Cross-Selling Sequentially Ordered Products: An Application to Consumer Banking Services

Shibo Li, Baohong Sun and Ronald T. Wilcox[1]

Contact Author

Ronald T. Wilcox
Darden Graduate School of Business Administration
University of Virginia
P.O. Box 6550
Charlottesville, VA 22906-6550
Phone: (434) 243-5558; Fax (434) 243-7680
Email: WilcoxR@darden.virginia.edu

[1] Shibo Li is a doctoral candidate at the Graduate School of Industrial Administration, Carnegie Mellon University. Baohong Sun is an Assistant Professor of Marketing at Kenan-Flagler Business School of University of North Carolina at Chapel Hill. Ronald T. Wilcox is an Associate Professor of Business Administration at the Darden Graduate School of Business Administration, University of Virginia. The authors would like to thank Peter Boatwright, Ron Goettler, Frenkel ter Hofstede, Ajay Kalra, Puneet Manchanda, Alan Montgomery, and Kannan Srinivasan for valuable comments.

Cross-Selling Sequentially Ordered Products: An Application to Consumer Banking Services

Abstract

In service and high-technology industries, we often observe consumers sequentially purchasing multiple products and services from the same provider. These sequential purchases can take place over an extended period of time and can be naturally ordered in terms of complexity and functionality. This commonly observed situation offers significant opportunities for companies carrying multiple products and services to "cross-sell" other products and services to their existing customer base.

In this paper, we propose a dynamic multivariate probit model that incorporates households' purchase decisions about all of the currently available products and services and investigates the sequential acquisition pattern of these items. Using data obtained from a large Midwestern bank, we demonstrate how the model can be used to predict to whom and when to cross-sell which new products and services.

While recent research has focused on the use of scanner data to probe consumer purchase behavior for frequently purchased package products, this research contributes to the literature by being the first paper to investigate consumers' sequential acquisition decisions for multiple products and services, a behavior that is common in service and consumer technology industries. At a practical level, natural ordering and the increased predictive accuracy that flow from it can enable managers to develop and execute better cross-selling tactics.

1. Introduction

We often observe consumers purchasing multiple products and services from the same provider over time.[1] These products can be naturally ordered in terms of complexity and functionality leading to the empirical regularity that the purchase of certain products often precedes the purchase of others to meet a consumer's evolving demand. For example, we often observe that a person generally establishes a checking account with a given bank before establishing a brokerage account. A woman may repeatedly patronize a salon or spa for only a haircut before moving on to purchasing facial treatments. A consumer may also sequentially purchase local and long distance telephone service, cable television service, and Internet access from the same company. A person who purchases a Palm Pilot may acquire an Internet connection, additional memory chips, and software from the same provider in the future. The common thread running through each of these examples is that consumers are more likely to purchase some product or subset of products before others. The authors have directly observed this phenomenon in the market for consumer banking services, but strongly suspect that it is prevalent in many other environments. We term the over-time development of consumers' complementary demand for multiple products and services as "sequential or natural ordering" among these products and services.

Markets especially prone to this empirical regularity include those in which consumers need to purchase multiple products or services to satisfy their evolving wants, those in which consumers face some uncertainty about the quality of the product or service offering or markets in which some consumer learning is required to receive the full benefit of the product. In such markets, sequentially purchasing multiple products or services from the same provider can enhance the relationship with the provider, lower switching costs associated with moving to a new provider, lower uncertainty with respect to additional product purchases, and, in some cases, ensure proper technical compatibility with products the consumer already owns.

The existence of sequentially developed demand for naturally ordered products offers great opportunities for companies carrying multiple products and services to

[1]Throughout this paper we will refer to products and services interchangeably. The model we develop is equally applicable to both.

"cross-sell" other products and services to their existing customer base. These companies are eager to explore additional profitable opportunities with their current customer base by cross-selling because it is often cheaper to cross-sell to their existing customers than to attract new ones. Further, customer retention is enhanced with the cross-selling of multiple accounts or services as customer switching cost increases with multiple relationships (Kamakura, Ramaswami and Srivastava 1991). Indeed, many companies have come to realize that their current customers are by far the best prospects for new or existing products and services. Many large providers have invested significant capital in the information technology necessary to develop large scale customer transactional databases and to implement new marketing tactics made possible by the intelligence gleaned from these databases (*Information Week* 2001; *Network World* 2001). This new marketing intelligence has substantially increased the ability of many providers to cross-sell additional products and services to current customers. Careful observation of customers' current and past purchasing behavior can lead to inferences about other products and services that they might want to purchase now or in the future. One of the challenges faced by these providers is how to develop the best ways to make these inferences and maximize the potential value of firms' marketing information technology investment.

There now exists a reasonably mature stream of the marketing literature that explores similarities in purchase patterns across categories for frequently purchased packaged goods. Work in this area centers on similar levels of brand loyalty across several categories for the same household (Cunningham 1956; Massy, Frank, and Lodahl 1968; Wind and Frank 1969), and uncovering underlying household demographic characteristics that drive similarities in cross-category choice behavior (Blattberg, Peacock, and Sen's 1976). More recently, research has linked similarities in category-level price sensitivities to households' demographic profile (Ainslie and Rossi 1999), shopping patterns (Kim, Srinivasan, and Wilcox 1999), and observable category characteristics (Fader and Lodish 1990; Raju 1992, and Narasimhan, Neslin, and Sen 1996). This body of research is generally descriptive in nature. One empirical paper that focuses on modeling and predicting multi-category purchases is Manchanda, Ansari, and Gupta (1999). This research develops a multivariate probit model to investigate the complementarity arising from cross-category promotions and coincidence effect arising from unobserved factors.

However, there have been relatively few studies over the past four decades that probe consumers' sequential purchase of items (Pyatt, 1964; Paroush, 1965; McFall, 1969; Hebden and Pickering, 1974; Kasulis, Lusch and Stafford, 1979; Clarke and Soutar, 1982; Dickson, Lusch and Wilkie, 1983; Hauser and Urban 1986; Mayo and Qualls, 1987). This research focuses most of its attention on explaining the existence of sequential acquisition patterns. Some basic reasons proposed by this literature include "logical ordering" and resource constraints. To our knowledge, the earliest paper that formally models sequential ordering and the cross-selling opportunities that arise from it is Kamakura, Ramaswami and Srivastava (1991). Their research applies latent trait analysis to position financial services and investors along a common continuum. Using this approach, they obtain the estimates of the ordering of financial services as well as the latent financial maturity for each household based on (only) the current ownership of financial services. Because their model focuses more on inferring acquisition order of financial services from a one-time measurement of (non-) ownership information across households, it does not model consumer purchase decisions that are made periodically and hence does not accommodate the development of complementary demand over time. In addition, their purpose is to predict what type of consumer is more beneficial to target in the future rather than when an individual should be targeted. Recently, Kamakura and Kossar (2001) develop a split hazard rate model and focus on predicting each customer's (physician's) time of adopting a new product (drug) based on the timing of their past adoptions of multiple products. Knott, Hayes and Neslin (2002) present four next-product-to-purchase models (discriminant analysis, multinomial logit, logistic and neural net) that can be used to predict what is to be purchased next and when. Using only cross-sectional data, both of these papers are aimed at inferring adoption (time) from past adoption (time) of similar products. While useful, the utilization of only cross-sectional data constrains these authors' ability to model the development of sequential demands for multiple products and services over time.

In this paper, we propose a dynamic multivariate probit model that allows households to make periodic purchase decisions spanning all of the available products and services.

We also investigate the sequential acquisition pattern of these products and services and the fundamental determinants that drive customer sequential needs. Further, we demonstrate how this model can be used to predict to whom and when to cross-sell which new products and services. Specifically, our model differs from Kamakura, Ramaswami, and Srivastava (1991), Kamakura and Kossar (2001), Knott, Hayes and Neslin (2002), and Manchanda, Andsari, and Gupta (1999) in the following ways. First, our dynamic multivariate probit model can be applied to panel data and thus can be used to investigate consumers' acquisition pattern and the development of complementary demand over time. Second, because our model allows consumers to make periodic purchase decisions over all of the available products/services, it can be used to determine when to cross sell. Third, although our model emphasizes modeling over-time cross selling, it is general enough to accommodate same-time cross selling as in Manchanda, Andsari, and Gupta (1999). Finally, our model allows the inclusion of explanatory variables that help explain the development of sequential demand for products and services from the same provider. In doing so, it provides behavioral explanations for why and how consumers' demands for different products/services develop over time and predicts to whom and when cross-selling is likely to be most productive. Methodologically, it is the first paper that studies consumer sequential acquisition decisions for products that are naturally ordered, a behavioral generalization that is common in service and consumer technology industries.

The rest of the paper is organized as follows. In Section 2, we develop the dynamic multivariate probit model. In Section 3, we apply this model to customer-level account data provided to us a by a large Midwestern bank. We then discuss both modeling issues as well as some substantive findings that arise from this application. Finally, Section 4 discusses our findings, points to the limitations of our research, and lays out some directions for future research.

2. **Model Development**

In this section, we develop a dynamic multivariate probit model to capture a customer's over-time acquisition pattern for multiple products and services from a single provider. We assume that household $i = 1, ..., I$ makes purchase decisions on product set

J, naturally ordered at each time period t from 1 to J. Product j is more likely to be purchased before product l for $l>j$ than visa versa and so forth up the ordering hierarchy. Assume the utility of a typical household choosing product $j = 1, 2, \ldots, J$ at occasion t is given by:

$$\begin{aligned}
U_{i1t} &= \alpha_{0i} + \alpha_{11} + \beta_{1i1}\sum_{l=1}^{t-1} y_{i,1,l} + \beta_{2i1}\sum_{l=1}^{t-1} y_{i,2,l} + \ldots + \beta_{Ji1}\sum_{l=1}^{t-1} y_{i,J,l} + X_{ijt}^{'}\gamma_{i1} + \varepsilon_{i1t} \\
U_{i2t} &= \alpha_{0i} + \alpha_{12} + \beta_{1i2}\sum_{l=1}^{t-1} y_{i,1,l} + \beta_{2i2}\sum_{l=1}^{t-1} y_{i,2,l} + \ldots + \beta_{Ji2}\sum_{l=1}^{t-1} y_{i,J,l} + X_{ijt}^{'}\gamma_{i2} + \varepsilon_{i2t} \quad (1) \\
&\cdots \\
U_{iJt} &= \alpha_{0i} + \alpha_{1J} + \beta_{1iJ}\sum_{l=1}^{t-1} y_{i,1,l} + \beta_{2iJ}\sum_{l=1}^{t-1} y_{i,2,l} + \ldots + \beta_{JiJ}\sum_{l=1}^{t-1} y_{i,J,l} + X_{ijt}^{'}\gamma_{iJ} + \varepsilon_{iJt}
\end{aligned}$$

