Abstract: Customer base analysis is an essential tool to measure and develop relationships with customers. While various models have been proposed in a noncontractual setting, they focus primarily on analyzing transactional patterns associated with a single product category or a firm-level activity, such as the times at which purchases are made at a particular retailer. This research proposes a modeling framework for customer base analysis in a multi-category context. Specifically, we model the time between a customer's purchases at the firm and the product categories that comprise her shopping basket arising from multi-category choice decisions. The proposed model uses a latent space approach that parsimoniously captures the dynamics of multi-category shopping behavior due to the interplay between purchase timing and shopping basket composition. We also account for interdependence among multiple categories, temporal dependence across category choices, and latent customer attrition. Using category-level transaction data, we show that the proposed model offers excellent fit and performance in predicting customer purchase patterns across multiple categories. The forecasts and inferences afforded by our model can assist managers in tailoring marketing efforts across categories.
A Multi-category Customer Base Analysis

================================================================================

ARTICLE INFO

Article history:
First received in April 16, 2013 and was under review for 5 months.

Area Editor: Olivier Toubia

================================================================================
A Multi-category Customer Base Analysis

Highlights

- We develop a modeling framework for customer base analysis in a multi-category context.
- The model predicts customer purchase patterns across multiple product categories.
- We allow for the association between the purchase timing and shopping basket composition.
- The model can be easily scaled to a large number of categories through the latent space approach.
A Multi-category Customer Base Analysis

Chang Hee Park, Young-Hoon Park, and David A. Schweidel

December 2013

*Chang Hee Park is Assistant Professor of Marketing at the School of Management, the State University of New York at Binghamton, P.O. Box 6000, Binghamton, NY 13902 (Phone: 607-777-6068, Email: cpark@binghamton.edu). Young-Hoon Park is AMOREPACIFIC Professor of Management and Associate Professor of Marketing at the Samuel Curtis Johnson Graduate School of Management, Cornell University, 361 Sage Hall, Ithaca, NY 14853 (Phone: 607-255-3217, Email: yp34@cornell.edu). David A. Schweidel is Associate Professor of Marketing at the Goizueta Business School, Emory University, GA 30322 (Phone: 404-727-5275, Email: dschweidel@emory.edu). The authors would like to thank the company, which wishes to remain anonymous, that provided the data used in this study. This paper stems from the first author’s dissertation work; he extends special thanks to his dissertation committee (Vishal Gaur, Sachin Gupta, and Vithala Rao) for their feedback and encouragement. All correspondence should be addressed to the first author.
A Multi-category Customer Base Analysis

Abstract

Customer base analysis is an essential tool to measure and develop relationships with customers. While various models have been proposed in a noncontractual setting, they focus primarily on analyzing transactional patterns associated with a single product category or a firm-level activity, such as the times at which purchases are made at a particular retailer. This research proposes a modeling framework for customer base analysis in a multi-category context. Specifically, we model the time between a customer’s purchases at the firm and the product categories that comprise her shopping basket arising from multi-category choice decisions. The proposed model uses a latent space approach that parsimoniously captures the dynamics of multi-category shopping behavior due to the interplay between purchase timing and shopping basket composition. We also account for interdependence among multiple categories, temporal dependence across category choices, and latent customer attrition. Using category-level transaction data, we show that the proposed model offers excellent fit and performance in predicting customer purchase patterns across multiple categories. The forecasts and inferences afforded by our model can assist managers in tailoring marketing efforts across categories.

Keywords: Customer base analysis, Timing models, Multi-category choice models, Latent space models, Bayesian estimation
1 Introduction

Customer base analysis is an essential tool to manage and develop relationships with customers by predicting their future purchase patterns (Schmittlein and Peterson 1994). The uses of the analyses range from aggregate-level sales trajectories to individual-level customer valuation (Fader and Hardie 2009). The “buy ‘til you die” framework is a widely adopted modeling approach for customer base analysis in a noncontractual setting. This class of models are based on the assumption that customers repeatedly make transactions at a firm while they are in an active state (alive), and at some unobserved time point they become permanently inactive (dead). Though many variants of “buy ‘til you die” models have been proposed (e.g., Schmittlein, Morrison, and Colombo 1987; Fader, Hardie, and Lee 2005; Fader, Hardie and Shang 2010), they primarily focus on analyzing transactional patterns associated with a single product category or a firm-level activity, such as the times at which purchases are made at a particular retailer.

