105.03	90.01	14.30%	2708	22.90%
women's dress	55.32	41.73	24.57%	1621	30.54%
men's casual	40.00	34.80	13.00%	10264	7.38%
men's outerwear	118.49	104.21	12.05%	1707	15.47%
kids' casual	21.43	19.47	9.15%	2429	9.06%
kids' outerwear	56.54	49.91	11.73%	1149	14.44%

The return policy by the retailer is similar to the mailing-catalog company in the pilot study. Customers can return any item for any reason using a prepaid UPS or USPS return label and paying $6.50 for return shipping.

3.5.2 A joint model of demand and return

To account for both the perceived value and incremental customer effect, we need a model to jointly estimate individual consumer's purchase and return decision. Since consumer heterogeneity is the basis of the incremental customer effect, we specify the model parameters at individual consumer level. This model not only allows us to better predict consumers' returns under price promotions, but also provides managerial insight into how to design targeted promotions to prevent returns in the first place.

We assume that at each time *t* the consumer *i* makes a purchase / no purchase decision in each category. If a purchase is made the consumer then makes a return / no return decision after receiving the product. This two-stage decision process is typical in the remote purchase environment.

To model the purchase and return decision, we use a binary Probit specification. It is assumed that consumer *i*'s utility to make a purchase in category *k* at time *t* is:

$$b_{ikt} = \beta_{ik} + \beta_{ik}^{p} \overline{P}_{kt} + \overline{\beta} D_t + \varepsilon_{ikt} \qquad (3.5)$$

where β_{ik} captures consumer *i*'s expected utility in category *k*, $\overline{P}_{kt}$ is the average price index of category *k* at time *t*, and D_t are year and month dummies to control for any time trend. The random error ε_{ikt} is assumed to be i.i.d distributed as N(0, π_b^2)[2]. If b_{ikt} >0, the consumer decides to purchase, denoted as $B_{ikt} = 1$. Otherwise, the consumer decides not to purchase, denoted as $B_{ikt} = 0$. The price index is calculated as the ratio of selling price to full price. Using

[2] For identification reasons, π_r^2 is normalized to 1.

the same dataset, Anderson et al. (2006) found that using this “discount” index better fits the data than using a simple price index. Notice that it controls for differences in the base price across categories.

The return or keep decision is made only after a purchase. It is assumed that the utility to make a return in category *k* given a purchase at time *t* is:

$$r_{ikt} = \omega_{ik} + \omega_{ik}^{p} \bar{p}_{ikt} + \omega_{i}^{item} \bar{x}_{ikt} + \varpi_{i}^{order} \bar{o}_{ikt} + \eta_{ikt} \tag{3.6}$$

where ω_{ik} is consumer *i*’s actual utility in category k, $\bar{p}_{it}$ is the average price index paid, $\bar{x}_{it}$ and $\bar{o}_{it}$ are the average of item characteristics and order characteristics conditional on purchases at time *t*. The item characteristics include dummy variables identifying differences in color, pattern, season and ordering characteristics (the shipping cost, use of the Internet channel, and payment through either coupons or gift cards). Notice that product size is a customer characteristic and so it is captured by the intercept ω_{ik}.

The variables used in the return utility specification are conditional on the product and how it is purchased. The random error η_{ijt} is assumed to be i.i.d. normally distributed as N(0, π_r^2)[3]. It is also assumed that after controlling for the deterministic part of the utility η_{ijt} and ε_{ikt} are independent. If r_{ikt} >0, the consumer decides to return, denoted as $R_{ikt} = 1$ (otherwise, $R_{ikt} = 0$). As modelers, we observe consumers’ purchase incidence B_{ikt} and return incidences R_{ikt}.

[3] For identification reasons, π_b^2 is normalized to 1.

The interdependence between a consumer's purchase and return decision is modeled through the correlations between the demand parameters β_i s and the return parameters ϖ_i s. Let θ_i denote for the vector of individual level parameters $(\{\beta_{ik}\},\{\beta_{ik}^p\},\{\omega_{ik}\},\{\omega_{ik}^p\},\omega_i^{item},\varpi_i^{order})'$ and it is assumed to vary based on observable customer demographics Z_i:

$$\theta_i = \Pi Z_i + \upsilon_i$$

where ν_i captures unobservable heterogeneity and it is assumed to be distributed as:

$$\nu_i \sim \text{MVN}\,(\mathbf{0},\ \Omega)$$

Variable Z_i includes customer demographics, such as gender of the head of household, marriage status, number of kids, age, and household income. To estimate the large set of individual level parameters, we use a hierarchical Bayesian approach. In particular, we use the MCMC procedure to simulate the posterior distributions of the model parameters (Allenby and Rossi 1999). The details for the prior specification, joint posterior distribution and the estimation algorithm are described in Appendix 3.1.

The model allows us to estimate both the perceived value effect and the incremental customer effect. The perceived value effect is captured by the price coefficient ϖ_{ik}^p in the return equation (3.6) and is expected to be positive. The incremental customer effect arises when customers who purchase at different prices have different overall return propensity. This effect can be captured by the correlation between the demand price sensitivity β_{ik}^p and the overall return utility r_{ikt}. If more price sensitive consumers are associated with a lower tendency to return, then we expect the aggregate return rate to decrease with price. Otherwise, the return rate will increase with price.

3.5.3 Results

To evaluate the model fit, we calculate the in-sample and out-of-sample hit rates for return incidences in all categories using estimates from the joint demand and return model. We also estimate a benchmark model which only considers the return decisions (similar to Hess and Mayhew 1997). As shown in Table 3.3, our model outperforms the benchmark model in both in-sample and out-of-sample predictions.

Table 3.3: Model Fit Comparison

Categories	***In-Sample Hit Rate***		***Out-of-Sample Hit Rate***	
	Our Demand-Return Model	**Benchmark Return Model**	**Our Demand-Return Model**	**Benchmark Return Model**
Women's Casual	79.79%	61.07%	74.43%	55.44%
Women's Outerwear	77.21%	61.40%	74.62%	58.74%
Women's Dress	65.84%	57.49%	59.59%	53.80%
Men's Casual	90.68%	59.80%	81.10%	56.80%
Men's Outerwear	81.30%	62.10%	77.09%	60.50%
Kids' Casual	86.55%	56.19%	80.69%	52.54%
Kids' Outerwear	72.07%	59.57%	69.15%	57.86%

We then discuss how the perceived value effect and the incremental customer effect are manifested in our model. Recall that the perceived value effect is captured by the price coefficient in the return equation. In the last column of Table 3.4, we report the mean and standard deviation of θ_i across consumers.