Purchase decision variables y_{ijt} are determined by

$$y_{ijt} = \begin{cases} 1 & \text{if } U_{ijt}>0 \\ 0 & \text{otherwise.} \end{cases} \qquad (2)$$

for all $i=1, \ldots, I$, $j=1, \ldots, J$ and $t=1, \ldots, T$. U_{ijt} is the latent utility of product j for household i at time t. α_{0i} and α_{1j} are scalar coefficients capturing consumers' intrinsic preference for all the products provided by the service provider. The term $\sum_{l=1}^{t-1} y_{i,j,l}$ is the cumulative number of purchases of product j for consumer i up to time $t-1$. This captures the experience or satisfied needs for product j and represents the maturity of consumer demand at the beginning of time t. This is consistent with the findings of Kamakura and Kossar (2001) and Knott, Hayes, and Neslin (2002), that current ownership is a strong predictor for future purchases. If we define β as the matrix $\begin{pmatrix} \beta_{1i1} & \beta_{2i1} & \cdots & \beta_{Ji1} \\ \beta_{1i2} & \beta_{2i2} & \cdots & \beta_{Ji2} \\ & \cdots & & \\ \beta_{1iJ} & \beta_{2iJ} & \cdots & \beta_{JiJ} \end{pmatrix}$, the diagonal elements measure how the cumulative past experience, familiarity, or knowledge about a product affects the purchase decision of the same product. The off-diagonal elements measure whether the cumulative purchases of other products effect the purchase of the given product. In the financial service industry, the estimated β describes how the fulfillment of customer financial needs drives the demand for a new product along the naturally ordered continuum (if there is any). The vector $\boldsymbol{X}_{ijt}$ ($K\times 1$) is a vector of product and consumer-related covariates that captures the behavioral reasons that effect consumers' purchases of products $j=1, \ldots, J$ over time from the same provider. $\boldsymbol{\gamma}_{ij}$ is a Kx1 vector of coefficients for $\boldsymbol{X}_{ijt}$. The factors that affect customer purchase decisions but are unobservable to researchers are denoted by $\boldsymbol{\varepsilon}_{it}=(\boldsymbol{\varepsilon}_{i1t}, \boldsymbol{\varepsilon}_{i2t}, \ldots, \boldsymbol{\varepsilon}_{iJt})'$, a $J\times 1$ vector of error terms that are assumed to be normally distributed with mean **0** and variance-covariance matrix $\boldsymbol{\Sigma}$ ($J\times J$ matrix). We allow the unobserved part of the utilities to be correlated. Formally:

$$\varepsilon_{it} \sim MVN[0,\ \Sigma] \tag{3}$$

The $\boldsymbol{\Sigma}$ matrix was termed purchase "coincidence" in Manchanda et al., (1999) and captures the contemporaneous complementary demands for multiple products taking place on the same shopping trip.[2] There are many possible reasons that cause simultaneous purchase of different products at the same purchase incidence. As one example, banks may bundle checking and saving accounts together at a discount price.

Note that the application of this model does not require the manager's *a priori* ordering of the products. The estimated coefficient pattern of $\boldsymbol{\beta}_{jil}$ for $l=1, \ldots, J$ and $j=1,\ldots, J$ will recover the natural ordering inherent in the data. For example, if the estimated β matrix looks like $\begin{pmatrix} + & - & \cdots & - & - \\ + & + & \cdots & - & - \\ & & \cdots & & \\ + & + & \cdots & + & + \end{pmatrix}$, this implies that cumulative ownership of product j for all $j=1, \ldots, J$ increases the probability of purchasing products l for $l>j$ and decreases the

[2] The covariance components of **Σ** capture the unobserved correlation among categories and were described as purchase co-incidence by Manchanda, Ansari, and Gupta (1999). The magnitude of the components of **Σ** not only indicate the strength of the unobserved purchase incidence correlations among categories but also reflects the effect of natural ordering in that categories that are closer in the sequential ordering are more likely to be correlated than those categories that are further apart.

probability of purchasing products l for $l<j$.[3] This represents a prototypical natural ordering $1<2<\ldots<J$. Thus, our model is flexible enough to handle a situation where no manager's initial beliefs about the ordering is available. Also note that our model allows consumers to make periodic purchase decisions of all the products, not just those that are currently higher in the purchase hierarchy.

Previous research suggests that consumers are heterogeneous in their reaction to marketing mix variables (e.g., Kamkura and Russell 1989; Gönül and Srinivasan 1996; Krishna, Currim, and Shoemaker 1991). It is also known that ignoring unobserved consumer heterogeneity might lead to biased estimates of β (Heckman 1981; Allenby and Rossi 1999). In order to take into account consumer heterogeneity, we adopt the hierarchical Bayesian model by letting α_{0i}, $\boldsymbol{\beta_{ji} = (\beta_{ji1}, \beta_{ji2}, \ldots, \beta_{jiJ})'}$ and $\gamma_{ki=}(\gamma_{ki1}, \gamma_{ki2}, \gamma_{kiJ})'$ be linear functions of household specific variables.

$$\begin{cases} \alpha_{0i} = D_i'\eta + v_i \\ \beta_{ji} = D_i'\mu_j + e_{ji,} \\ \gamma_{ki} = D_i'\omega_k + \lambda_{\mathrm{ki}} \end{cases} \tag{4}$$

for $j=1, \ldots, J$ and $k = 1, \ldots, K$. $\boldsymbol{D_i}$ is a $m\times 1$ matrix containing $m-1$ household specific variables and a "1" for the intercept. η, μ_j, and ω_k are $m\times 1$ vectors with each element measuring the effect of D_i variable on α_{0i} , β_{ji}, and γ_{ki}, respectively. We assume v_i ~ N[0, σ_v^2], e_{ji} ~ MVN[**0**, $\boldsymbol{\Lambda_e}$] and λ_{ki} ~ MVN[**0**, $\boldsymbol{\Lambda_\lambda}$].

To solve the identification problem, we fix the scale of the utilities by dividing each utility by its corresponding standard deviation and set one α_{0i} to zero. The identification restrictions thus transform **Σ** to a correlation matrix. In subsequent sections, we use the same set of symbols (**U**, **Σ**, and **Φ**) to denote utilities and identified

[3] There are situations when the signs of elements in β matrix do not show a clear pattern. Then some partial ordering can be inferred. For example, a 3x3 β matrix $\begin{pmatrix} + & - & - \\ + & + & + \\ - & - & + \end{pmatrix}$ implies 1<2 and 3<2. But there is no clear ordering between products 1 and 3 which indicates a possible substitute relationship between 1 and 3.

parameters. Define **I** as a J x 1 vector of ones, $C_{it} = \begin{pmatrix} \sum_{l}^{t-1} y_{i,1,l} & \sum_{l}^{t-1} y_{i,2,l} & \cdots & \sum_{l}^{t-1} y_{i,J,l} \\ \sum_{l}^{t-1} y_{i,1,l} & \sum_{l}^{t-1} y_{i,2,l} & \cdots & \sum_{l}^{t-1} y_{i,J,l} \\ & \cdots & & \\ \sum_{l}^{t-1} y_{i,1,l} & \sum_{l}^{t-1} y_{i,2,l} & \cdots & \sum_{l}^{t-1} y_{i,J,l} \end{pmatrix}$ and

X_{it} as J x K matrix. Then $\boldsymbol{Z_{it}} = \begin{pmatrix} \mathrm{I} & \mathrm{I} & C_{i1} & X_{i1t} & 0 & 0 & 0 & 0 & 0 \ldots 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & \mathrm{I} & \mathrm{I} & C_{i2} & X_{i2t} & 0 \ldots 0 & 0 & 0 & 0 & 0 \\ & & & \cdots & & & & & & & & & \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \ldots 0 & \mathrm{I} & \mathrm{I} & C_{iJ} & X_{iJt} \end{pmatrix}$ is a J x F matrix where $F= J$ x $(2+J+K)$. Define $\boldsymbol{\Phi_i} = \{\alpha_{0i}, \alpha_{11}, \beta_{i1}, \gamma_{i1}', \ldots, \alpha_{0i}, \alpha_{1J}, \beta_{iJ}, \gamma_{iJ}'\}'$ as an F x 1 coefficient vector. Then the utilities for any given consumer on any given purchase occasion can be compactly expressed as $\mathbf{u}_{it} = \mathbf{Z}_{it}\boldsymbol{\Phi}_i + \boldsymbol{\varepsilon}_{it}$, where $\mathbf{u}_{it}$ and $\boldsymbol{\varepsilon}_{it}$ are $J\times 1$ vectors. Given the error structure we impose, our model is a canonical multivariate Probit specification, and hence the probability of category purchase for each household at time t is given by:

$$\Pr(Y_{it} = y_{it}) = \int_{M_1} \cdots \int_{M_J} \frac{1}{\sqrt{(2\pi)\det(\Sigma)^{1/2}}} \exp(-\frac{1}{2}\varepsilon_{it}' \Sigma^{-1} \varepsilon_{it}) d\varepsilon_{it} \tag{5}$$

where $\boldsymbol{\varepsilon}_{it} = \mathbf{u}_{it} - \mathbf{Z}_{it}\boldsymbol{\Phi}_i$, $M_j = (-\infty, 0)$ if $y_{ijt} = 0$ and $(0, \infty)$ otherwise. $\boldsymbol{Y_{it}}$ is the observed profile (J×1 vector) of binary choices of y_{ijt} for household i at time t.

This likelihood involves computation of high-order multidimensional integrals making classical inference based on maximum likelihood estimators difficult. A hierarchical Bayesian approach is introduced to estimate the model, accounting for both observed and unobserved heterogeneity (Gelfand and Smith 1990, Allenby and Rossi 1999). This type of estimation has become reasonably commonplace in the marketing literature, and for that reason we relegate the details of the estimation to the appendix.

3. Data, Empirical Application and Discussion

3.1 Data Description

To demonstrate the efficacy of our approach, we applied this model to household-level data collected from a large midwestern bank. Specifically, the bank gave us access to holding and transaction information of all 20 financial products it offered for 1201 randomly selected households from July 1997 to June 1998. This data included monthly observations on which accounts/products the customer had purchased or held as an investment. We also had access to demographic information for each of these customers. Finally, the bank provided us with the results of a customer satisfaction survey completed by each of the customers in the sample just before July 1997. Because our data was at the household level, we also observed repeat purchases. A brief description of the variables used in this paper is shown in Table 1.

[Insert Table 1 About Here]

Given the one-year observation period, we observed very few purchases for some of the accounts. For example, in our sample, nobody purchased the "Trust" service offered by the bank over a prescribed time period. Since our focus is to demonstrate how our proposed model can be used to predict cross-selling probabilities instead of investigating the exact purchase order of the 20 products, we grouped together similar accounts to ensure a sufficient number of purchases in each category. The categorization we used is based on Kamakura, Ramaswami, and Srivastava (1991). Thus, we divided all of the accounts offered by the bank into three categories. We named these categories Convenience Services, Cash Reserves, and Advanced Services without implying any order.[4] Convenience Services include very commonly held products such as checking accounts and mortgages. Cash Reserves represents more intensive use of bank services and includes products such as CD's and IRA's, and Advanced Services includes more complex products such as Brokerage Services and Annuities.

[4] This nomenclature follows Kamakura, Ramaswami, and Srivastava (1991) quite closely.