Oftentimes, however, customers’ shopping behavior involves the purchase of multiple product categories at a single transaction. In such cases, marketers can gain a more comprehensive understanding of its customer base by analyzing customers’ purchase patterns within individual categories. The inferences would be of particular importance to firms making marketing and operational decisions at the category level. For example, the analyses could be used for customer targeting and scoring (e.g., Reinartz and Kumar 2003; Schweidel and Knox 2013) on a category-by-category basis. If a customer is predicted to be more actively purchasing a specific category than others, sending the customer marketing offers relevant to the category could more effectively facilitate the realization of the future transactions within this category. Such category-level marketing efforts could be further improved by recognizing the relationships that may exist across categories, both during the current shopping trip (i.e., what categories tend to be co-purchased) and across shopping trips (i.e., are category purchases on the previous trip informative of category
purchases on the current trip). For example, cross-category dependencies may suggest promoting some product categories to customers to increase their likelihood of making purchases in multiple categories, thereby increasing their total expenditures and value to the retailer. The objective of this research is therefore to propose a modeling framework for multi-category customer base analysis in a noncontractual setting.

Modeling customer purchase patterns across categories requires a number of careful considerations. First, a customer’s arrival process to the firm and her category purchase decisions are likely to be interdependent. Consider a customer whose shopping trips to the store are driven by multiple categories. Depending on the nature of the relationship among the categories for the customer, the combination of product categories in her shopping basket may be informative of how long it will be until her next shopping trip to a store. For example, if a customer were to purchase beef and pork from a butcher to prepare meals, the time until the customer returns to the butcher to make another purchase may be longer than if she were to just have purchased one of the meats on her previous trip. Preparing a meal with one of the meats may delay the consumption of the other, resulting in a longer time to consume both meats before making the next shopping trip. A customer may return to a store sooner following a trip in which she bought several items together if those items are jointly consumed, such as ingredients going into a salad. In turn, the customers’ interarrival time to the store may affect their category purchase decisions upon a visit. Second, in a multi-category context, a customer’s purchase decisions on a given shopping trip may be correlated across categories. Accounting for such associations is therefore important to capture the simultaneous choice of categories in a shopping basket. Third, repeatedly observed shopping behavior from the same customer could be correlated over time, as a customer’s current choice decisions are likely to be associated with past outcomes both within and across categories. Finally, latent customer attrition needs to be considered in a noncontractual setting.

In this research, we develop a model for multi-category customer base analysis which accom-
modates the aforementioned complications by generalizing the “buy ‘til you die” framework to a multi-category context. Toward this end, we jointly model customers’ interarrival process to the firm, multi-category purchase decisions, and latent attrition while explicitly accounting for the shopping dynamics arising from the interplay of purchase timing and incidence across categories. The key challenge in modeling the interarrival process and its association with multi-category purchase behavior is that the number of possible compositions of shopping baskets increases exponentially as the number of categories increases, and so does the complexity of the model.\textsuperscript{1} To address this issue parsimoniously, we employ a latent space approach that can be generalized to a large number of product categories with ease. In modeling the shopping basket composition, we account for interdependent category choices within a transaction and the correlation of repeatedly observed outcomes over time both within and across categories. We also allow customers’ category purchase decisions to depend on the timing of their past purchase of the category. Therefore, our model captures the sequential interaction between shopping basket choice and interarrival times.