Table 3.4: Individual Parameter θ_i and Posterior Mean and Standard Deviation of Π in the Individual Customer Joint Demand and Return Model

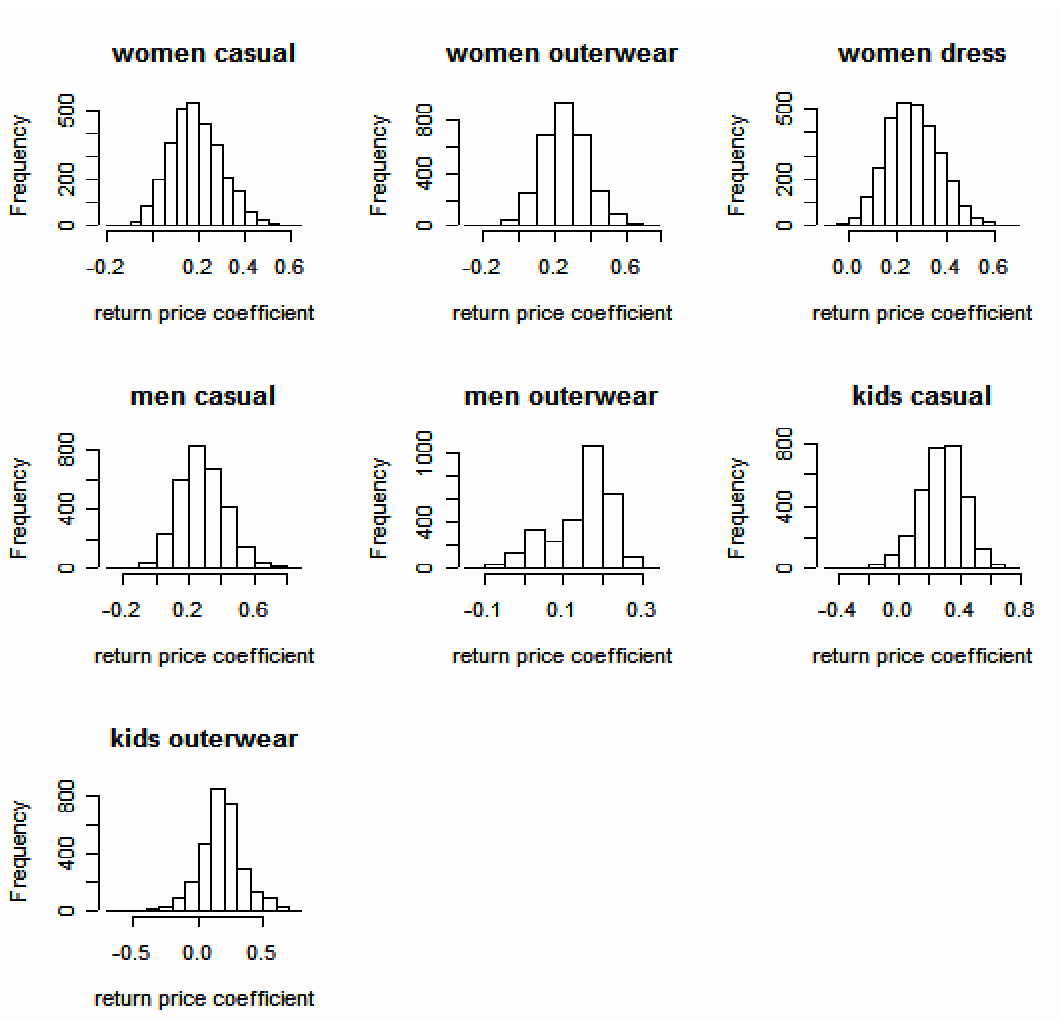
	Variables	Constant	Gender	Married	# kids	Income	Age	θ_i
Demand								
Intercept	women's	-1.62*	0.27*	-0.03	-0.03*	0.11*	-0.08	-1.40
	casual	(.06)	(.03)	(.02)	(.01)	(.01)	(.08)	(.22)
	women's	-1.06*	-0.10*	-0.11*	-0.09*	-0.02	-0.37*	-1.47
	outerwear	(.12)	(.02)	(.02)	(.01)	(.02)	(.15)	(.16)
	women's	-1.68*	-0.01	0.07*	-0.02	0.00	-0.15	-1.72
	dress	(.06)	(02)	(.04)	(.01)	(.02)	(.15)	(.16)
	men's	-1.49*	-0.17*	-0.04*	0.05*	0.06*	0.20*	-1.46
	casual	(.05)	(.03)	(.02)	(.02)	(.01)	(.09)	(.14)
	men's	-1.22*	-0.07*	-0.02	-0.06*	0.07*	-0.29*	-1.42
	outerwear	(.07)	(.03)	(.02)	(.01)	(.02)	(.07)	(.10)
	kids' casual	-1.36*	0.16*	-0.03	0.08*	-0.10*	-0.45*	-1.56
		(.06)	(.01)	(.03)	(.02)	(.02)	(.09)	(.18)
	kids'	-1.29*	0.06*	0.11*	0.07*	0.03	-0.53*	-1.40
	outerwear	(.11)	(.03)	(.03)	(.01)	(.05)	(.09)	(.20)
Price	women's	-1.32*	0.02	0.05*	0.00	-0.13*	-0.06	-1.41
Index	casual	(.06)	(.02)	(.01)	(.01)	(.01)	(.08)	(.14)
	women's	-2.38*	0.32*	0.11*	0.07*	0.04*	0.65*	-1.59
	outerwear	(.05)	(.03)	(.02)	(.02)	(.02)	(.10)	(.21)
	women's	-1.57*	0.31*	-0.01	-0.02	0.00	-0.12	-1.41
	dress	(.12)	(.03)	(.02)	(.02)	(.03)	(.16)	(.24)
	men's	-1.29*	-0.09*	0.08*	-0.08*	-0.04*	0.04	-1.36
	casual	(.06)	(.02)	(.02)	(.02)	(.01)	(.10)	(.14)
	men's	-1.63*	-0.15*	0.07*	0.04*	-0.05*	0.44*	-1.44
	outerwear	(.09)	(.02)	(.02)	(.01)	(.02)	(.13)	(.13)
	kids' casual	-1.76*	-0.06*	0.11*	0.12*	0.09*	0.02	-1.54
		(.10)	(.03)	(.03)	(.02)	(.03)	(.18)	(.24)
	kids'	-1.62*	0.12*	-0.05*	0.15*	0.02	0.07	-1.42
	outerwear	(0.09)	(0.01)	(0.02)	(0.01)	(0.03)	(0.14)	(.16)
Return								
Intercept	women's	-0.35*	0.18*	0.00	-0.02	0.04	-0.88*	-0.7
	casual	(.07)	(.03)	(.04)	(.02)	(.02)	(.14)	(.18)
	women's	-0.58*	0.04	-0.15*	-0.01	0.13*	-0.01	-0.56
	outerwear	(.06)	(.03)	(.02)	(.02)	(.02)	(.08)	(.12)
	women's	0.54*	-0.15*	0.06*	-0.15*	-0.09*	-0.65*	-0.05
	dress	(.10)	(.03)	(.02)	(.03)	(.02)	(.17)	(.16)
	men's	-0.77*	-0.01	-0.03	0.03	-0.03	-0.42*	-1.06
	casual	(.11)	(.03)	(.02)	(.02)	(.02)	(.13)	(.08)
	men's	-0.31*	0.12*	0.06	0.08*	0.01	-0.31*	-0.31
	outerwear	(.08)	(.04)	(.03)	(.02)	(.03)	(.14)	(.11)
	kids' casual	-0.67*	0.30*	-0.03	-0.05*	0.09*	-0.20*	-0.52
		(.09)	(.02)	(.02)	(.02)	(.04)	(.09)	(.15)
	kids'	0.04	-0.06*	-0.19*	-0.02	0.19*	-0.60*	-0.34
	outerwear	(.16)	(.03)	(.03)	(.03)	(.02)	(.21)	(.15)
Price	women's	0.00	-0.02	-0.04	0.05*	-0.01	0.33*	.18
Index	casual	(.11)	(.03)	(.04)	(.01)	(.02)	(.11)	(.11)
	women's	-0.15*	0.08	0.04	0.06*	0.02	0.45*	.25
	outerwear	(.06)	(.06)	(.03)	(.01)	(.02)	(.09)	(0.13)
	women's	0.35*	-0.16*	0.03	-0.01	-0.06*	0.08	.26
	dress	(.07)	(.04)	(.03)	(.02)	(.02)	(.12)	(0.11)

Table 3.4 (Continued)