[Insert Table 2 About Here]

3.2 Empirical Application

Formally, we write the utility function of any household *i* for bank service *j* at time *t* as:

$$
\begin{aligned}
U_{i1t} &= \alpha_{0i} + \alpha_{11} + \beta_{1i1}\sum_{l=1}^{t-1} y_{i,1,l} + \beta_{2i1}\sum_{l=1}^{t-1} y_{i,2,l} + \beta_{3i1}\sum_{l=1}^{t-1} y_{i,3,l} \\
&\quad + \gamma_{1i1} AGE1_i + \gamma_{2i1} AGE2_i + \gamma_{3i1} AGE3_i + \gamma_{4i1} INCOME_i + \gamma_{5i1} COMPET_{i1} \\
&\quad + \gamma_{6i1} OVERSAT_i + \gamma_{7i1} SWIT_i + \varepsilon_{i1t} \\
U_{i2t} &= \alpha_{0i} + \alpha_{12} + \beta_{1i2}\sum_{l=1}^{t-1} y_{i,1,l} + \beta_{2i2}\sum_{l=1}^{t-1} y_{i,2,l} + \beta_{3i2}\sum_{l=1}^{t-1} y_{i,3,l} \\
&\quad + \gamma_{1i2} AGE1_i + \gamma_{2i2} AGE2_i + \gamma_{3i2} AGE3_i + \gamma_{4i2} INCOME_i + \gamma_{5i2} COMPET_{i2} \qquad (6) \\
&\quad + \gamma_{6i2} OVERSAT_i + \gamma_{7i2} SWIT_i + \varepsilon_{i2t} \\
U_{i3t} &= \alpha_{0i} + \alpha_{13} + \beta_{1i3}\sum_{l=1}^{t-1} y_{i,1,l} + \beta_{2i3}\sum_{l=1}^{t-1} y_{i,2,l} + \beta_{3i3}\sum_{l=1}^{t-1} y_{i,3,l} \\
&\quad + \gamma_{1i3} AGE1_i + \gamma_{2i3} AGE2_i + \gamma_{3i3} AGE3_i + \gamma_{4i3} INCOME_i + \gamma_{5i3} COMPET_{i3} \\
&\quad + \gamma_{6i3} OVERSAT_i + \gamma_{7i3} SWIT_i + \varepsilon_{i3t}
\end{aligned}
$$

As stated earlier, the error term is assumed to be normally distributed with mean **0** and variance-covariance matrix **Σ**. For identification purposes, α_{13} is set to zero.

Our model includes some basic household demographic information. $AGE1_i$, $AGE2_i$ and $AGE3_i$ are defined as dummy variables indicating if the account owner *i*'s age is less than 19, between 19 and 33, and between 34 and 60, respectively. $INCOME_i$ is household *i*'s income. The reason for including these variables is to control for the effects of household lifecycle and income on purchase propensity. We believed that older individuals and those with relatively higher incomes were likely to have already purchased some of the more basic financial services prior to the time period our data covers. Similarly, overall satisfaction of the service provider plays an important role in determining household's future purchases of similar products from the same service provider. We define $OVERSAT_i$ as the overall satisfaction score of household *i* and $COMPET_{ij}$ as a dummy variable such that it is equal to one when household *i* has opened an account in category *j* with another bank in the past six months. It controls for competition.

We also included household-level switching cost ($SWIT_i$) in our analysis. It is defined as one if the account owner's profession is white collar, and the household has at least one non-adult child, and the household owns more than the average number of accounts with this bank. The logic of this inclusion follows closely the reasoning of Blattberg, Buesing, Peacock, and Sen (1978), and more primitively the work of Becker (1965). In short, we believe that households that have a high cost of time, as evidenced by increasing levels of education or the presence of children, will tend to spend less time shopping around for banking services. Further, individuals that have a relatively complex relationship with the bank (have a large number of accounts there) will also suffer greater inconvenience if they choose to switch banking service providers. While switching grocery stores is relatively easy, switching service providers can often entail a significantly higher degree of discomfort. The reasons for this greater inconvenience are reasonably self-evident, but include such unpleasantness as filling out all of the paperwork necessary for opening up a series of accounts at a new bank. Therefore, our measure of switching cost takes into account both intrinsic time costs as well as extrinsic costs associated with the customer's current level of relationship with the bank.

We now write the utility coefficients as a function of other household-level observables. This allows for household characteristics to influence the weight certain factors will play in the utility function. Formally, we specify the coefficient heterogeneity as:

$$\begin{cases} \alpha_{i0} = \eta_0 + \eta_1 EDUCAT_i + \eta_2 GENDER_i + v_i \\ \beta_{ji} = \mu_{0j} + \mu_{1j} EDUCAT_i + \mu_{2j} GENDER_i + e_{ji} \\ \gamma_{ki} = \omega_{0k} + \omega_{1k} EDUCAT_i + \omega_{2k} GENDER_i + \lambda_{ki} \end{cases} \tag{7}$$

where $EDUCAT_i$ is account owner *i*'s education level. It is coded such that the value one indicates some high school education and five indicates post-graduate education etc. $GENDER_i$ is defined as one if the account owner is male and zero otherwise. We believe that better educated and male household heads are more knowledgeable and better informed about the more advanced financial products.[5] They are also more confident in

[5] For an explanation of the relationships among gender, confidence, and financial decision making see Barber and Odean (2001).

managing risky investments. Thus, we conjecture that the adoption speed or movement along the financial maturity continuum may be faster for households that are better educated and with a male head. For similar reasons, these types of households may be less sensitive to variables like satisfaction and switching cost. Thus, we chose to include *EDUCAT* as part of the utility weight in the heterogeneity expression. We note that one could reasonably argue that level of education (*EDUCAT*) should be included directly in the utility function, forming part of the composite variable $SWIT_i$. Certainly Becker (1965) argued that education was related to the cost of time. However, we were particularly interested in the influence of level of education on the impact of the β matrix on purchase probability.[6]

Our final task is to specify the top level of the hierarchy, the prior distributions of the parameters defined in the heterogeneity expression. We chose to specify very diffuse priors. Those interested in the exact specification may consult the appendix.

In the interest of comparing our model to other plausible specifications, we estimate four separate models. The first model assumes the category choices are independent. They are simply three individual binary probit models, one for each category. This model ignores potential unobserved correlation across categories and natural ordering of the products. We include this model because this type is commonly adopted by industry to predict probabilities of cross-selling. This model is also similar to Kamakura, Ramaswami, and Srivastava (1991) and Knott, Hayes, and Neslin (2002) in the sense that they assume consumers make independent decisions about the purchase of different products. The second model is a multivariate probit model, which allows for random errors in the utility functions to be correlated across different products. This model is similar to Manchanda, Ansari, and Gupta (1999). It only captures the contemporaneous (non-over-time) complementary relationship among products on the same purchase occasion. We include this model as one benchmark model because this is the most closely related model existing in the marketing literature. The third model adds to the second model the switching cost a customer might incur by changing service providers. We believe that switching cost, which indicates the customer relationships built over the year as well as the time value to the households, plays an important role in providing the company with opportunities to cross-sell their other products. These Becker-type effects are important in markets where changing product or service providers is particularly costly in terms of the time needed to complete the change. The fourth model is our proposed model. It is a multivariate probit model with dynamic effects and switching cost to allow for the possibility that the temporal ordering of purchases affects the utility of product options not yet exercised. It captures the existence of natural ordering effects that are caused by the over-time sequential demand of products offered by the same provider.

[Insert Table 3 about here]

The Independent model was estimated using the Probit procedure in SAS. The remaining three models were estimated using an MCMC (Gibbs Sampler) approach. For each model we allowed a "burn-in" period of 30,000 iterations and captured the final 5,000 to compute posterior moments. We use the Bayes factor criterion for model comparison.[6] As shown in Table 3, we find that all the models allowing products to be correlated (Model 2, 3, and 4) fit the data statistically better than the independent probit models (Model 1). Model 3 fits better than Model 2 indicating that switching cost plays an important role in describing demand. It reveals that switching cost provides consumers (part of the) incentive to acquire multiple products from the same provider over time. Our proposed Model (Model 4) fits data better than any of the competing models demonstrating that adding dynamic effects to capture the natural ordering of the products improves model fit. In the remainder of the paper, we focus our discussion on our proposed model.

[6] See Kass and Raferty (1995) for a discussion of Bayes factors. A fit statistic for Bayesian model comparison can be generated by taking the log of the ratio of the marginal densities of the two competing models assuming the priors for the two models are equally likely. Here, marginal density is defined as the harmonic mean of the likelihood over the "sample" period. Specifically, if the log of this ratio is greater than two, it is taken to be "decisive" evidence for the better fit of the model in the numerator relative to that in the denominator. In the model above, we can simply take the absolute value of the difference between the two competing models' log marginal densities to obtain log of the Bayes factor.

Table 3 also reports the estimated coefficients for our proposed model. All estimated coefficients are significant in the classical sense except that of AGE1 in category 2. We first study the ordering among products represented by the coefficients of the ownership of all the products. This effect is also borne out by the significance of each element of coefficient matrix β.[7] Referring to the highlighted area of Table 3, we can see that they are all significant and present the pattern $\begin{pmatrix} + & - & - \\ + & + & - \\ + & + & + \end{pmatrix}$ suggesting that the cumulative ownership of product *j* makes consumers more likely to purchase products *l>j* and less likely to purchase products *l<j*. This implies an acquisition sequence of Category 1, Category 2, and Category 3. For example, consumers who own a checking account in Category 1 are more likely to apply for a debit card in Category 1. Once their needs for financial convenience are met, they will start to invest in CDs, life insurance, stocks and second homes. Once they start to acquire CDs and life insurance, they are less likely to purchase checking accounts and debit cards. Instead, they were more willing to invest in stocks or real estate than another home. This reflects the development of financial needs that are consistent with their lifecycles. The ownership(s) of different products indicates the household's financial maturity developed over time and the signs in β matrix indicate an natural ordering of convenience service, cash reserves and advanced financial services. The natural ordering that was recovered by our model is exactly what we *a priori* expected. It is also consistent with our conversations with bank managers as well as the hierarchy suggested by Kamakura, Ramaswami, and Srivastava (1991).

All of the included demographic variables are significant. Recall that we included age to take into account household lifecycle. Interestingly, account owners less than 19 years old are less likely to repeat purchase convenience products during our observation period, but instead accounts in advanced financial services are opened under their names. This is probably because teenagers are not financially independent. But their parents may choose to invest in more advanced financial products hoping for a higher long-term return. Young families (account owners are between 19 and 33 years of age) are more

[7] Although we use a Bayesian estimation procedure, we will follow the convention of reporting the posterior standard deviations of the parameters so that classical inference statistics may be computed.

financially mature and aggressively seek services related to cash reserves and advanced financial services.

Middle-aged families do not actively acquire new products and services from this bank. We speculate that this is due to the fact that these households are already quite mature in their holding of financial products and hence have less need to increase the complexity of their financial holding. It is also possible that these same individuals begin to look outside of the bank, to brokerage houses for example, to meet some of their emerging financial needs. For these, and perhaps other, reasons our results indicate that middle-aged households are not active in the market for additional banking services.

The inclusion of income in our model arises out of some reasonably obvious intuition. Higher income families are more likely to have the resources to invest in advance financial services. We find this to be the case.

Some additional substantive results are exactly as expected. Higher income households tend to be more likely to invest in advanced financial products. Owning accounts with other financial institutions takes away business from the bank. Overall satisfaction with the bank increases the bank's ability to cross-sell its products to existing customers, and the service quality has a higher impact on the future demand for advanced financial services and convenience services than for cash reserves. This result is intuitive in that advanced financial services, brokerage for example, require much more interaction with the bank than say a cash reserves account which requires almost no human interaction.