We apply our model to category-level customer data from a leading beauty care company in Korea. Though the firm’s product categories do not have clear demand relationships with each other, we observe systematic variation in customers’ interarrival times related to the combination of categories purchased at their prior transaction. Through a series of benchmark models, we assess the need to incorporate the association of shopping basket choice and interarrival times as well as cross-category dependence both within and across transactions. We find that ignoring the effect of shopping basket composition on the arrival process impairs the prediction of firm-level transactional patterns. Moreover, omitting the cross-category dependence in category choice decisions results in the overestimation of the purchases of shopping baskets containing only one category and the underestimation of the purchases of shopping baskets consisting of multiple

\textsuperscript{1}Throughout the paper, we use the terms, “the composition of shopping baskets” and “shopping basket composition” to refer to the combination of product categories purchased in the same shopping basket on a shopping trip.
categories. Our proposed model offers superior fit and performance in predicting individual customers’ purchase patterns of shopping baskets, in addition to informing our understanding of the relationships that exist among multiple aspects of customer behavior in a multi-category transactional context. The results can be used by marketers for the customization of targeting and communication strategies. For example, firms can identify those customers most likely to make purchases within each of the categories, leveraging their purchase histories across all categories, thereby allowing marketers to tailor their communications based on the categories in which a customer is likely to purchase activity.

The remainder of this article is organized as follows. Section 2 gives an overview of prior literature related to our research. In section 3, we describe the data used in our empirical analysis. Section 4 provides a detailed specification of our model. In section 5, we illustrate the performance of the model and discuss inferences afforded by investigating purchase patterns of shopping baskets. Finally, section 6 concludes and suggests future research directions.

2 Related Work

The current research draws on prior work that has examined customers’ purchase patterns and multi-category behavior. We briefly review relevant literature on both streams of research and discuss the contribution of our work relative to them.

Predicting customers’ purchase patterns from a timing perspective has long been of interest to marketing researchers. Much research in this domain has relied on the assumption of an exponential interarrival process due to its parsimony and performance (e.g., Schmittlein, Bemmaor, and Morrison 1985; Gupta 1991), which underlies the negative binomial distribution (NBD) model (e.g., Ehrenberg 1988; Morrison and Schmittlein 1988). Several researchers have relaxed the assumption of a memoryless timing process, such as by taking steps to allow for nonstationarity in the customer arrival process (e.g., Fader, Hardie, and Huang 2004; Moe and Fader 2004; Schweidel
and Fader 2009), incorporating time-varying explanatory variables (e.g., Gupta 1988, 1991), and considering alternative baseline timing processes (e.g., Jain and Vilcassim 1991; Allenby, Leone, and Jen 1999; Seetharaman and Chintagunta 2003). An important extension of timing models for customer base analysis in transactional settings was the addition of latent attrition, resulting in “buy ‘til you die” models such as the Pareto/NBD (Schmittlein, Morrison, and Colombo 1987) and the BG/NBD (Fader, Hardie, and Lee 2005). These models have served as the foundation for analyzing transactional patterns within a single product category or a firm-level activity in a non-contractual setting (e.g., Batislam, Denizel, and Filiztekin 2007; Wübben and Wangenheim 2008; Abe 2009; Singh, Borle, and Jain 2009; Ma and Büschken 2011; Van Oest and Knox 2011).

Subsequent research has generalized the univariate models of customers’ interarrival times to a multivariate context. Prior research has used multivariate distributions, such as the Farlie-Gumbel-Morgenstern distributions (e.g., Chintagunta and Haldar 1998) and the Sarmanov distributions (e.g., Park and Fader 2004; Schweidel, Fader, and Bradlow 2008) to correlate the timing processes of different categories or transactional activities. Though these models reveal the association between the timing processes, perhaps owing to the model complexity that would accompany generalizing the multivariate distributions beyond two activities, such research has largely been restricted to bivariate applications. Moreover, they did not take into account customer attrition and thus have limitations for customer base analysis. Our model, using a latent space approach (e.g., Bradlow and Schmittlein 2000; DeSarbo and Wu 2001; Li, Sun, and Wilcox 2005), generalizes to a multivariate context in a parsimonious manner. This is valuable from a managerial perspective, as it allows firms to assess the relationships that may exist across a broader range of product categories and assess their impact on customers’ multi-category purchase patterns.