	Variables	*Constant*	*Gender*	*Married*	*# kids*	*Income*	*Age*	θ_i
Price	men's	-0.09	-0.12*	-0.10*	0.10*	0.00	0.77*	.28
Index	casual	(.16)	(.03)	(.05)	(.01)	(.02)	(.19)	(0.15)
	men's	0.20*	0.16*	-0.05	-0.01	0.04	-0.29*	.15
	outerwear	(.05)	(.02)	(.03)	(.03)	(.03)	(.08)	(0.08)
	kids'	0.79*	-0.15*	-0.05	-0.09*	-0.10*	-0.40*	.28
	casual	(.09)	(.03)	(.06)	(.02)	(.03)	(.09)	(0.15)
	kids'	0.15	-0.01	-0.11*	-0.16*	0.20*	0.02	.18
	outerwear	(.10)	(.05)	(.02)	(.03)	(.03)	(.11)	(0.17)
Order	internet	-0.10	0.13*	-0.01	-0.07*	0.02	-0.46*	-.32
Characte		(.09)	(.03)	(.03)	(.02)	(.03)	(.13)	(0.14)
ristics	shipping	0.11	0.08	-0.17*	-0.02	-0.30*	-0.02	-.23
	cost	(.09)	(.07)	(.04)	(.01)	(.05)	(.11)	(0.22)
	gift card /	0.14	-0.33*	0.02	-0.12	-0.01	-0.22	-.30
	coupon payment	(.11)	(.02)	(.02)	(.01)	(.01)	(.15)	(0.18)

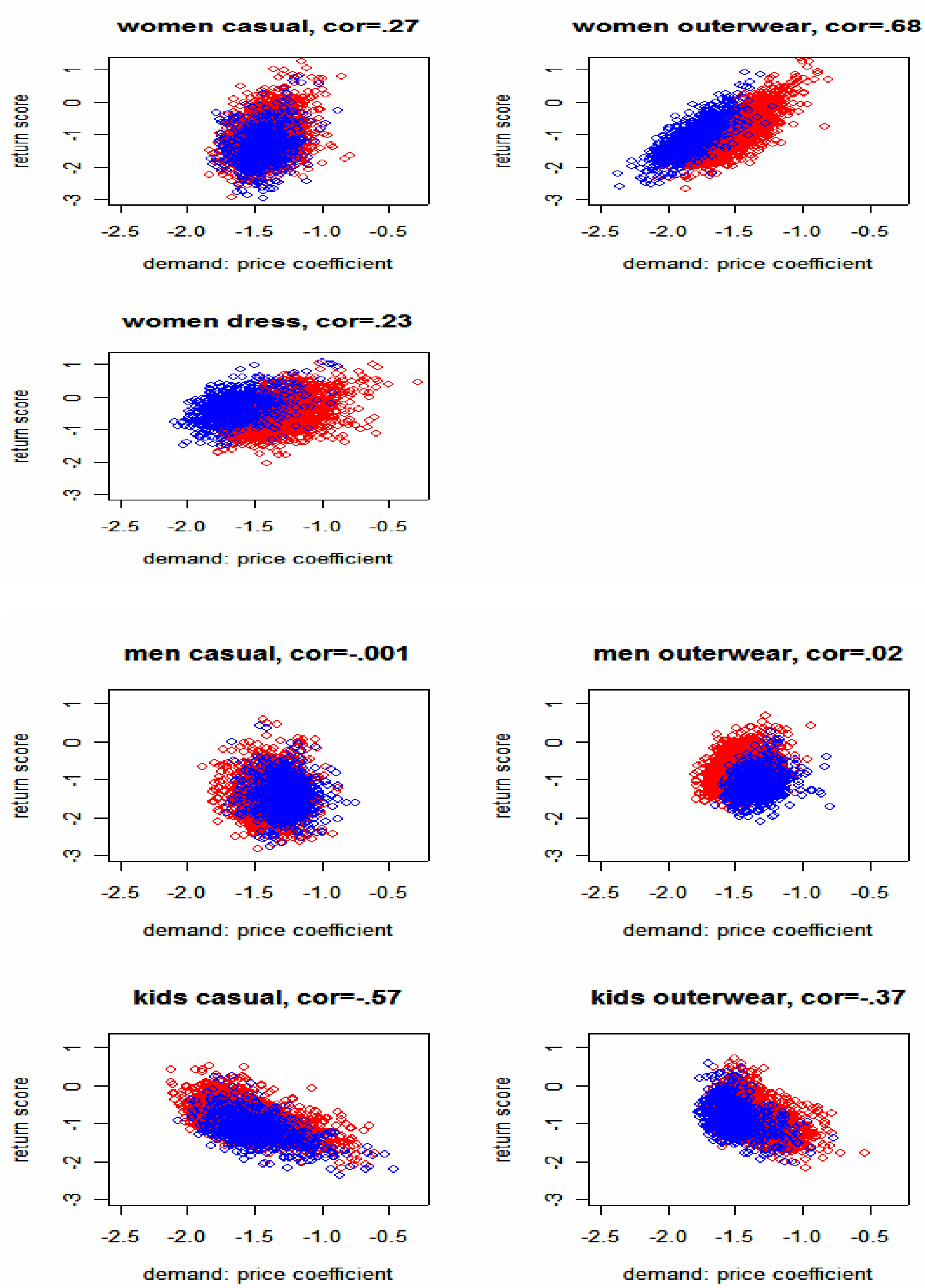
The perceived value effect is captured by the price sensitivity in the return equation. In Figure 3.1, we show the histograms of the return price sensitivity by categories. The perceived value effect is positive as expected for over 90% of the customers in all categories. Interestingly, consumers' return price sensitivities are found to be much smaller than their demand price sensitivities across all categories. This effect can be explained by the endowment effect (Kahneman et. al. 1990) in the behavioral literature. Consumers' feeling of owning the products after purchasing may increase their willingness to pay for the products. In a similar remote purchase context, Wood (2001) also finds the endowment effect using lab experiments.

Figure 3.1: Histograms of the Perceived Value Effect across Categories



We then discuss the incremental customer effect. While the endowment effect may reduce consumers' propensity to return after purchase, it is not sufficient to induce the incremental customer effect if the endowment effect affects the return propensity of consumers who purchase at different price levels in the same way. In addition to price, there could be other factors influencing the return decision. A more comprehensive indicator of the incremental customer effect is the correlation between consumers' demand price sensitivity β_{ik}^{p} and the average return propensity $\bar{r}_{ik}$. To calculate consumer i's $\bar{r}_{ik}$, we first calculate r_{ikt} at each time and then take the average over time. For the illustration purpose, we calculate r_{ikt} for a representative item and fix the item characteristics and order characteristics at the mode (a blue, solid color, all-year-round item, order by phone/mail, shipping cost $4.45, and pay 6% using gift card or coupon). In Figure 3.2, we show the scatter plots of β_{ik}^{p} and $\bar{r}_{ik}$ by categories and customer demographic variable: gender. Interestingly, the correlation varies systematically across product categories. The correlations in women's categories are significantly positive (0.27 for women's casual, 0.68 for women's outerwear, 0.23 for women's dress), in men's categories are not significant (-0.001 for men's casual and 0.02 for men's outerwear) and in kids' categories are significantly negative (-0.57 for kids' casual and -0.37 for kids' outerwear). This confirms our hypotheses that there exists the incremental customer effect and the sign of the effect is ambiguous. Customers who purchase discounted women's products turn out to have lower return propensity, whereas who purchase discounted kids' products turn out to have higher return propensity.