Switching costs play a powerful role in influencing households in all three product categories. If time costs are high, customers are unlikely to switch banks even in the face of some amount of dissatisfaction. Following the logic above, we would expect this effect to be more potent for convenience and advanced services than for cash reserves and indeed it is. While not the primary focus of this research, this result does raise some interesting questions about which customers are most important for the bank to satisfy. While conventional wisdom would certainly dictate placing a great deal of emphasis on satisfying the bank's wealthier customers our results indicate that some customers may be "trapped" by the bank owing to the substantial implicit costs a given customer might face in switching from one bank to another.

[Insert Table 4 Here]

Table 4 provides the estimated correlations among the unobserved part of the utilities. This is the "incidence correlation" termed in Manchanda, Ansari, and Gupta (1999). The positive correlations among Y1 and Y2 and Y1 and Y3 indicate that there are complementary demands for convenience products, cash reserve and advanced financial products at each purchase occasion. For example, a household may open checking and saving accounts together simply because they are bundled. The insignificant correlation between Y2 and Y3 implies that no such purchase complementarity exist between these two products. Note that these correlations capture the current inter-product relationship. This is different from the sequential demand for naturally ordered products driven by financial maturity and knowledge.

[Insert Table 5 About Here]

We now turn our attention to the consumer heterogeneity expression and examine how demographic variables like education and gender of the heads of households affect the acquisition pattern of sequentially ordered products. Table 5 reports the effects of education and gender on utility weight heterogeneity. Most of the estimates are significant. These two included variables describing knowledge and risk bearing appear to be important in characterizing the speed of adoption for financial products. For example, the positive coefficients of education and gender in heterogeneity equation for β_{1ij} (β_{1ij} measures the effect of cumulative ownership of category 1 products on the purchase probability of category j for $j=1, 2,$ and 3) indicate that owners of category 1 products with greater education or male household heads are even more likely to purchase products in category 2 and 3 than owners with lower education or female household heads. This is consistent with our observations that households with higher education or male household heads move relatively faster along the financial maturity continuum. Similarly, we can see that, once owning cash reserves, those households are less likely to invest more in convenience services and cash reserve accounts. Instead, they are more likely to seek advanced financial services. These results show that households with greater education or male-headed households are less likely to repeat purchase products they already own but are more aggressive in pursuing products that are more advanced in the sequential ranking. The coefficients of education and gender in the heterogeneity equation for γ_{i5} tell us that those households are less likely to be cross-sold once they open accounts with the bank's competitors. We conjecture that this is due to the fact that banks are facing intensive competition from specialized professional fund management companies like Charles Schwab and low-cost Internet brokerage firms like TDWaterhouse, which compete for demand of advanced financial services related products. More highly educated or male heads of households are, or believe themselves to be, more knowledgeable and confident about financial products and hence are more likely to take advantage of these opportunities.[8] Interestingly, education enhances the effects of satisfaction and switching cost on purchase probabilities in general, but male household heads seem to care less about customer service and switching cost in determining future purchases than female household heads. In summary, we found that the maturing of demand for financial products is much faster for households with educated or male household heads.

[Insert Table 6 About Here]

Because the focus of this application is to predict the probability of acquiring new products in order to help bank managers more efficiently allocate their targeting efforts, we now compare the predictive ability of our proposed model with other benchmark models. We did cross-time prediction by dividing our original sample of 1201 households into an estimation sample and a holdout sample. The estimation sample has three quarters of the households, and the holdout sample has the remaining quarter. We first applied our proposed model to the estimation sample and then used the estimated parameters and explanatory variables from the holdout sample to calculate the predicted purchase probabilities for each individual product. For demonstration purposes, in Table 6 we

[8] See Barber and Odean (2001).

reported the cross-selling probabilities for the first two periods of ten randomly selected households. We can see that the cross-selling probability predictions generated from our proposed model are quite consistent with the actual purchase data. These provide managers with information on whom and when to target in order to cross-sell products in the three categories. We expect to get more accurate predictions on when to target if we have a longer and wider time span (e.g., yearly data for 20 years).

[Insert Figure 1a and Figure 1b About Here]

In order to better test the predicting accuracy of our proposed model, we conducted three exercises. Figure 1a and Figure 1b depict the mean absolute error between the purchase probability and the actual purchase realization for each of the four models across the three products using the estimation sample and holdout sample, respectively.[9]

None of the models are particularly poor at predicting category choice.[10] The worst predictive accuracy arises from the Independent Model (Model 1), especially for convenience products. The mean absolute difference between the predicted probability and the actual realization is about 5 percent. Including "co-incidence" to allow for co-purchase of multiple products and heterogeneity (Model 2) increases the prediction accuracy. Thus, the approach suggested by Manchanda, Ansari, and Gupta (1999), although developed in studying complementary demands (shopping basket) for frequently purchased products, does indeed capture unobserved factors that cause multiple purchases at the same purchase occasion (e.g., bundling in our application). Comparing Model 2 and Model 3 with Model 4, we find that a multivariate probit model with switching cost and sequential ordering effects provides the most accurate description of households' future purchase decisions. The predictive accuracy of our proposed model is fairly remarkable. For example, the mean absolute error rate of convenience services is less than 0.5 percent. While we do not claim that all applications will achieve such a high

[9] See Jeffreys (1961).

[10] The Models are numbered consistently with Table 1. Thus, Model 1 is the Independent Model and so forth.

degree of accuracy, we believe that the evidence does weigh in favor of including order effects in attempts to formally model choice in this type of environment.

[Insert Figure 2 About Here]

We also checked the robustness of our proposed model to various cut-off points for determining purchase vs. non-purchase prediction. These cut-off points constitute a map between the predicted probabilities generated by the models and a predicted buy/no-buy decision. For example, a cut-off of 0.5 means that we code any predicted probability in the closed interval [0.5,1] as a predicted purchase and any predicted probability below 0.5 as a non-purchase. Figure 2 depicts the hit rate of the holdout sample for each model for a set of cut-off points (0.1, 0.3, 0.5, 0.7). It indicates that all four models have high hit rates for a cut-off point of 0.5, the point that is traditionally used to compute hit rates. However, the Independent Model offers significantly less precise cross-selling prediction when we lower the cut-off points.

[Insert Figure 3 About Here]

Perhaps more importantly, at least from the perspective of managerial applicability, is to explore how each of the models performs in a situation in which a consumer actually made a purchase. As we noted before, our data set contains a great number of non-purchase observations. Following the logic presented earlier, computing the hit rate for the entire holdout sample may cloud the true ability of the models to predict purchases because the hit rates are inflated by the high percentage of non-purchase occasions. To disentangle the idiosyncrasies of our data with the phenomenon we wish to measure, we constructed a holdout sample that is a subset of our original holdout sample. In particular, we partitioned the sample into observations that included at least a single purchase and those that did not. We then used the data, our full estimation sample, to examine the hit rates for the various models for our holdout sample subset that contained purchase observations. It helps to keep in mind that what we are now asking these models to do is a difficult task. We are asking the models, which have been estimated with a large percentage of non-purchase occasions in the data and hence will naturally tend to predict low purchase probabilities, to recognize a situation with a high potential for purchase and, as such, assign a relatively high purchase probability. Figure 3 presents the results of this analysis.

The results of this analysis are more striking. Focusing on the 0.5 cut-off point, we discover that the independent model's performance in predicting actual purchase occasions is miserable. Specifically, out of the 115 purchase occasions in our new holdout sub-sample the Independent Model was able to correctly predict precisely none of them. The Independent Model was not able to overcome its tendency to assign low purchase probabilities even in situations where purchases ultimately did occur. Allowing for correlation in the unobserved utilities and heterogeneity (Model 2) increases the ability of the model to detect purchase at a rate of about 20 percent. Including order effects and switching cost (Model 4) boosts the predictive ability to 30 percent. Allowing the model the flexibility to capture these market realities markedly increases its ability to detect the likelihood of an impending purchase.

4. Discussion, Limitations, and Directions for Future Research

In this paper, we propose a dynamic multivariate probit model to investigate the development of consumer demands for multiple products that are sequentially ordered. This research developed a model useful for predicting product and service acquisition in markets where consumers have sequentially ordered demands. Unlike the market for frequently purchased packaged goods, we believe that there are many products and services that are naturally and temporally ordered, especially in the service and the consumer technology industries. Our model was designed to leverage these reoccurring purchase patterns and in so doing increase the predictive accuracy of our attempts to model product and service choice. We demonstrated our approach on data collected from a large Midwestern bank and found that including these proposed effects significantly improved predictive performance. We expect that there are many other service environments in which including information on natural ordering would yield valuable insights.

Being the first cross-selling model that applies to panel data, some of the obvious advantages of our proposed model is its ability to incorporate periodic consumer decision making and to provide a behavioral explanation for the over-time development of consumer demand for multiple products (from the same provider). Unlike choice models that are independent for each category, or even choice models that take into account correlations in category utilities during joint purchase occasions, our model explicitly allows for the cumulative purchases of other categories to capture the demand maturity and help predict the purchase probabilities of a particular category. Second, our model is general enough to accommodate both same-time complementary demands ("purchase coincidence" as in Manchanda et al.,) and over-time complementary demands ("sequential demand") for multiple products. It also can be applied to both repeatedly purchased product categories and non-repeatedly purchased product categories. Third, the purchase probabilities we estimate are by their nature conditional probabilities, based on the purchases previously made by a particular customer. Fourth, this model allows us to predict which product a consumer is likely to buy next, when s/he is going to buy, and what drives them away from purchasing the next product. In this way, our model provides rich guidance to managers charged with allocating marketing dollars towards customers with the greatest incremental profit potential.

Our banking application was limited and can be expanded in several ways. First, while we were able to obtain relatively detailed information on customer-level account activity, we did not have access to data detailing the marketing activity that each individual in our sample was exposed to over this time period. If such information were available, we could more formally explore the impact of natural ordering on cross-selling tactics and opportunities. This would be a valuable contribution to our knowledge of these markets. Second, our data only covered a time span of one year. In an environment such as banking in which a customer relationship may last for many years, but new service acquisitions are relatively infrequent, one year may not be sufficient to capture the richness of the phenomenon. We expect that data that covers a longer time span would yield even greater differences between our modeling approach and those that do not account for temporal ordering. Also, a longer and wider time span (e.g., yearly) can enable us to better predict when is the best time to cross-sell a household a certain financial product. Third, the model we proposed is a purchase incidence model with no ability to capture the magnitude of the sale. Clearly, a product or service provider is not only interested in whether or not she or he is likely to sell an additional service to an existing customer but also in the amount of money such a sale is likely to generate. In many markets, this may vary widely from customer to customer. Future research should explore how the ordering of purchase incidence and purchase expenditures may inform future purchase behavior.

In summary, the market for services is a huge and growing segment of the U.S. economy. We have seen the benefits of predictive modeling in the arena of packaged goods. The state of our knowledge in this market has grown dramatically over the past decade. Conversely, disproportionately little attention has been paid to the market for services. There is no doubt that this neglect is at least partially explained by the relative difficulty of obtaining quality and timely data from their providers. Data from service providers tends to be more convoluted than data generated by scanner technology. Yet, in spite of these challenges, the benefits of exploring this marketplace are remarkable. We hope that this research is one step in that direction.