In addition to generalizing univariate multi-event timing models, our work is related to the research on multi-category choice behavior (e.g., Fader and Lodish 1990; Narasimhan, Neslin, and Sen 1996). Recognizing that a customer’s purchase decisions across categories are not indepen-
dent, multivariate choice models have incorporated heterogeneity and coincidence in purchase
decisions (e.g., Manchanda, Ansari, and Gupta 1999; Russell and Petersen 2000; Chib, Seetharaman, and Strijnev 2002; Moon and Russell 2008). Our research differs from extant multi-category choice models in two notable ways. First, rather than focusing solely on the incidence decisions, we jointly model the shopping basket composition and the interarrival process, allowing for their interactive effect on purchase patterns across categories. Though extant research has considered both multi-category choices and the arrival process (e.g., Wagner and Taudes 1986; Gupta 1988; Chintagunta 1999), prior work has not explored the relationship between shopping basket composition and the interarrival process. Second, we incorporate latent attrition in customers’ relationship with the firm to accommodate the importance of accounting for the latent attrition in forecasting future transactional activities (Fader and Hardie 2009). In doing so, our modeling framework enables us to forecast both firm- and category-level metrics of managerial interest.

3 Data

Our data come from a leading beauty care company in Korea. The data set consists of the transaction histories of 2,870 customers who purchased beauty care products of a high-end brand designed for women at department stores in Korea from January 2008 to December 2010. For our purpose, we categorized 79 cosmetics products (i.e., SKU) purchased by the customers during the data period into the following four categories: (1) basics, (2) creams, (3) serums, and (4) makeups. This categorization is also used on the firm’s dashboard for marketing planning and performance evaluation. Basics consist of lotions and toners which are frequently purchased together as must-have skin care items for women. Creams and serums are more advanced and topical (e.g., anti-aging and anti-wrinkle) skin care products, and makeups contain products for facial makeup and cleansing.

Table 1 summarizes the descriptive statistics of our data. On average, customers made 9.06
transactions with the firm and purchased 1.64 different categories per transaction. Out of the total transactions by the customers, 54% were the purchase of only one category, 30% involved two categories, 13% involved three categories, and the remaining 3% were the purchase of all four categories. Among the categories, creams were most frequently purchased (4.62 times per consumer) and makeups were least purchased (2.46 times per consumer). Looking at the combinations of categories purchased together on a shopping trip, we find that customers purchased 5.10 different compositions of shopping baskets through their 9.06 transactions over the data period.

Insert Table 1 about here

To explore the customers’ shopping patterns, we examine the relationship between the number of transactions and the average number of categories purchased per transaction across the customers in Fig. 1. We find a significant negative correlation (-0.29) between the number of transactions and the number of categories purchased per transaction. This indicates that while some customers purchase more categories on fewer transactions, others shop more frequently and buy less per transaction, illustrating the need to simultaneously consider the frequency of visits (i.e., the interarrival process) and shopping basket composition.

Insert Fig. 1 about here

Given customers’ purchases of multiple categories per transaction (an average of 1.64 different categories per transaction), for each shopping basket, in Table 2 we present the frequency with which it is purchased, the number of customers purchasing it and the time taken until the next transaction since they purchased it. The first column of the table lists all possible basket compositions denoted by a vector whose four elements indicate the purchase of basics, creams, serums, and makeups in the order, respectively. For example, \{1,0,0,0\} in the table denotes a shopping basket containing basics only. The table shows that the number of observations varies considerably
across shopping baskets, ranging from 430 for \( \{1,0,1,1\} \) to 4,489 for \( \{0,1,0,0\} \). The most frequently purchased shopping basket \( \{0,1,0,0\} \) contains creams only and was purchased by 62% of the customers, while only 12% of customers purchased the shopping basket \( \{1,0,1,1\} \). Furthermore, the time taken until the next transaction, conditional on a transaction of a specific shopping basket, ranges from 66.3 days to 95.3 days. Though we generally observe longer times until the next transaction following the purchase of shopping baskets with more categories, we find variation in the time until the next transaction following the purchase of shopping baskets containing the same number of categories. As customers purchased multiple different compositions of shopping baskets over the data period (an average of 5.10 different compositions of shopping baskets) and the interarrival times vary considerably with the combination of categories purchased, this suggests the need to consider the association between basket composition and the interarrival process in forecasting customers’ future purchase patterns across categories.