Figure 3.2: Scatter Plots of β_{ik}^{p} vs. r_{ikt} and Their Correlations by Categories



The systematic variation in the incremental customer effect across categories may be attributed to product or customer or characteristics. We explore these two potential sources respectively. First, we examine customer characteristics. Recall that in our model we include several consumer demographics to control for the observable heterogeneity. The posterior mean and standard deviation of parameter Π are shown in Table 3.4. Although the demographic variables are significant in the demand and return equations, we do not observe a systematic variation in these coefficients across categories. In particular, as shown in Figure 3.2, the correlations are similar for female and male customers across categories. Secondly, we examine how product characteristics drive the incremental customer effect. Since $r_{ikt} = \omega_{ik} + \omega_{ik}^{p}\bar{p}_{ikt} + \omega_{i}^{item}\bar{x}_{ikt} + \varpi_{i}^{order}\bar{o}_{ikt} + \eta_{ikt}$, we can decompose the correlation into terms involving product characteristics:

$$cor(\beta_{ik}^{p},\bar{r}_{ik}) = \frac{cor(\beta_{ik}^{p},\omega_{ik}) + cor(\beta_{ik}^{p},\omega_{ik}^{p}\bar{\bar{p}}_{ik}) + cor(\beta_{ik}^{p},\omega_{i}^{item}\bar{\bar{x}}_{ik}) + cor(\beta_{ik}^{p},\varpi_{i}^{order}\bar{\bar{o}}_{ik})}{std(\beta_{ik}^{p})std(\bar{r}_{ik})}.$$

The decomposed correlations are shown in Table 3.5. No single product characteristics or order characteristics solely drives the systematic incremental customer effect across categories. Price, color, and pattern seem to contribute the most to the correlation between the demand price sensitivity β_{ik}^{p} and the overall return propensity $\bar{r}_{ik}$. These results suggest that the incremental customer effect arises mainly because consumers who purchase at different price levels hold different preferences for these product characteristics when returning the products.

Table 3.5: Decomposing Correlations between β_{ik}^{p} and $\overline{r}_{ik}$

Correlations	β_{ik}^{p}						
	women's casual	**women's outerwear**	**women's dress**	**Men's casual**	**Men's outerwear**	**Kids' casual**	**Kids' Outerwear**
Intercept	0.030	0.083	-0.244	-0.043	-0.084	-0.064	-0.013
Price	0.014	0.134	-0.031	-0.044	-0.120	-0.281	-0.183
Internet Dummy	-0.002	0.000	-0.001	-0.002	-0.001	-0.002	-0.001
Shipping Cost	0.005	-0.004	-0.002	0.003	-0.006	-0.018	-0.008
Color Dummy 1	-0.003	0.000	0.000	-0.001	0.000	-0.001	-0.001
Color Dummy 2	-0.001	0.000	0.000	-0.001	0.000	0.000	0.000
Color Dummy 3	0.163	0.216	0.215	0.097	0.223	-0.099	-0.122
Color Dummy 4	-0.003	0.000	0.000	-0.001	0.000	-0.001	0.000
Color Dummy 5	-0.001	-0.001	-0.001	-0.001	-0.001	-0.001	-0.001
Color Pattern Dummy	0.083	0.278	0.318	-0.018	0.005	-0.090	-0.020
Season Dummy 1	-0.002	0.000	-0.001	-0.001	0.000	-0.002	0.000
Season Dummy 2	-0.003	0.000	-0.001	-0.002	0.000	-0.003	-0.002
Coupon Payment	-0.007	-0.015	-0.018	0.012	0.002	-0.007	-0.016
Total Correlation	**0.272**	**0.689**	**0.234**	**-0.001**	**0.018**	**-0.569**	**-0.367**

In summary, we show that price influence returns in two ways: the perceived value effect and the incremental customer effect. The perceived value effect is found to be positive as expected, i.e. consumers are less likely to return discounted products. The incremental customer effect is found to vary across categories. Consumers who purchase at lower prices in women's categories have lower return propensities, in men's categories similar return propensities and in kids' categories higher return propensities. Therefore, the overall effect of price on returns is ambiguous depending on the perceived value effect and the incremental customer effect.

3.5.4 Simulation study and managerial implications

In this section, we conduct a simulation study to measure the impact of the perceived value effect and the incremental effect on returns.

Based on the model estimated earlier, we simulate demand and return for each product when its price is discounted by 30% (a typical discount in the dataset). In Table 3.6, we show the simulated aggregate demand and return by price paid for each category (rescaled to annual units sold and returned). While the decreased price does not change the demand from the full-price buyers, it reduces their returns in all categories. This is due to the perceived value effect. In addition, the decreased price also attracts incremental customers. The simulation shows that the incremental customers have lower return rates (since β_{ik}^{p} and r_{ikt} are positively correlated) than the full-price buyers in women's categories, similar return rates in men's categories and higher return rates in kids' categories.

Table 3.6: Simulated Demand, Return, and Return Rate at Full Price and at 70% Price

	Full Price			70% Price					
	Existing Customers			Existing Customers			Incremental Customers		
	Demand D_0	Return R_0	Return Rate $r_0 = \frac{R_0}{D_0}$	Demand D'_0	Return R'_0	Return Rate $r'_0 = \frac{R'_0}{D_0}$	Demand D_1	Return R_1	Return Rate $r_1 = \frac{R_1}{D_1}$
women's casual	974	184.75	18.97%	974	169.5	17.40%	1028.5	159.5	15.51%
women's outerwear	299	72	24.08%	299	65	21.74%	500	92	18.40%
dress	160.5	60.75	37.85%	160.5	56.25	35.05%	184.75	59.25	32.07%
men's casual	664.75	58.25	8.76%	664.75	49.25	7.41%	774.75	58	7.49%
men's outerwear	198.25	40	20.18%	198.25	38	19.17%	333.5	64	19.19%
kids' casual	129	17.5	13.57%	129	16.5	12.79%	197.25	29.5	14.96%
kids' outerwear	105	16.75	15.95%	105	16	15.24%	161.5	27	16.72%

As in Section 3.4, the change in total return and return rate can be decomposed into the perceived value effect and the incremental customer effect. In Table 3.7, we show the decomposition of the change in total number of return. If the return rate is constant, we should expect the simulated change in total returns (column 6) to be close to the straw-man model prediction (column 7). However, we observe that the straw-man model overestimates in some categories, but underestimates in the others: it overestimates by 35% for women's casual, 42% for women's outerwear, 28% for women's dress, 39% for men's casual, 9% for men's outerwear, and underestimates by 6% for kids' casual and 2% for kids' outerwear. These discrepancies can be further decomposed into the perceived value effect and the incremental customer effect, where the perceived value effect is calculated as the change in the number of returns of the full-price buyers and the incremental customer effect is calculated as the incremental demand multiplied by the difference in the return rate of incremental customer and the full-price customer. The last two columns in Table 3.7 show that the incremental customer effect has a larger effect than the perceived value effect across categories.

Table 3.7: Decomposing the Change in the Number of Returns

	Total demand (full)	Total return (full)	Total Demand (70%)	Total Return (70%)	Change in Return $R'_0+R_1-R_0$	Straw man $D_1\cdot r_0$	Perceived Value	Incremental Customer $D_1(r_1-r_0)$
	D_0	R_0	D_0+D_1	R'_0+R_1			R'_0-R_0	
women's casual	974.00	184.75	2002.50	329.00	144.25	195.09	-15.25	-35.59
women's outerwear	299.00	72.00	799.00	157.00	85.00	120.40	-7.00	-28.40
dress	160.50	60.75	345.25	115.50	54.75	69.93	-4.50	-10.68
men's casual	664.75	58.25	1439.50	107.25	49.00	67.89	-9.00	-9.89
men's outerwear	198.25	40.00	531.75	102.00	62.00	67.29	-2.00	-3.29
kids' casual	129.00	17.50	326.25	46.00	28.50	26.76	-1.00	2.74
kids' outerwear	105.00	16.75	266.50	43.00	26.25	25.76	-0.75	1.24

We can also decompose the change in return rates into the perceived value and incremental customer effect, as shown in Table 3.8. This decomposition confirms that the change in return rates is significant, especially in women's categories. Moreover, it confirms that the incremental customer effect is the driving force across all categories and it even offsets the perceived value effect and increase the return rate slightly in the kids' categories.