Appendix

1. Additional Details for the General Model

1A. Prior Distribution

In order to complete the hierarchical Bayesian model, we need to specify the priors for α_l, η, $\boldsymbol{\mu}$, $\boldsymbol{\omega}$, $\boldsymbol{\Sigma}$, σ_v^2 and Λ_e, Λ_λ. Priors are given below (see the details of the specification in part 2A of the appendix):

$$\alpha_1 \sim MVN(\tau_0, \Omega)$$

$$\eta \sim N(\eta_0, \sigma_\eta^2)$$

$$\boldsymbol{\mu} \sim MVN\,[\boldsymbol{\theta_\mu}, \boldsymbol{\Theta_\mu}]$$

$$\boldsymbol{w} \sim MVN\,[\boldsymbol{\theta_w}, \boldsymbol{\Theta_w}]$$

$$\sigma_v^2 \sim InverseGamma(\gamma, \rho)$$

$$\boldsymbol{\Lambda_e^{-1}} \sim Wishart[\nu_1, (\nu_1 \mathbf{R_1})^{-1}]$$

$$\boldsymbol{\Lambda_\lambda^{-1}} \sim Wishart[\nu_2, (\nu_2 \mathbf{R_2})^{-1}]$$

the off-diagonal elements of $\boldsymbol{\Sigma}$ ~ truncated multivariate normal.

1B. Estimation

In the Bayesian framework, the inference is based on the joint posterior distribution of the unknown parameters. We use the MCMC method (Gibbs sampler) to simulate a sufficient number of random draws from the full conditional distributions (see part 2B of the Appendix) for our model. The sequence of draws generates a Markov chain whose stationary distribution is the joint posterior density of all unknowns. This involves the following sampling algorithm:

a. Generate α_{i0} draws from p($\alpha_{i0}^{(k+1)}$ | $\{\mathbf{U}^k_{it}\}$, $\{Z_{ijt}\}$, $\{\alpha_1^{(k)}\}$, $\boldsymbol{\beta}_{il}^{(k)}$, $\boldsymbol{\gamma}_{il}^{(k)}$, $\boldsymbol{\Sigma}^{(k)}$, $\eta^{(k)}$, $\{D_i\}$, $\sigma_v^{2(k)}$) for i = 1 to I.

b. Generate $\boldsymbol{\beta}_{il}$ draws from p($\boldsymbol{\beta}_{il}^{(k+1)}$ | $\{\alpha_{i0}^{(k+1)}\}$, $\{\mathbf{U}^k_{it}\}$, $\{Z_{ijt}\}$, $\{\alpha_1^{(k)}\}$, $\boldsymbol{\gamma}_{il}^{(k)}$, $\boldsymbol{\Sigma}^{(k)}$, $\boldsymbol{\mu}^{(k)}$, $\{D_i\}$, $\boldsymbol{\Lambda_e}^{(k)}$) for i = 1 to I and l = 1, 2, 3.

c. Generate $\boldsymbol{\gamma}_{il}$ draws from p($\boldsymbol{\gamma}_{il}^{(k+1)}$ | $\{\alpha_{i0}^{(k+1)}\}$, $\{\mathbf{U}^k_{it}\}$, $\{Z_{ijt}\}$, $\{\alpha_1^{(k)}\}$, $\boldsymbol{\beta}_{il}^{(k+1)}$, $\boldsymbol{\Sigma}^{(k)}$, $\boldsymbol{w}^{(k)}$, $\{D_i\}$, $\boldsymbol{\Lambda_\lambda}^{(k)}$) for i = 1 to I and l = 1, 2, …, 7.

d. Generate α_1 draws from p($\alpha_1^{(k+1)}$ | {α_{i0} $^{(k+1)}$},{$\boldsymbol{\beta}_{il}$ $^{(k+1)}$}, {$\boldsymbol{\gamma}_{il}$ $^{(k+1)}$}, {$\mathbf{U}^{k}_{it}$}, {Z_{ijt}}, $\boldsymbol{\Sigma}^{(k)}$, τ_0, Ω).

e. Generate a η draw from p(η $^{(k+1)}$|{$\boldsymbol{\beta}_{il}$ $^{(k+1)}$}, {$\boldsymbol{\gamma}_{il}$ $^{(k+1)}$}, {α_{i0} $^{(k+1)}$}, {$\alpha_1^{(k+1)}$}, {D_i}, $\sigma_v^{2(k)}$, η_0, σ_η^2).

f. Generate a $\boldsymbol{\mu}$ draw from p($\boldsymbol{\mu}^{(k+1)}$|{$\boldsymbol{\beta}_{il}$ $^{(k+1)}$}, {$\boldsymbol{\gamma}_{il}$ $^{(k+1)}$}, {α_{i0} $^{(k+1)}$},{$\alpha_1^{(k+1)}$}, {D_i}, $\boldsymbol{\Lambda}_{\mathbf{e}}^{(k)}$, $\boldsymbol{\theta}_{\mu}$, $\boldsymbol{\Theta}_{\mu}$).

g. Generate a $\boldsymbol{w}$ draw from p($\boldsymbol{w}^{(k+1)}$|{$\boldsymbol{\beta}_{il}$ $^{(k+1)}$}, {$\boldsymbol{\gamma}_{il}$ $^{(k+1)}$}, {α_{i0} $^{(k+1)}$},{$\alpha_1^{(k+1)}$}, {D_i}, $\boldsymbol{\Lambda}_{\boldsymbol{\lambda}}^{(k)}$, $\boldsymbol{\theta}_{w}$, $\boldsymbol{\Theta}_{w}$).

h. Generate a σ_η^2 draw from p($\sigma_\eta^{2\,(k+1)}$|{α_{i0} $^{(k+1)}$}, {D_i}, γ, ρ).

i. Generate a $\boldsymbol{\Lambda}_{\mathbf{e}}^{-1}$ draw from p($\boldsymbol{\Lambda}_{\mathbf{e}}^{-1(k+1)}$|{$\boldsymbol{\beta}_{il}$ $^{(k+1)}$}, {D_i}, $\boldsymbol{\mu}^{(k+1)}$, $\boldsymbol{R_1}$, $\boldsymbol{v_1}$).

j. Generate a $\boldsymbol{\Lambda}_{\boldsymbol{\lambda}}^{-1}$ draw from p($\boldsymbol{\Lambda}_{\boldsymbol{\lambda}}^{-1(k+1)}$|{$\boldsymbol{\gamma}_{il}$ $^{(k+1)}$}, {D_i}, $\boldsymbol{w}^{(k+1)}$, $\boldsymbol{R_2}$, $\boldsymbol{v_2}$).

k. Generate $\mathbf{U}_{it}$ draws from p($\mathbf{U}_{ijt}^{(k+1)}$ | {$\boldsymbol{\beta}_{il}$ $^{(k+1)}$}, {$\boldsymbol{\gamma}_{il}$ $^{(k+1)}$}, {α_{i0} $^{(k+1)}$}, {$\alpha_1^{(k+1)}$}, $\boldsymbol{\Sigma}^{(k)}$, y_{it}) for $i = 1$ to I and $t = 1$ to T_i with $\sum_{i=1}^{I} T_i = T$.

l. Generate a $\boldsymbol{\Sigma}$ matrix from p($\boldsymbol{\Sigma}^{(k+1)}$ | {$\boldsymbol{\beta}_{il}$ $^{(k+1)}$}, {$\boldsymbol{\gamma}_{il}$ $^{(k+1)}$}, {α_{i0} $^{(k+1)}$}, {$\alpha_1^{(k+1)}$}, {$\mathbf{U}_{it}^{(k+1)}$},{y_{it}}, $\boldsymbol{\Sigma}_0$) using a Metropolis-Hastings Hit-and-Run algorithm, where $\boldsymbol{\Sigma}_0 = \mathbf{0}$.

Where k indicates the kth step of the algorithm. The initial draws from the chain are discarded because they reflect a "burn-in" period in which the chain has not converged. A sample of draws obtained after convergence is used to make posterior inferences about model parameters of interest.

2. Additional Details for the Model in Our Application

2A. Prior Specification

The priors are set as diffuse priors. The hyper-parameters in the priors are set as follows:

$\alpha_1 \sim MN(\tau_0, \Omega)$, where $\tau_0 = 0_{(J-1)\times 1}$ and $\Omega = diagonal(1000)$. Recall that α_{13} is set to zero

$\eta \sim N(\eta_0, \sigma_\eta^2)$, where $\eta_0 = 0$ and $\sigma_\eta^2 = 1000$.

$\mu \sim MVN(\theta_\mu, \Theta_\mu)$, where $\theta_\mu = 0, \Theta_\mu = diagonal(1000)$

$\omega \sim MVN(\theta_\omega, \Theta_\varpi)$, where $\theta_\omega = 0, \Theta_\varpi = diagonal(1000)$

$\sigma_v^2 \sim InverseGamma(\gamma, \rho)$, where $\gamma = 0.001, \rho = 0.001$

$\Lambda_e^{-1} \sim Wishart(v_1, (v_1 R_1)^{-1})$, where $v_1 =$ number of categories $+1$, $v_1 = 4$ in this application and $\mathrm{R}_1 = diagonal(0.001)$

$\Lambda_\lambda^{-1} \sim Wishart(v_2, (v_2 R_2)^{-1})$, where $v_2 =$ number of categories $+1$, $v_2 = 4$ in this application and $\mathrm{R}_2 = diagonal(0.001)$

$f(vec^*(\Sigma) \mid \Sigma_0) \propto \exp\{-\frac{1}{2}(vec^*(\Sigma) - vec^*(\Sigma_0))`(vec^*(\Sigma) - vec^*(\Sigma_0))\}$, where $vec^*(\Sigma)$ is a $(\mathrm{J}(\mathrm{J}-1)/2)\times 1$ vector of free elements of Σ. $\Sigma_0 = 0$ which is a $\mathrm{J}\times\mathrm{J}$ matrix of zeros.

2B. Full Conditional Distributions and Estimation

The full conditional distribution for $\alpha_{i0} \sim N(\overline{\alpha_{i0}}, \sigma_\alpha^{2*})$

where $\overline{\alpha_{i0}} = (\sum_{t=1}^{T_i} U_{ijt}^* \big/ \Sigma_{11} + D_i \eta / \sigma_v^2) \Big/ (T_i / \Sigma_{11} + 1/\sigma_v^2)$, $\Sigma_{11} = 1$,

$U_{ijt}^* = U_{ijt} - \alpha_{1j} - C_{ij(t-1)}`\beta_{ji} - X_{ijt}`\gamma_{ij}$,

$\sigma_\alpha^{2*} = 1/(T_i / \Sigma_{11} + 1/\sigma_v^2)$, Σ_{11} is the element in the first row and first column of Σ which is 1, for $\mathrm{i} = 1, 2, ..., \mathrm{I}$.

The full conditional distribution for $\beta_{il} \sim MVN(\overline{\beta_{il}}, V_{il}^{-1})$

where $\overline{\beta_{il}} = V_{il}^{-1}(\sum_{t=1}^{T_i} C'_{ilt}\Sigma^{-1}U_{it}^* + \Lambda_{el}{}^{-1}D_i\mu_l)$, $\mathrm{V}_{il} = \sum_{t=1}^{T_i} C'_{ilt}\Sigma^{-1}C_{ilt} + \Lambda_{el}{}^{-1}$,

and $U_{it}^* = U_{it} - \alpha_0 - \alpha_1 - \overline{C}_{i(t-1)}`\beta_i^* - X_{it}`\gamma_i$, where $\overline{C}_{i(t-1)}$ excludes C_{ilt} and β_i^* excludes β_{il} for $\mathrm{i} = 1, ..., \mathrm{I}$ and $\mathrm{l} = 1, 2, 3$.