Insert Table 2 about here

### 4 Model Development

Our proposed model consists of the following three components: (1) a customer’s arrival process to a firm, (2) latent attrition, and (3) her choice decisions across multiple categories conditional on arrival. We first present the specification of the model components and discuss how they interact with each other. We then describe our computational approach to estimating the model.

#### 4.1 Timing Model

During the period \((0, T]\), where 0 corresponds to the beginning of the model calibration period and \(T\) is the censoring point that corresponds to the end of the data period, we observe customer \(i\) who makes \(J_i\) transactions at a firm at times \(t_{i1}, t_{i2}, \ldots, t_{iJ_i}\). In each transaction at the firm, the customer makes purchase decisions across \(K\) product categories. We define a binary variable \(C_{ijk}\)
to indicate whether or not customer $i$ purchases category $k$ at her $j$th transaction. Customer $i$’s shopping basket at the $j$th transaction can be represented by a vector of category choice outcomes, $C_{ij} = \{C_{ij1}, C_{ij2}, \ldots, C_{ijK}\}$, where $C_{ijk}$ equals 1 if customer $i$ purchases category $k$ at her $j$th transaction and 0 otherwise. It is important to note that, in our offline shopping context, we have $\sum_{k=1}^{K} C_{ijk} \geq 1$ for any transaction, because a customer’s shopping trip to the firm is observed only when she purchases at least one category.\(^2\)

We first model customer $i$’s arrival process to the firm by specifying the timing behavior for a set of $J_i - 1$ interarrival times, $t_{i2} - t_{i1}$, $t_{i3} - t_{i2}$, $\ldots$, $t_{iJ_i} - t_{iJ_i-1}$, and the right-censored observation $T - t_{iJ_i}$. We assume that customer $i$’s interarrival time between her $j$th and $(j + 1)$th transactions follows an exponential distribution with arrival rate $\lambda_{ij}$. Then, the density function for the interarrival times and the survival function for the right-censored observations are given by:\(^3\)

\[
\begin{align*}
    f(t_{i,j+1}|t_{ij}; \lambda_{ij}) &= \lambda_{ij} \exp\{-\lambda_{ij}(t_{i,j+1} - t_{ij})\}, \\
    S(T|t_{i,J_i}; \lambda_{iJ_i}) &= \exp\{-\lambda_{iJ_i}(T - t_{iJ_i})\}.
\end{align*}
\]  

(1)

When a customer repeatedly shops across multiple product categories, her shopping basket composition at a transaction may be related to the time until her next shop. To take into account this effect, we specify $\lambda_{ij}$ as:

\[
\lambda_{ij} = \lambda_i f(C_{ij}) \exp(X'_{ij} \alpha),
\]  

(2)

where $\lambda_i$ is customer $i$’s baseline arrival rate, $f(\cdot)$ is a function of the prior shopping basket choice ($C_{ij}$), $X_{ij}$ is a vector of covariates which may affect the purchase timing behavior, and $\alpha$ is a vector of the corresponding parameters.

At the heart of our specification of $\lambda_{ij}$ in Eq. (2) is the function $f(C_{ij})$ which captures the association between the customer’s shopping basket choice and the time until her next shopping

\(^2\)This research focuses on modeling customer purchase patterns across categories at an offline firm. However, our model can be easily modified to a context where customer visits may involve no purchases (i.e., $\sum_{k=1}^{K} C_{ijk} = 0$).