Table 3.8: Decomposing the Change in the Return Rates

	return rate (full) $\frac{R_0}{D_0}$	return rate (70% price) $\frac{R'_0+R_1}{D_0+D_1}$	% Change in Return Rate $\frac{\frac{R'_0+R_1}{D_0+D_1}-\frac{R_0}{D_0}}{\frac{R_0}{D_0}}$	% Due to Perceived Value $\frac{\frac{R'_0-R_0}{D_0+D_1}}{\frac{R_0}{D_0}}$	% Due to Incremental Customer $\frac{\frac{R_1-R_0\frac{D_1}{D_0}}{D_0+D_1}}{\frac{R_0}{D_0}}$
women's casual	18.97%	16.43%	-13.38%	-4.01%	-9.37%
women's outerwear	24.08%	19.65%	-18.40%	-3.64%	-14.76%
dress	37.85%	33.45%	-11.62%	-3.44%	-8.17%
men's casual	8.76%	7.45%	-14.97%	-7.13%	-7.84%
men's outerwear	20.18%	19.18%	-4.93%	-1.86%	-3.07%
kids' casual	13.57%	14.10%	3.93%	-2.26%	6.19%
kids' outerwear	15.95%	16.14%	1.15%	-1.76%	2.91%

The simulation results confirm the importance of coordinating marketing promotions and managing product returns. To maximize profits, a retailer needs not only to predict product returns, but also prevent returns in the first place. Our findings suggest that promoting the right products to the right customers can help prevent returns. Specifically, the result of the incremental customer effect tells us that the retailer may promote women's and men's products to price sensitive customers without worrying about inflating returns, but promoting kids' products to price sensitive consumers may inflate returns. The individual level estimates of our model suggest that retailers may be able to mitigate returns by targeting promotions to specific customers.

3.6 Conclusion

It is a widely accepted assumption in the inventory management literature that a product's return rate is invariant to its selling price. In this paper, we theoretically and empirically investigate this assumption. The findings reject this constant return rate assumption and suggest a mechanism that improves our understanding of the relationship between prices, sales and returns.

While our model explores the effect of price on returns, our model framework can be applied to study the effect of other policies on returns, such return policy and shipping costs. As customizing marketing offerings to consumers become more and more convenient for catalog retailers, our model provides a tool to design more effective customized offerings based on the understanding of customer and product heterogeneity and evaluate the impact of these offerings on product returns.

Chapter 4

Conclusions and Future Directions

This dissertation addresses two issues in retailing: managing brand equity and managing product returns. The primary objective is to develop methodology for utilizing the readily available customer transaction database to understand consumer behavior and evaluate marketing strategies. Our analyses also help retailers to understand the different effects across customers and products in order to design effective customized strategies.

In the first essay, we study consumers' learning behavior and its implication on brand equity formation. While there is an extensive body of literature on brand equity, its formation process has not been fully studied. To our best knowledge, there are no empirical studies to quantify this formation process. To bridge this gap, we build an individual customer level learning model to describe the brand equity formation process. Our model allows a retailer to track each customer's brand equity over time. Furthermore, it measures the impact of each product category in influencing the brand equity. The ability to track individual customer's brand equity and to identify the key category in the brand equity evolution provides useful guidelines for retailers to manage customers and products.

In the second essay, we examine the effect of price on consumers' product return behavior. A widely accepted assumption in the operation literature is that returns occur as a fixed proportion of demand. However, no empirical studies have tested this assumption. In two empirical studies we show that this assumption is not valid in some situations and that the effect of price on return rates is ambiguous depending on two effects: the perceived value effect and the incremental customer effect. We also show how these effects vary by categories and customers.

In terms of future research, this dissertation is only a first step to understand consumer behavior in today's retail industry. As retailers are offering more products in more channels, consumers' behavior may change accordingly. While this dissertation focus on the remote channel (mail, phone and internet), there is a great need to understand consumers' behavior across the multiple channels: remote channel and store channel. The multiple channels also pose big challenges for retailers to coordinate their marketing and operation strategies across these channels. A superior knowledge of how to design channel specific strategies and how to optimally integrate these channels will give retailers a significant competitive advantages in today's fast-moving retail industry.

References

Ackerberg, Daniel A. (2003), "Advertising, Learning, and Consumer Choice in Experience Good Markets: An Empirical Examination", *International Economic Review,* Vol. 44, No. 3 (Aug), 1007-1040

Ailawadi, Kusum L. and Donald R. Lehmann, and Scott A. Neslin (2003), "Revenue Premium as an Outcome Measure of Brand Equity", *Journal of Marketing*, 67(Oct), 1-17

Allenby, G. M. and P.E. Rossi (1999), "Marketing Models of Consumer Heterogeneity", *Journal of Econometrics,* 89, 57-78

Anderson, Eric and Duncan Simester (2003), "Effects of $9 Price Endings on Retail Sales: Evidence from Field Experiments," *Quantitative Marketing and Economics,* 1 (1), 93-110

Anderson, Eric, Karsten Hansen and Duncan Simester (2006), "The Option Value of Returns", working paper

Anderson, John R. (1983), "The Architecture of Cognition", Cambridge, MA, *Harvard University Press*

Bettman, James R. and C. W. Park (1980), "Effects of Prior Knowledge and Experience and Phase of Choice Process on Consumer Decision Processes: A Protocol analysis", *Journal o f Consumer Research,* Vol. 7, 234-248

Bettman, James R. and Mita Sujan (1987), "Effects of Framing on Evaluation of Comparable and Noncomparable Alternatives by Expert and Novice Consumers", *Journal of Consumer Research,* Vol. 14 (Sep), 141-153

Che, Yeon-Koo (1996), "Customer Return Policies for Experience Goods", *Journal of Industrial Economics,* XLIV (1), March, 17-24

Chib, Siddhartha (1995), "Marginal Likelihood from the Gibbs Output", *Journal of the American Statistical Association,* 90(432), 1313-1321

Ching, Andrew (2002), "Consumer Learning and Heterogeneity: Dynamics of Demand for Prescription Drugs after Patent Expiration", University of Toronto

Coscelli, Andrea and Matthew Shum (2004), "An Empirical Model of Learning and Patient Spillovers in New Drug Entry", *Journal of Econometric,* 122, 213-246

Crawford, Gregory S. and Matthew Shum (2005), "Uncertainty and Learning in Pharmaceutical Demand," *Econometrica*, Vol. 73 (Jul), 1137

Davis, S., E. Gerstner, M. Hagerty (1995), "Money Back Guarantees in Retailing: Matching Products to Consumer Tastes", *Journal of Retailing,* 71(1) 7-22

Davis, S., M. Hagerty, and E. Gerstner (1998), "Return Policies and the Optimal Level of 'Hassle'," *Journal of Economics and Business*, 50, 445-60

de Brito, M.P. and E.A. van der Laan (2002), "Inventory Management with Product Returns: The Impact of (Mis)information," Economic Institute Report EI 2002-29

DeGroot, M., "Optimal Statistical Decisions", New York: McGraw Hill, 1970

Eliashberg, J. and R. Steinberg (1987), "Marketing-Production Decisions in An Industrial Channel of Distribution", *Management Science*, 33(8), 981-1000

Erdem, Tulin and Micheal P. Keane (1996), "Decision Making Under Uncertainty: Capturing Dynamic Brand Choice Processes in Turbulent Consumer Goods Markets", *Marketing Science*, 15, 1-20