The full conditional distribution for $\gamma_{il} \sim MVN(\overline{\gamma_{il}}, V'^{-1}_{il})$

where $\overline{\gamma_{il}} = V'^{-1}_{il}(\sum_{t=1}^{T_i} X'_{ilt}\Sigma^{-1}U_{it}^* + \Lambda_{\lambda l}{}^{-1}D_i\omega_l)$, $\mathrm{V'}_{il} = \sum_{t=1}^{T_i} X'_{ilt}\Sigma^{-1}X_{ilt} + \Lambda_{\lambda l}{}^{-1}$,

and $U_{it}^* = U_{it} - \alpha_0 - \alpha_1 - C_{i(t-1)}`\beta_i - \overline{X}_{it}`\gamma_i^*$, where $\overline{X}_{it}$ excludes X_{ilt} and γ_i^* excludes γ_{il} for $\mathrm{i} = 1, ..., \mathrm{I}$ and $\mathrm{l} = 1, 2, ..., 7$.

The full conditional distribution for $\alpha_{1(J-1)\times(J-1)} \sim MVN(\overline{\alpha}, P^{-1})$

$$\text{where } \overline{\alpha} = P^{-1}(\Omega^{-1}\tau_0 + \sum_{i=1}^{I}\sum_{t=1}^{T_i}\Sigma^{-1}U_{it}^{*}), U_{it}^{*} = U_{it} - \alpha_0 - C_{i(t-1)}`\beta_i - X_{it}`\gamma_i,$$

$$\text{and P} = \sum_{i=1}^{I}\sum_{t=1}^{T_i}\Sigma^{-1} + \Omega^{-1}$$

The full conditional distribution for $\eta_{\mathrm{k}} \sim N(\overline{\eta}, \Theta^{*})$

$$\text{where } \overline{\eta} = (\sum_{i=1}^{I} D_{ik}\alpha_{i0}^{*} \Big/ \sigma_v^2 + \eta_0 \big/ \sigma_\eta^2) \Big/ (\sum_{i=1}^{I} D_{ik}^2 \Big/ \sigma_v^2 + 1/\sigma_\eta^2),$$

$$\Theta^{*} = 1 \Big/ (\sum_{i=1}^{I} D_{ik}^2 \Big/ \sigma_v^2 + 1/\sigma_\eta^2), \alpha_{i0}^{*} = \alpha_{i0} - \eta_1 D_{i1} - \eta_2 D_{i2} \text{ if k} = 0,$$

$$\alpha_{i0}^{*} = \alpha_{i0} - \eta_0 - \eta_2 D_{i2} \text{ if k} = 1, \alpha_{i0}^{*} = \alpha_{i0} - \eta_0 - \eta_1 D_{i1} \text{ if k} = 2.$$

The full conditional distribution for $\mu_{\mathrm{sk}} \sim MVN(\overline{\mu_{\mathrm{sk}}}, M_0^{-1})$

$$\text{where } \overline{\mu_{\mathrm{sk}}} = M_0^{-1}(\sum_{i=1}^{I} D'_{is}\Lambda_e^{-1}\beta_{ik}^{*} + \Theta_\mu^{-1}\theta_\mu),$$

$$M_0 = \sum_{i=1}^{I} D'_{is}\Lambda_e^{-1}D_{is} + \Theta_\mu^{-1}, \beta_{ik}^{*} = \beta_{ik} - \mu_1 D_{i1} - \mu_2 D_{i2} \text{ if s} = 0,$$

$$\beta_{ik}^{*} = \beta_{ik} - \mu_0 - \mu_2 D_{i2} \text{ if s} = 1, \beta_{ik}^{*} = \beta_{ik} - \mu_0 - \mu_1 D_{i1} \text{ if s} = 2,$$

$$\text{for k} = 1, 2, 3.$$

The full conditional distribution for $\omega_{\mathrm{sk}} \sim MVN(\overline{\omega_{\mathrm{sk}}}, M_1^{-1})$

$$\text{where } \overline{\omega_{\mathrm{sk}}} = M_1^{-1}(\sum_{i=1}^{I} D'_{is}\Lambda_\lambda^{-1}\gamma_{ik}^{*} + \Theta_\omega^{-1}\theta_\omega),$$

$$M_1 = \sum_{i=1}^{I} D'_{is}\Lambda_\lambda^{-1}D_{is} + \Theta_\omega^{-1}, \gamma_{ik}^{*} = \gamma_{ik} - \omega_1 D_{i1} - \omega_2 D_{i2} \text{ if s} = 0,$$

$$\gamma_{ik}^{*} = \gamma_{ik} - \omega_0 - \omega_2 D_{i2} \text{ if s} = 1, \gamma_{ik}^{*} = \gamma_{ik} - \omega_0 - \omega_1 D_{i1} \text{ if s} = 2,$$

$$\text{for k} = 1, 2, \ldots, 7.$$

The full conditional distribution for $\sigma_v^2 \sim InverseGamma(\overline{\gamma}, \overline{\rho})$

$$\text{where } \overline{\gamma} = \gamma + I/2,$$

$$\overline{\rho} = \frac{1}{2}\sum_{i=1}^{I}(\alpha_{i0} - \eta' D_i)^2 + \rho$$

The full conditional distribution for $\Lambda_{ek}^{-1} \sim Wishart(\nu_1 + I, W_1^{-1})$

$$\text{where } W_1 = \sum_{i=1}^{I} (\beta_{ik} - \mu_k ' D_i)(\beta_{ik} - \mu_k ' D_i)' + v_1 R_1, \text{ for k} = 1,2,3.$$

The full conditional distribution for $\Lambda_{\lambda k}^{-1} \sim Wishart(\nu_2 + I, W_2^{-1})$

$$\text{where } W_2 = \sum_{i=1}^{I} (\gamma_{ik} - \omega_k ' D_i)(\gamma_{ik} - \omega_k ' D_i)' + v_2 R_2, \quad \text{for k} = 1,2,...,7.$$

The full conditional distribution for $\mathrm{U_{it}} \sim \text{Truncated}\, MVN(Z_{it}`\theta_i, \Sigma)$

The truncation region is determined by the purchase incidences across categories. A mini-Gibbs sampler (Geweke 1991) is used to generate the truncated multivariate normal draws. That is, J Gibbs sampling steps are applied such that $\mathrm{U_{ijt}}$ is generated from its corresponding univariate truncated conditional normal density. This density is truncated from below 0 if a purchase in category j is made and otherwise is truncated from above by 0. An standard inverse CDF method is used to generate the univariate truncated conditional normal densities.

A Metropolis Hit-and-Run algorithm (Manchanda, Ansari and Gupta 1999, Dey and Chen 1996, Chen and Schmeiser 1993) is used to generate Σ. A candidate matrix Σ^c is generated by using a random walk chain such that $\Sigma^c = \Sigma + \Psi$, where Σ is the current value of the correlation matrix, and Ψ is a matrix such that $\Psi_{ii} = 0$ and $E(\Psi_{ij}) = 0.$ The algorithm involves the following steps:

a. Generate J(J-1)/2 i.i.d standard normal deviates, $q_{12,} q_{13,...,} q_{(J-1)J}$,
b. Generate a deviate s from $N(0, \sigma_s^2)$ which is truncated to the interval $(-\phi/\sqrt{2}, \phi/\sqrt{2})$, where ϕ is the smallest eigen value of Σ, σ_s^2 is a tuning constant to ensure that candidate matrix Σ^c is not rejected disproportionately.
c. Compute

$$\Psi_{ij} = \frac{sq_{ij}}{(\sum_{j=1}^{J-1} \sum_{k=j+1}^{J} q_{jk}^2)}$$

for i < j, $\Psi_{ii} = 0$, and $\Psi_{ij} = \Psi_{ji}$ for i > j.

The Metroposlis-Hastings acceptance probability is:

$$\min\left\{1, \frac{L(\Sigma^c \mid \beta, O, \{U_{it}\}, y)\, prior(\Sigma^c)(\Phi(\frac{\phi^C}{\sqrt{2}\sigma_s}) - \Phi(\frac{-\phi^C}{\sqrt{2}\sigma_s}))}{L(\Sigma \mid \beta, O, \{U_{it}\}, y)\, prior(\Sigma)(\Phi(\frac{\phi}{\sqrt{2}\sigma_s}) - \Phi(\frac{-\phi}{\sqrt{2}\sigma_s}))}\right\}$$

where L(.) is the likelihood function and $\Phi(.)$ is the standard normal cumulative distribution.

Table 1: The Operationalization of The Instruments

Variables	Definitions	Mean or Freq
AGE1	1 if the account owner's age is less than 19, 0 otherwise	0.8%
AGE2	1 if the account owner's age is between 19 and 33, 0 otherwise	4.2%
AGE3	1 if the account owner's age is between 34 and 60, 0 otherwise	24.1%
AGE4	1 if the account owner's age is above 60, 0 otherwise	70.9%
INCOME	Household's income (1-7) where 1 -<$15,000, 2 - $15,000-$24,999, 3 - $25,000-$34,999, 4 - $35,000 - $49,999, 5 - $50,000 - $74,999, 6 - $75,000 - $99,999, 7 - $100,000+	3.34
COMPET	1 if the household opens some accounts in another bank during the last six months, 0 otherwise	13.3%
OVERSAT	Household's overall satisfaction (1 to 7)	4.33
SWIT	1 if the account owner's profession is white collar and the household has at least one non-adult child and the household owns more than average number of accounts with this bank, 0 otherwise	6.6%
EDUCAT	The household head's education measured on a 1-5 scale (1 = some high school , …, 5 = post graduate	1.21
GENDER	1 if the account owner is a male, 0 otherwise	82.5%

Table 2: Categorization of Financial Services

Categories	Financial Services	Total Purchases During Observation Period
Category 1: Convenience Services	All Checking Accounts, All Saving Accounts, ATM Cards, Debit Cards, All loans and Mortgage, Bank by Phone	381
Category 2: Cash Reserve	IRA's, Money Market Accounts, CD's, Life Insurance, Pension Plan	143
Category 3: Advanced Financial Services	Time Deposits, Annuities, Corporate Stocks, Cash Management Account, Mutual Funds, Travel/Entertainment Card (Luxury), Tax Shelters, Corporate/Government Bonds, Real Estate other than Home	17