\(^3\)We tested alternative distributions such as Weibull and Erlang-2 distributions, and found no significant improvement in model performance.
trip. The function $f(C_{ij})$, as well as the covariates $X_{ij}$, allow for nonstationarity in the timing process and for a customer’s subsequent decisions to be informed by her past decisions. We next discuss alternative specifications of $f(C_{ij})$.

### 4.1.1 An Additive Model

A simple approach that incorporates the effects of previously purchased categories into the arrival rate $\lambda_{ij}$ would be to include $C_{ijk}$s as additive terms with exponential base in $f(C_{ij})$ (to ensure $\lambda_{ij} > 0$). Formally,

$$f(C_{ij}) = \exp(a_1C_{ij1} + a_2C_{ij2} + \cdots + a_KC_{ijK}),$$

(3)

where $a_k$ is a coefficient which captures the effect of customer $i$’s purchase of category $k$ on her next arrival rate.

This specification accounts for the “main effects” of each purchased category on the timing process. If we were to restrict $a_k = a$ for all $k$, this would be equivalent to counting up the number of categories in the shopping basket. However, Eq. (3) does not take into account the interplay that may exist among multiple categories, thereby ignoring the possible joint effects of categories on the interarrival process. To account for this, one could include all possible interaction terms of categories purchased, or equivalently assume different arrival rates following the purchase of each combination of categories. While this approach allows one to account for the dependence of the arrival rate on the composition of the previous shopping basket, the possible shopping basket compositions (and hence model parameters) increase exponentially with the number of categories. To address this issue parsimoniously, we next propose a modeling approach that places product categories in a latent space (e.g., Bradlow and Schmittlein 2000; Li, Sun, and Wilcox 2005) and use this to parameterize their relationship to the interarrival process.
4.1.2 A Latent Space Model

We model the effect of shopping basket composition on the arrival rate using the Euclidean distances between the categories and the origin in the space. To formalize this, we specify $f(C_{ij})$ as:

$$f(C_{ij}) = \left\| \sum_{k=1}^{K} P_k C_{ijk} \right\|^{-1},$$

(4)

where $P_k$ is the position of category $k$ in an $n$-dimensional latent space, represented by $n$ coordinates to be estimated, and $\| \cdot \|$ denotes the Euclidean norm of a point (i.e., the Euclidean distance between the point and the origin).

We illustrate the intuition behind Eq. (4) with an example of two categories on a two-dimensional space. $P_1$ and $P_2$ in Fig. 2 represent categories 1 and 2, respectively, on the space where the positions of the categories are denoted by their $x$- and $y$-coordinates. The evaluation of the Euclidean norm in Eq. (4) and the inverse relationship between the arrival rate and interarrival time indicate that when a customer purchases category $k$ only, the effect of the purchase on the interarrival time until the next transaction is given by $\|P_k\| = (x_k^2 + y_k^2)^{1/2}$. Thus, a larger (smaller) value of $\|P_k\|$ implies that the customer’s interarrival time after purchasing category $k$ tends to be longer (shorter).

Insert Fig. 2 about here

Next, when a customer purchases both categories 1 and 2, its effect on the interarrival time is given by the Euclidean norm of the summed vector, $\|P_1 + P_2\| = \{(x_1 + x_2)^2 + (y_1 + y_2)^2\}^{1/2}$. Note that the Euclidean norm varies depending on the angle, $\delta$, between the two vectors, $P_1$ and $P_2$. For a given length of the vectors, as $\delta$ decreases, $\|P_1 + P_2\|$ increases and thus the interarrival time after purchasing both categories becomes larger, which we would anticipate for product categories in which consuming one may delay the consumption of the other. In contrast, larger values of $\delta$ would reflect shorter expected interarrival times, which would occur when consuming one
category is expected to accelerate the consumption of another. Thus, \( \delta \) reflects the combined effect of the categories on the arrival rate.\(^4\) The same logic would apply to cases with more than two categories.