Erdem, Tulin (1998), "An Empirical Analysis of Umbrella Branding", *Journal of Marketing Research,* Vol. XXXV (Aug), 339-351

Farquhar, Peter H. (1989), "Managing Brand Equity", *Marketing Research,* 1(Sep), 24-33

Fleischmann, M., R. Kuik, and R. Dekker (2002), "Controlling Inventories with Stochastic Item Returns: A Basic Model," *European Journal of Operational Research*, 138, 63-75

Green, Paul E. and V. Srinivasan (1978), "Conjoint Analysis in Consumer Research: Issues and Outlook", *Journal of Consumer Research,* 5 (Sep), 103-123

Green, Paul E. and V. Srinivasan (1990), "Conjoint Analysis in Marketing", *Journal of Consumer Research*, 5 (Sep), 103-123

Guide, Jr., V.D.R. (2000), "Production Planning and Control for Remanufacturing: Industry Practice and Research Needs," *Journal of Operations Research*, 18 (4), 467-83

Hess, James D. and Glenn E. Mayhew (1997), "Modeling Merchandise Returns in Direct Marketing," *Journal of Direct Marketing*, 11 (2), 20-35.

Hess, James, Wujin Chu, and Eitan Gerstner (1996), "Controlling Product Returns in Direct Marketing," *Marketing Letters,* 7 (Oct), 307-17

Hess, James D. Marilyn T. Lucas (2004), “Doing the Right Thing or Doing the Thing Right: Allocating Resources between Marketing Research and Manufacturing”, *Management Science,* 50 (4), 521-526

Ho, Teck H. And Christopher S. Tang (2004), “Introduction to the Special Issue On Marketing and Operations Management Interfaces and Coordination,” *Management Science*, 50(4), 429-430

John, Deborah Roedder, Brarbara Loken and Christopher Joiner (1998), “The Negative Impact of Extensions: Can Flagship Products Be Diluted?”, *Journal of Marketing,* Vol 62 (Jan), 19-32

Kamakura, Wagner A., Sridhar Ramaswami and Rajenda K. Srivastava (1991), “Qualification of Prospects for Cross-Selling in the Financial Services Industry”, *International Journal of Research in Marketing*, 329-49

Kardes, Frank R., Steven S. Posavac, and Maria L. Cronley (2004), “Consumer Inference: A Review of Process, Bases, and Judgment Contexts”, *Journal of Consumer Psychology,* 14(3), 230-256

Karmarkar, Uday S. (1996), “Integrative Research in Marketing and Operations Management”, *Journal of Marketing Research,* 33(2), 125-133

Keller, Kevin Lane (1993), “Conceptualizing, Measuring, and Managing Customer-Based Brand Equity”, *Journal of Marketing,* Vol. 57 (1), 1-22

Kiesmuller, Gudrun and Erwin van der Laan (2001), “An Inventory Model with Dependent Product Demands and Returns”, *International Journal of Production Economics,* 72(1), 73-87

Kim, Byung-Do, Mengze Shi, and Kannan Srinivasan (2004), “Managing Capacity through Reward Programs”, *Management Science,* 50 (4), 503-520

Knott Aaron, Andrew Hayes, and Scott A. Neslin (2002), “Next-Product-to-Buy Models for Cross-Selling Applications”, *Journal of Interactive Marketing*, 16(3), 59-75

Kulp, Susan Cohen, Hau L. Lee, and Elie Ofek, (2004), “Manufacturer Benefits from Information Integration with Retail Customers”, *Management Science,* 50(4), 431-444

Li, Shibo, Baohong Sun, and Ronald T. Wilcox (2005), “Cross-Selling Sequentially Ordered Products: An Application to Consumer Banking Services”, *Journal of Marketing Research*, 2005(5), 233-239

Narayannan, Sridhar and Puneet Manchanda (2006), “Heterogeneous Learning and the Targeting of Marketing Communication for New Products”, University of Chicago, GSB

Park, Chan Su, and V. Srinivasan (1994), “A Survey-Based Method for Measuring and Understanding Brand Equity and Its Extendibility”, *Journal of Marketing Research,* May, 271-288

Rangaswamy, Arvind, Raymond R. Burke, and Terence A. Oliva (1993), “Brand Equity and the Extendibility of Brand Names”, *International Journal of Research in Marketing*, 1993(10), 61-75

Rogers, D.S. and R.S. Tibben-Lembke (2001), "An Examination of Reverse Logistics Practices," *Journal of Business Logistics*, 22 (2), 129-248

Rossi, P. R. McCulloch, G. Allenby (1996), “On the Value of Household Purchase Informatioon in Target Marketing”, *Marketing Science,* 15(4), 321-340

Savaskan, R., S. Bhattacharya, and Van Wassenhove (2004), "Closed-loop Supply Chain Models with Product Remanufacturing," *Management Science,* 50 (2), 239-52.

Srinivasan, V., Chan Su Park, Dae Ryun Chang (2005), “An Approach to the Measurement, Analysis, and Prediction of brand equity and Its Sources”, *Management Science,* 51(9), 1433-1448

Tang, Christopher S., Kumar Rajaram and Aydin Alptekinoglu (2004), “The Benefit of Advance Booking Discount Programs: Model and Analysis”, *Management Science,* 50(4), 465-478

Tanner, Martin and W. Wong (1987), “The Calculation of Posterior Distributions by Data Augmentation”, *Journal of the American Statistical Association*, 85, 528-550

Thaler, Richard (1980), “Toward a Positive Theory of Consumer Choice”, *Journal of Economic Behavior and Organization,* 1 (March), 39-60

Toktay, L.Beril, E.A. van der Laan, and M.P. de Brito (2003), "Managing Product Returns: the Role of Forecasting," *Reverse Logistics: Quantitative Models for Closed-loop Supply Chains*, R. Dekker, M. Fleishchmann, K. Inderfurth, L. N. Van Wassenhove, Editors.

Woodruff, Robert B., Emest R. Cadotte, and Roger L. Jenkins (1983), “Modeling Consumer Satisfaction Process Using Experience-Based Norms”, *Journal of Marketing Research,* 20(Aug), 296-304

Wright, Alice A. and John G. Lynch, JR. (1995), “Communication Effects of Advertising Versus Direct Experience When both Search and Experience Attributes Are present”, *Journal of Consumer Research,* Vol. 21 (Mar), 708-718

Appendix 2.1

Derivation of the Recursive Updating Equation for Σ_{it+1}

$$\Sigma_{it+1} = (\tilde{\Sigma}_{it}^{-1} + \begin{bmatrix} T_{it}^{E}\pi_i^{-2} + T_{it}^{S}\sigma_i^{-2} & \mathbf{0} \\ \mathbf{0}' & 0 \end{bmatrix})^{-1}$$

$$= (I_{(K+1)*(K+1)} + \tilde{\Sigma}_{it} \begin{bmatrix} T_{it}^{E}\pi_i^{-2} + T_{it}^{S}\sigma_i^{-2} & \mathbf{0} \\ \mathbf{0}' & 0 \end{bmatrix})^{-1}\tilde{\Sigma}_{it}$$

$$= (I_{(K+1)*(K+1)} + \begin{pmatrix} \Sigma_{it}^{K*K} & \tilde{W}_{it} \\ \tilde{W}_{it}' & \delta_{bit}^2 \end{pmatrix} \begin{bmatrix} T_{it}^{E}\pi_i^{-2} + T_{it}^{S}\sigma_i^{-2} & \mathbf{0} \\ \mathbf{0}' & 0 \end{bmatrix})^{-1} \begin{pmatrix} \Sigma_{it}^{K*K} & \tilde{W}_{it} \\ \tilde{W}_{it}' & \delta_{bit}^2 \end{pmatrix}$$