Table 3. Estimation Results

	Independent Models			Correlation/No Order			Correlation/SWIT			Correlation/Order		
	Y1	Y2	Y3	Y1	Y2	Y3	Y1	Y2	Y3	Y1	Y2	Y3
Intercept	-2.06	-2.42	-3.44	-6.07	-6.45	-7.84	-1.75	-1.75	-1.75	-1.92	-1.92	-1.92
	(7.27)	(11.9)	(663)	(0.09)	(0.10)	(0.05)	(0.05)	(0.05)	(0.05)	(0.04)	(0.04)	(0.04)
CUMY1	0.23	0.07	0.16							1.82	1.66	3.29
	(0.02)	(0.04)	(0.08)							(0.15)	(0.06)	(0.04)
CUMY2	-0.05	0.18	-0.002							-1.78	1.29	3.10
	(0.03)	(0.02)	(0.11)							(0.09)	(0.08)	(0.08)
CUMY3	-0.05	-0.10	0.16							-0.16	-1.17	0.80
	(0.03)	(0.08)	(0.03)							(0.05)	(0.07)	(0.07)
AGE1	-0.19	-0.19	-0.18	-0.99	-0.40	-0.37	-0.23	-0.15	0.09	-0.16	-0.19	0.37
	(165)	(271)	(3216)	(0.06)	(0.09)	(0.12)	(0.14)	(0.08)	(0.13)	(0.07)	(0.13)	(0.05)
AGE2	0.09	-0.05	-0.82	-0.05	-0.04	0.87	0.33	-0.65	-0.57	-0.84	1.04	0.78
	(0.02)	(0.05)	(3110)	(0.05)	(0.07)	(0.08)	(0.13)	(0.05)	(0.07)	(0.12)	(0.04)	(0.09)
AGE3	0.09	-0.07	0.04	-2.93	-2.16	-0.14	-2.43	-3.44	-3.39	-3.54	-3.33	-3.23
	(0.03)	(0.04)	(0.10)	(0.04)	(0.05)	(0.06)	(0.07)	(0.06)	(0.04)	(0.07)	(0.05)	(0.09)
INCOME	0.10	-0.12	0.08	0.33	-0.23	-0.18	-0.87	-0.46	-0.90	-0.18	-0.25	0.08
	(0.03)	(0.04)	(0.11)	(0.05)	(0.10)	(0.07)	(0.13)	(0.09)	(0.08)	(0.09)	(0.11)	(0.04)
COMPET	-0.08	0.07	0.03	0.13	0.51	0.68	-2.87	-3.28	-3.20	-2.89	-1.93	-0.79
	(0.03)	(0.03)	(0.09)	(0.15)	(0.05)	(0.07)	(0.07)	(0.09)	(0.06)	(0.04)	(0.04)	(0.06)
OVERSAT	-0.06	-0.09	0.07	0.28	0.10	0.73	0.12	-0.06	0.11	0.48	0.34	0.64
	(0.03)	(0.03)	(0.11)	(0.09)	(0.07)	(0.04)	(0.06)	(0.09)	(0.09)	(0.06)	(0.09)	(0.09)
SWIT	0.08	-0.09	-0.13				1.93	1.19	1.20	0.98	0.69	1.03
	(0.02)	(0.04)	(0.11)				(0.15)	(0.09)	(0.08)	(0.07)	(0.04)	(0.16)
α_1							-2.5	-1.34	0	-4.3	-3.5	0
							(0.41)	(0.36)		(0.17)	(0.64)	
Log-Marginal Density	-1244	-586	-76.4		-64.9			-60.1			-53.2	

* The numbers in the parentheses for the independent models in the table are standard errors, and those in the parentheses for the other models are posterior standard deviations.

Table 4: Unobserved Correlation Across Three Categories

	Proposed Model		
	Category 1	Category 2	Category 3
Category 1	1	0.18	0.13
		(0.04)	(0.03)
Category 2		1	0.03
			(0.04)
Category 3			1

Table 5: Estimation Results for Heterogeneity Equations

Heterogeneity Equation - α_{i0}

	Intercept	Education	Gender
Estimates	-0.51 (0.05)	-0.51 (0.02)	-0.43 (0.02)

Heterogeneity Equation - β_{1i} : coefficient of $\sum_{l}^{t-1} y_{i,1,l}$

	Intercept	Education	Gender
β_{1i1}	1.19 (0.07)	0.15 (0.03)	0.24 (0.03)
β_{1i2}	1.16 (0.04)	0.07 (0.02)	0.23 (0.02)
β_{1i3}	1.93 (0.05)	0.53 (0.04)	0.39 (0.02)

Heterogeneity Equation - β_{2i} : coefficient of $\sum_{l}^{t-1} y_{i,2,l}$

	Intercept	Education	Gender
β_{2i1}	-1.34 (0.04)	-0.28 (0.03)	-0.06 (0.02)
β_{2i2}	0.91 (0.05)	-0.05 (0.02)	0.24 (0.02)
β_{2i3}	1.16 (0.05)	0.78 (0.02)	0.55 (0.03)

Heterogeneity Equation - β_{3i} : coefficient of $\sum_{l}^{t-1} y_{i,3,l}$

	Intercept	Education	Gender
β_{3i1}	0.48 (0.03)	-0.41 (0.01)	-0.08 (0.01)
β_{3i2}	-1.04 (0.04)	-0.07 (0.01)	-0.03* (0.02)
β_{3i3}	-0.12 (0.04)	0.34 (0.01)	0.28 (0.02)

Heterogeneity Equation - γ_{i1} : coefficient of AGE1

	Intercept	Education	Gender
γ_{1i1}	-0.53 (0.04)	-0.05* (0.03)	0.23 (0.02)
γ_{1i2}	-0.32 (0.06)	-0.08 (0.03)	0.12 (0.02)

γ_{1i3}	-0.16 (0.04)	-0.08 (0.03)	0.34 (0.03)

Heterogeneity Equation - γ_{i2} : coefficient of AGE2

	Intercept	Education	Gender
γ_{2i1}	-0.96 (0.07)	0.03* (0.02)	0.05 (0.02)
γ_{2i2}	0.06 (0.04)	-0.07 (0.03)	0.58 (0.01)
γ_{2i3}	-0.37 (0.07)	0.36 (0.03)	0.39 (0.01)

Heterogeneity Equation - γ_{i3} : coefficient of AGE3

	Intercept	Education	Gender
γ_{3i1}	-1.81 (0.03)	-0.63 (0.01)	-0.53 (0.02)
γ_{3i2}	-1.67 (0.03)	-0.69 (0.01)	-0.46 (0.01)
γ_{3i3}	-1.81 (0.04)	-0.43 (0.01)	-0.49 (0.02)

Heterogeneity Equation - γ_{i4} : coefficient of INCOME

	Intercept	Education	Gender
γ_{4i1}	-0.72 (0.04)	0.24 (0.03)	0.14 (0.02)
γ_{4i2}	-0.59 (0.06)	0.13 (0.03)	0.11 (0.02)
γ_{4i3}	-0.50 (0.03)	-0.01* (0.02)	0.33 (0.01)

Heterogeneity Equation - γ_{i5} : coefficient of COMPET

	Intercept	Education	Gender
γ_{5i1}	-1.57 (0.03)	-0.66 (0.02)	-0.29 (0.01)
γ_{5i2}	-1.24 (0.02)	-0.38 (0.03)	-0.13 (0.01)
γ_{5i3}	-0.90 (0.03)	-0.001* (0.03)	0.06 (0.02)

Heterogeneity Equation - γ_{i6} : coefficient of OVERSAT

	Intercept	Education	Gender
γ_{6i1}	-0.85 (0.06)	0.01* (0.02)	-0.21 (0.01)
γ_{6i2}	-0.43	0.07	-0.37

	(0.06)	(0.01)	(0.02)
γ_{6i3}	-0.84 (0.04)	0.12 (0.02)	-0.19 (0.02)

Heterogeneity Equation - γ_{i7} : coefficient of SWIT

	Intercept	Education	Gender
γ_{7i1}	0.99 (0.06)	0.15 (0.02)	-0.09 (0.02)
γ_{7i2}	0.53 (0.03)	0.22 (0.01)	-0.05 (0.02)
γ_{7i3}	0.96 (0.07)	0.17 (0.03)	-0.07 (0.03)

* Parameters with t statistics lower than 1.96.

Table 6: Selected 10 Households for Cross-Selling Efforts

Household	t	Actual Data			Predicted Probabilities		
		Category1	Category2	Category3	Category1	Category2	Category3
1	1	0* (6)	0 (0)	0 (0)	0.0736	0.0069	0.0017
	2	0 (6)	0 (0)	0 (0)	0.0501	0.0105	0.0013
2	1	0 (1)	0 (0)	0 (0)	0.0502	0.0105	0.0013
	2	0 (1)	0 (0)	0 (0)	0.0172	0.0066	0.0003
3	1	1 (2)	0 (0)	0 (0)	0.5287	0.0185	0.0003
	2	0 (3)	0 (0)	0 (0)	0.2311	0.0249	0.0012
4	1	0 (8)	0 (4)	0 (0)	0.2926	0.0141	0.0026
	2	1 (8)	0 (4)	0 (0)	0.5788	0.0170	0.0038
5	1	1 (3)	0 (4)	0 (0)	0.7227	0.0142	0.0006
	2	1 (4)	0 (4)	0 (0)	0.7358	0.0254	0.0041
6	1	1 (3)	0 (9)	0 (0)	0.6125	0.0335	0.0007
	2	0 (4)	0 (9)	0 (0)	0.1251	0.0639	0.0041
7	1	0 (6)	0 (9)	0 (0)	0.0181	0.0122	0.0036
	2	1 (6)	1 (9)	1 (0)	0.3981	0.3094	0.3986
8	1	0 (2)	0 (0)	0 (0)	0.0762	0.0045	0.0005
	2	1 (2)	0 (0)	0 (0)	0.6955	0.0218	0.0006
9	1	1 (5)	0 (3)	0 (1)	0.3003	0.0098	0.0038
	2	0 (6)	0 (3)	0 (1)	0.1271	0.0167	0.0089
10	1	1 (5)	0 (3)	1 (2)	0.3282	0.0068	0.3448
	2	0 (6)	0 (3)	0 (3)	0.0960	0.0079	0.0332

* The number 0 or 1 indicates whether the household purchases (1) some service in the corresponding category or not (0) during the sample period. The numbers in the parentheses are the numbers of accounts that the household opened in the corresponding category with the bank before the current period.

Figure 1a: Comparison of Mean Absolute Error Across Models in Estimation Sample

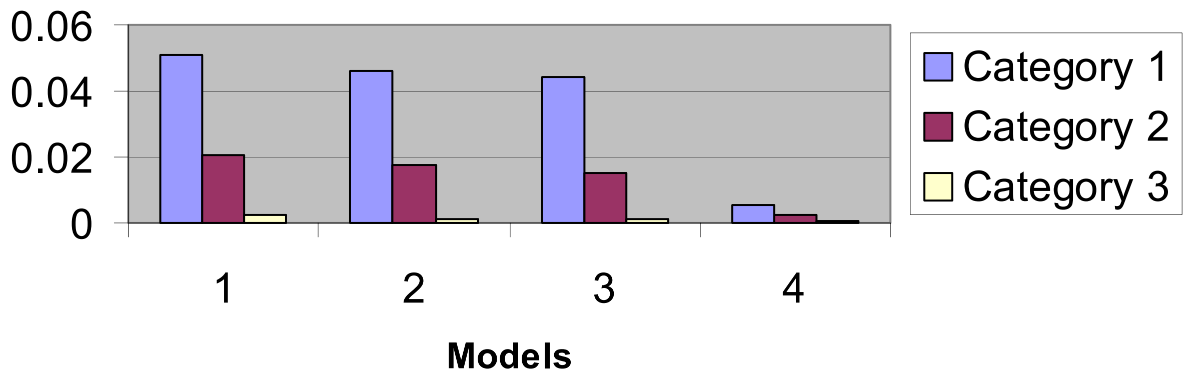


Figure 1b: Comparison of Mean Absolute Error Across Models in Holdout Sample

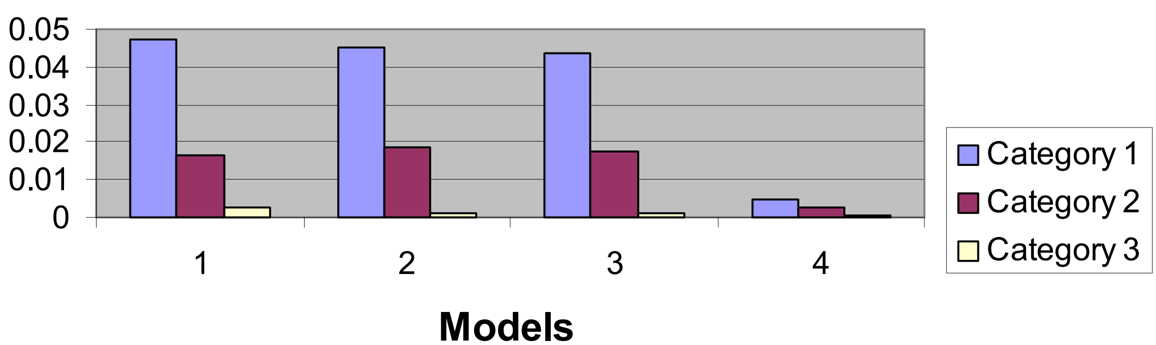


Figure 2: Comparison of Hit Rate 1 Across Models in Holdout Sample

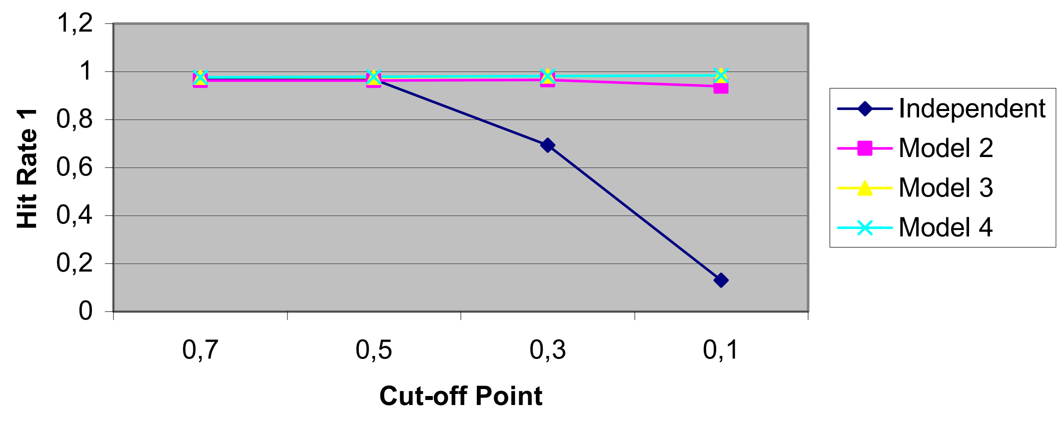
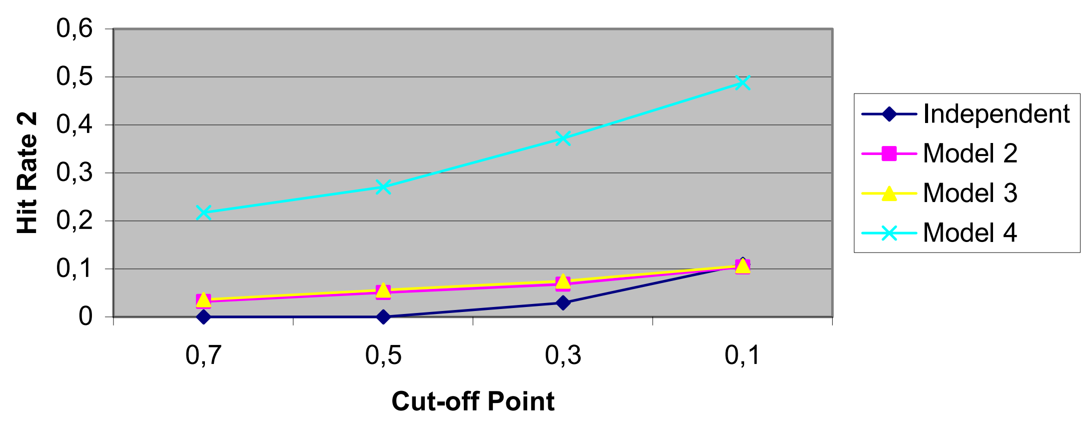


Figure 3: Comparison of Hit Rate 2 for Purchasers Across Models in Estimation Sample



References

Ainslie A., and P. Rossi (1998), "Similarities in Choice Behavior Across Product Categories," *Marketing Science*, 17(2): 91-106.

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

Barber B. and T. Odean (2001), "Boys will be Boys: Gender, Overconfidence, and Common Stock Investment," *Quarterly Journal of Ecomomics*, February 2001, Vol. 116, No. 1, 261-292

Becker, G. (1965), "A Theory of the Allocation of Time," *Economic Journal*, 75: 493-517.

Blattberg, R., T. Buesing, P. Peacock, and S. Sen (1978), "Identifying the Deal Prone Segment," *Journal of Marketing Research*, 15, 3: 369-377.

Blattberg, R., P. Peacock, and S. Sen (1976), "Purchasing Strategies Across Product Categories," *Journal of Consumer Research*, 3: 143-154.

Chiang J., and R. Wilcox (1998), "A Cross-Category Analysis of Shelf-Space Allocation, Product Variety, and Retail Margins," *Marketing Letters*, 8(2): 183-191.

Clarke, Y. and G. N. Soutar (1982), "Consumer Acquisition Patterns for Consumer Durables: Some Australian Evidence," *Journal of Consumer Research*, 8: 456-460.

Cunningham, R. (1956), "Brand Loyalty—What, Where, and How Much," *Harvard Business Review*, 34, January–February: 116-128.

Dickson, P. R., R. F. Lusch and W, L. Wilkie (1983), "Consumer Acquisition Priorities for Home Appliances: A Replication and Reevaluation," *Journal of Consumer Research*, 9(1): 432-435.

Erdem and Sun (2000), "An Empirical Investigation of Spillover Effects of Marketing Mix Strategy in Umbrella Branding," forthcoming, *Journal of Marketing Research*.

Fader, P., and L. Lodish (1990), "A Cross-Category Analysis of Category Structure and Promotional Activity for Grocery Products," *Journal of Marketing*, October: 52-65.

Gelfand, A E., and A.F. M. Smith (1990), "Sampling-Based Approaches to Calculating Marginal Densities," *Journal of the American Statistical Association*, 85, 410: 398-409.

Gonul, Fusun, and Kannan Srinivasan (1996), "Estimating the Impact of Consumer Expectation of Coupons on Purchase Behavior: A Dynamic Structural Model," *Marketing Science*, 15, 3: 262-279.

Hauser, J.R. and G. L. Urban (1986), "The Value Priority Hypotheses for Household Budget Plans," *Journal of Consumer Research*, 12: 446-462.

Hebden, J.J., and J. F. Pickering, "Patterns of Acquisition of Household Durables," *Oxford Bulletin of Economics and Statistics*, (1974): 67-94.

Heckman, James J., "Dummy Endogenous Variables in a Simultaneous Equation System," *Econometrica*, 46, 4 (1978): 931-959.

Heckman, James J., "The Incidental Parameters Problem and The Problem of Initial Conditions in Estimating a Discrete Time-Discrete Data Stochastic Process," in *Structural Analysis of Discrete Data with Econometric Applications*, edited by C. F. Manski and D. McFadden, (Cambridge: MIT Press 1981) 179-195.

Information Week, "Outlook for 2001," January 8, 2001.

Jeffries H., *Theory of Probability* (3^{rd} ed.) (Oxford, U.K., Oxford University Press1961).

Kamkura, Wagner A., and Gary Russel, "A Probabilistic Choice Model For Market Segmentation and Elasticity," *Journal of Marketing Research*, 26, 4, 1989: 379-391.

Kamakura W., S. Ramaswami, and R. Srivastava (1991), "Applying Latent Trait Analysis in the Evaluation of Prospects for Cross-Selling of Financial Services," *International Journal of Research in Marketing*, 8: 329-349.

Kamakura, Wagner A. and Bruce S. Kossar (2001), "Identifying Innovators for the Cross-selling of New Products," working paper, Duke University.

Kass, R., and A. Raferty (1995), "Bayes Factors," *Journal of the American Statistical Association*, 90, 430: 773-795.

Kasulis, J.J., R. F. Lusch and E. F. Stafford (1979), "Consumer Acquisition Patterns for Durable Goods," *Journal of Consumer Research*, 6(2): 47-57.

Kim, B., K. Srinivasan, and R. Wilcox (1999), "Identifying Price Sensitive Consumers: The Relative Merits of Demographic vs. Purchase Pattern Information," *Journal of Retailing*, 75(2): 173-193.

Knott Aaron, Andrew Hayes and Scott A. Neslin (2002), "Next-Product-to-Buy Models for Cross-Selling Applications," working paper, Dartmouth College.

Krishna, Aradhna, Imran S. Currim and Robert W. Showmaker (1991), "Consumer Perceptions of Promotional Activity," *Journal of Marketing*, 55, 2: 4-17.

Lum S., and B. Moyer, “Gross Domestic Product by Industry for 1997-99,” *Survey of Current Business*, December 2000: 24-35.

Manchanda, P., A. Ansari, and S. Gupta (1999), “The ‘Shopping Basket’: A Model for Multicategory Purchase Incidence Decisions,” *Marketing Science*, 18: 95-114.

Massy, W., R. Frank, and T. Lodahl, *Purchasing Behavior and Personal Attributes* (Philadelphia, University of Pennsylvania Press, 1968).

Mayo, M. C. and W. J. Qualls (1987), “Household Durables Goods Acquisition Patterns: A Longitudinal Study,” P. Anderson and M. Wallendorf (eds.), *Advances in Consumer Research*. (Ann Arabor. MI: Association for Consumer Research), 14.

McFall, John (1969), “Priority Patterns and Household Behavior,” *Journal of Marketing*: 50-55.

Narasimhan, C., S. Neslin, and S. Sen (1996), “Promotional Elasticities and Category Characteristics,” *Journal of Marketing*, 60: 17-30.

Network World, “Spending on the Rise,” January 29, 2001.

Paroush, Jacob (1965), “The Order of Acquisition of Household Durables,” *Econometrica*: 225-235.

Pyatt, F. Graham, “Priority Patterns and The Demand for Household Durable Goods,” (Cambridge, MA: Cambridge University Press, 1964).

Raju, J. (1992), “The Effect of Price Promotions on Variability in Product Category Sales,” *Marketing Science*, 11(3): 207-220.

Rossi, Peter E., Robert E. McCulloch and Greg M. Allenby (1996), “The Value of Purchase History Data in Marketing,” *Marketing Science*, 15, 4: 321-340.

Wall Street Journal, “ Citigroup is Ready to Realize the Benefits of Cross-Selling,” New York, March 15, 1999.

Wind Y., and R. Frank (1969), “Interproduct Household Loyalty to Brands,” *Journal of Marketing Research*, 6: 434-435.