$$= \begin{pmatrix} \Sigma_{it}^{K*K}(T_{it}^{E}\pi_i^{-2} + T_{it}^{S}\sigma_i^{-2}) + I_{K*K} & \mathbf{0} \\ \tilde{W}_{it}'(T_{it}^{E}\pi_i^{-2} + T_{it}^{S}\sigma_i^{-2}) & 1 \end{pmatrix}^{-1} \begin{pmatrix} \Sigma_{it}^{K*K} & \tilde{W}_{it} \\ \tilde{W}_{it}' & \delta_{bit}^2 \end{pmatrix}$$

$$= \begin{pmatrix} (\Sigma_{it}^{K*K}(T_{it}^{E}\pi_i^{-2} + T_{it}^{S}\sigma_i^{-2}) + I_{K*K})^{-1} & \mathbf{0} \\ -\tilde{W}_{it}'(T_{it}^{E}\pi_i^{-2} + T_{it}^{S}\sigma_i^{-2})(\Sigma_{it}^{K*K}(T_{it}^{E}\pi_i^{-2} + T_{it}^{S}\sigma_i^{-2}) + I_{K*K})^{-1} & 1 \end{pmatrix} \begin{pmatrix} \Sigma_{it}^{K*K} & \tilde{W}_{it} \\ \tilde{W}_{it}' & \delta_{bit}^2 \end{pmatrix}$$

$$= \begin{pmatrix} X_{11} & X_{12} \\ X_{21} & X_{22} \end{pmatrix}$$

where

$$X_{11} = (I_{K*K} + \Sigma_{it}^{K*K}(T_{it}^{E}\pi_i^{-2} + T_{it}^{S}\sigma_i^{-2}))^{-1}\Sigma_{it}^{K*K} = ((\Sigma_{it}^{K*K})^{-1} + T_{it}^{E}\pi_i^{-2} + T_{it}^{S}\sigma_i^{-2})^{-1} = \Sigma_{it+1}^{K*K}$$

$$X_{12} = (\Sigma_{it}^{K*K}(T_{it}^{E}\pi_i^{-2} + T_{it}^{S}\sigma_i^{-2}) + I_{K*K})^{-1}\tilde{W}_{it} = W_{it+1}$$

$$X_{21} = -\tilde{W}_{it}'(T_{it}^{E}\pi_i^{-2} + T_{it}^{S}\sigma_i^{-2})(\Sigma_{it}^{K*K}(T_{it}^{E}\pi_i^{-2} + T_{it}^{S}\sigma_i^{-2}) + I_{K*K})^{-1}\Sigma_{it}^{K*K} + \tilde{W}_{it}'$$

$$= \tilde{W}_{it}{}'(I_{K*K} - (T_{it}^{E}\pi_i^{-2} + T_{it}^{S}\sigma_i^{-2})(\Sigma_{it}^{K*K}(T_{it}^{E}\pi_i^{-2} + T_{it}^{S}\sigma_i^{-2}) + I_{K*K})^{-1}\Sigma_{it}^{K*K})$$

$$= \tilde{W}_{it}{}'(I_{K*K} - (I_{K*K} + (\Sigma_{it}^{K*K})^{-1}(T_{it}^{E}\pi_i^{-2} + T_{it}^{S}\sigma_i^{-2})^{-1})^{-1})$$

$$= \tilde{W}_{it}{}'((T_{it}^{E}\pi_i^{-2} + T_{it}^{S}\sigma_i^{-2})\Sigma_{it}^{K*K} + I_{K*K})^{-1}$$

$$= ((\Sigma_{it}^{K*K}(T_{it}^{E}\pi_i^{-2} + T_{it}^{S}\sigma_i^{-2}) + I_{K*K})^{-1}\tilde{W}_{it})'$$

$$= W_{it+1}'$$

$$X_{22} = -\tilde{W}_{it}'(T_{it}^{E}\pi_i^{-2} + T_{it}^{S}\sigma_i^{-2})(\Sigma_{it}^{K*K}(T_{it}^{E}\pi_i^{-2} + T_{it}^{S}\sigma_i^{-2}) + I_{K*K})^{-1}\tilde{W}_{it} + \delta_{bit}^{2}$$

Appendix 2.2

Derivation of the Recursive Updating Equation for $\bar{B}_{it+1}$ and $\bar{C}_{it+1}$

$$
\begin{aligned}
\bar{Q}_{it+1} &= \begin{pmatrix} \bar{C}_{it+1} \\ \bar{B}_{it+1} \end{pmatrix} \\
&= \Sigma_{it+1}(\tilde{\Sigma}_{it}^{-1}\bar{Q}_{it} + \begin{pmatrix} T_{it}^{E}\pi_i^{-2}E_{it} + T_{it}^{S}\sigma_i^{-2}S_{it} \\ 0 \end{pmatrix}) \\
&= (I_{(K+1)*(K+1)} + \tilde{\Sigma}_{it}\begin{pmatrix} T_{it}^{E}\pi_i^{-2} + T_{it}^{S}\sigma_i^{-2} & \mathbf{0} \\ \mathbf{0}' & 0 \end{pmatrix})^{-1}\begin{pmatrix} \bar{C}_{it} \\ \bar{B}_{it} \end{pmatrix} \\
&\quad + (I_{(K+1)*(K+1)} + \tilde{\Sigma}_{it}\begin{pmatrix} T_{it}^{E}\pi_i^{-2} + T_{it}^{S}\sigma_i^{-2} & \mathbf{0} \\ \mathbf{0}' & 0 \end{pmatrix})^{-1}\tilde{\Sigma}_{it}\begin{pmatrix} T_{it}^{E}\pi_i^{-2}E_{it} + T_{it}^{S}\sigma_i^{-2}S_{it} \\ 0 \end{pmatrix} \\
&= \begin{bmatrix} \tilde{\Sigma}_{it}^{K*K}(T_{it}^{E}\pi_i^{-2} + T_{it}^{S}\sigma_i^{-2}) + I_{K*K} & \mathbf{0} \\ \tilde{W}_{it}'(T_{it}^{E}\pi_i^{-2} + T_{it}^{S}\sigma_i^{-2}) & 1 \end{bmatrix}^{-1}\begin{pmatrix} \bar{C}_{it} \\ \bar{B}_{it} \end{pmatrix} \\
&\quad + \begin{bmatrix} \tilde{\Sigma}_{it}^{K*K}(T_{it}^{E}\pi_i^{-2} + T_{it}^{S}\sigma_i^{-2}) + I_{K*K} & \mathbf{0} \\ \tilde{W}_{it}'(T_{it}^{E}\pi_i^{-2} + T_{it}^{S}\sigma_i^{-2}) & 1 \end{bmatrix}^{-1}\begin{pmatrix} \Sigma_{it}^{K*K} & \tilde{W}_{it} \\ \tilde{W}_{it}' & \delta_{bit}^{2} \end{pmatrix}\begin{bmatrix} T_{it}^{E}\pi_i^{-2}E_{it} + T_{it}^{S}\sigma_i^{-2}S_{it} \\ 0 \end{bmatrix} \\
&= \begin{pmatrix} (\Sigma_{it}^{K*K}(T_{it}^{E}\pi_i^{-2} + T_{it}^{S}\sigma_i^{-2}) + I_{K*K})^{-1}\bar{C}_{it} \\ -\tilde{W}_{it}'(T_{it}^{E}\pi_i^{-2} + T_{it}^{S}\sigma_i^{-2})(\Sigma_{it}^{K*K}(T_{it}^{E}\pi_i^{-2} + T_{it}^{S}\sigma_i^{-2}) + I_{K*K})^{-1}\bar{C}_{it} + \bar{B}_{it} \end{pmatrix} \\
&\quad + \begin{pmatrix} (\Sigma_{it}^{K*K}(T_{it}^{E}\pi_i^{-2} + T_{it}^{S}\sigma_i^{-2}) + I_{K*K})^{-1}\Sigma_{it}^{K*K}(T_{it}^{E}\pi_i^{-2}E_{it} + T_{it}^{S}\sigma_i^{-2}S_{it}) \\ (\tilde{W}_{it}' - \tilde{W}_{it}'(T_{it}^{E}\pi_i^{-2} + T_{it}^{S}\sigma_i^{-2})(\Sigma_{it}^{K*K}(T_{it}^{E}\pi_i^{-2} + T_{it}^{S}\sigma_i^{-2}) + I_{K*K})^{-1}\Sigma_{it}^{K*K})(T_{it}^{E}\pi_i^{-2}E_{it} + T_{it}^{S}\sigma_i^{-2}S_{it}) \end{pmatrix} \\
&= \begin{pmatrix} \Sigma_{it+1}^{K*K}((\Sigma_{it}^{K*K})^{-1}\bar{C}_{it} + T_{it}^{E}\pi_i^{-2}E_{it} + T_{it}^{S}\sigma_i^{-2}S_{it}) \\ \bar{B}_{it} - \tilde{W}_{it}'(T_{it}^{E}\pi_i^{-2} + T_{it}^{S}\sigma_i^{-2})\Sigma_{it+1}^{K*K}(\Sigma_{it}^{K*K})^{-1}\bar{C}_{it} + \tilde{W}_{it}'(I_{K*K} - (T_{it}^{E}\pi_i^{-2} + T_{it}^{S}\sigma_i^{-2})\Sigma_{it+1}^{K*K})(T_{it}^{E}\pi_i^{-2}E_{it} + T_{it}^{S}\sigma_i^{-2}S_{it}) \end{pmatrix}
\end{aligned}
$$

Appendix 2.3:

Algorithm for drawing the model parameters from the joint posterior distribution

To obtain the posterior distributions of all the parameters, we take draws from the joint posterior distribution using Gibbs Sampling. The algorithm for drawing the parameters consists of the following steps:

(1) Draw U_{kit} from truncated normal distribution

$U_{kit} \sim Truncated \quad N(\bar{B}_{it} + \bar{C}_{kit} + \beta_{ki} X_{kit} + \bar{\beta}\bar{X}_t \ ,1)$, with the truncation such that U_{kit} >0 if $Y_{kit} > 0$, otherwise U_{kit} <0.

(2) Draw aggregate level parameters:

$$V_r \mid \{\theta_i\}, \{Z_i\}, \Pi, g, G \sim Inverse\,Wishart(g + N, [\sum_{i=1}^{N} (\theta_i - \Pi Z_i)(\theta_i - \Pi Z_i)' + G^{-1}]^{-1})$$

$$\pi \mid \{\theta_i\}, Z, \Omega, \Phi, \hat{\pi}, \sim N([\Omega^{-1} \otimes Z'Z + \Phi^{-1}]^{-1}[\Omega^{-1} \otimes Z'Z\hat{\theta} + \Phi^{-1}\hat{\pi}], [\Omega^{-1} \otimes Z'Z + \Phi^{-1}]^{-1})$$

Where

$$\hat{\theta} = vec[(Z'Z)^{-1} Z'\Theta],\ Z = \begin{pmatrix} Z_1' \\ \ldots \\ Z_N' \end{pmatrix},\ \Theta = \begin{pmatrix} \theta_1' \\ \ldots \\ \theta_N' \end{pmatrix}$$

(3) Draw individual specific parameters θ_i using RW Metropolis Hasting Algorithm

(4) Draw signals $\{ T_{it}^{E} \pi_i^{-2} E_{it} + T_{it}^{S} \sigma_i^{-2} S_{it} \}$ using RW Metropolis Hasting Algorithm

(5) Repeat the above steps until the draws converge

Appendix 3.1

Estimation Algorithm

To complete the model, we specify the prior of Π and Ω as:

$$\pi = vec(\Pi) \sim N(\hat{\pi},\Phi) \text{ and } \Omega \sim Wishart(g,G)$$

The joint posterior distribution is:

$$\begin{aligned}&Posterior \propto \\ &\prod_i \left[\begin{array}{l} \prod_{k=1}^{K}\prod_{t=1}^{T}\left[\begin{array}{l}\exp(-.5*(b_{ikt}-\beta_{ik}-\beta_{ik}^{p}\overline{P}_{kt}-\overline{\beta}D_t)^2 I(b_{ikt}>0)^{B_{ikt}} I(b_{ikt}<0)^{1-B_{ikt}} \\ \exp(-.5*(r_{ikt}-\omega_{ik}-\omega_{ik}^{p}\overline{p}_{ikt}-\omega_i^{item}\overline{x}_{ikt}-\varpi_i^{order}\overline{o}_{ikt})^2 I(r_{ikt}>0)^{R_{ikt}} I(r_{ikt}<0)^{1-R_{ikt}}\end{array}\right] \\ |\Omega|^{-1/2}\exp(-\frac{1}{2}(\theta_i-\Pi Z_i)'\Omega^{-1}(\theta_i-\Pi Z_i))\end{array}\right] \\ &|\Phi|^{-1/2}\exp(-\frac{1}{2}(\pi-\hat{\pi})'\Phi^{-1}(\pi-\hat{\pi})')\frac{\Omega^{(g-K-1)/2}}{|G|^{g/2}}\exp(-\frac{1}{2}tr(G^{-1}\Omega))\end{aligned}$$

We take draws from the joint posterior distribution using Gibbs Sampling. The algorithm for drawing the parameters consists of the following steps:

(1) Draw b_{ikt} from truncated normal distribution

$b_{ikt} \sim Truncated \quad N(\beta_{ik}+\beta_{ik}^{p}\overline{P}_{kt}+\overline{\beta}D_t,1)$, with the truncation such that b_{ikt} >0 if $B_{ikt}=1$ and b_{ikt} <0 if $B_{ikt}=0$.

(2) Draw r_{ikt} from truncated normal distribution

$r_{ikt} \sim Truncated \quad N(\omega_{ik}+\omega_{ik}^{p}\overline{p}_{ikt}+\omega_i^{item}\overline{x}_{ikt}+\varpi_i^{order}\overline{o}_{ikt},1)$, with the truncation such that r_{ikt} >0 if $R_{ikt}=1$ and r_{ikt} <0 if $R_{ikt}=0$.

(3) Draw aggregate level parameters:

$$\Omega \mid \{\theta_i\}, \{Z_i\}, \Pi, g, G \sim Inverse\,Wishart(g+N, [\sum_{i=1}^{N}(\theta_i - \Pi Z_i)(\theta_i - \Pi Z_i)' + G^{-1}]^{-1})$$

$$\pi \mid \{\theta_i\}, Z, \Omega, \Phi, \hat{\pi}, \sim N([\Omega^{-1} \otimes Z'Z + \Phi^{-1}]^{-1}[\Omega^{-1} \otimes Z'Z\hat{\theta} + \Phi^{-1}\hat{\pi}], [\Omega^{-1} \otimes Z'Z + \Phi^{-1}]^{-1})$$

where

$$\hat{\theta} = vec[(Z'Z)^{-1}Z'\Theta],\ Z = \begin{pmatrix} Z_1' \\ \dots \\ Z_I' \end{pmatrix},\ \Theta = \begin{pmatrix} \theta_1' \\ \dots \\ \theta_I' \end{pmatrix}$$

(4) Draw individual specific parameters θ_i using Random Walk Metropolis Hasting Algorithm

(5) Repeat the above steps until the draws converge