Note that when a one-dimensional space is employed, each category is represented by a single coordinate (i.e., \( x \)-coordinate) on the space. Hence, similar to the case in Eq. (3), the specification of \( f(C_{ij}) \) in Eq. (4) only accounts for the main effects of the categories using the same set of covariates with a different functional form (i.e., \( f(C_{ij}) = | \sum_{k=1}^{K} x_k C_{ijk} |^{-1} \)). The latent space can be extended to more than one dimension to provide increased flexibility. Selection of the number of dimensions in the latent space would depend on the goal of the analysis. A two-dimensional space would allow for a relatively simple representation of categories in the space. However, a higher dimensional space may provide better model fit as the number of categories increases. For identification, the number of dimensions cannot exceed the number of categories considered.

The Euclidean norms of a set of vectors are invariant under rotation and reflection of the space. To identify the model, we impose the following restrictions. For a two-dimensional space, we restrict the \( x \)-coordinate of category 1 to be positive and the \( y \)-coordinate of the category to zero. This takes into account rotation of the space. We also restrict the \( y \)-coordinate of category 2 to be positive so as to account for reflection of the space over the \( y \)-axis. Finally, because \( \lambda_{ij} \) is proportional to the product of \( \lambda_i \) and \( \| \sum_{k=1}^{K} P_k C_{ijk} \|^{-1} \), there are infinite number of combinations of \( \lambda_i \) and \( P_k \) that give the same value of \( \lambda_{ij} \). To consider this, we set the \( x \)-coordinate of category 1 to one. The remaining coordinates in the latent space are treated as model parameters to be inferred from the data. Similar identification conditions can be derived for a general case with an

\(^4\)It is worth noting that one can also use the Euclidean norm without taking an inverse in Eq. (4), that is, \( f(C_{ij}) = \| \sum_{k=1}^{K} P_k C_{ijk} \| \). In such case, a smaller (larger) \( \delta \) indicates a larger (smaller) arrival rate and thus a shorter (longer) interarrival time after purchasing both categories. We take an inverse of the Euclidean norm so that purchasing categories which are closer to each other in the space likely leads to a longer interarrival time until the next transaction, which we believe is more intuitively appealing for the interpretation of the latent space.
\( n \)-dimensional space as follows. First, we restrict the first coordinate of category 1 to one and all other coordinates of the categories to zero. For category \( k \) such that \( 1 < k \leq n \), we restrict the \( k \)th coordinate of the category to be positive and the \( m \)th coordinate to be zero for all \( m \) such that \( m > k \). No identification condition is required on the coordinate of category \( k > n \). Using a latent space model allows us to parameterize the effects of shopping basket composition parsimoniously as \( n \) increases.

### 4.2 Customer Attrition

As our empirical context is transactional (i.e., noncontractual) in nature, customers may cease shopping at the firm and silently defect at an unobserved time. The failure to incorporate latent attrition can result in biased parameter estimates for the arrival process. We therefore incorporate latent attrition by assuming that customer \( i \) becomes permanently inactive after a transaction with probability \( r_i \), consistent with the individual-level attrition process of the BG/NBD model (Fader, Hardie, and Lee 2005).

\[
P(\text{Customer } i \text{ drops out after her } j \text{th transaction}) = r_i (1 - r_i)^{j-1}.
\]

### 4.3 Shopping Basket Choice Model

We next model a customer’s multi-category choice behavior in her transaction at the firm. Our proposed choice model is a modified version of the conditional logit model, which has been em-

\[ r_{ij} = \left[ 1 + \exp \left( -\varphi_i \left\| \sum_{k=1}^{K} C_{ijk} Q_k \right\| \right) \right]^{-1}, \]

where \( \varphi_i \) represents a customer-specific baseline dropout tendency and \( Q_k \) is a point representing category \( k \) on the latent space. Similar to the timing model of customer arrivals, the probability of customer attrition is determined based on the relative positions of the purchased categories in the space. However, we have found neither improvement in model fit nor meaningful results, and therefore employed the simple geometric process of the attrition at the firm level. The lack of improvement in model fit may not be surprising because the attrition process is not observed.
ployed in a few marketing studies to account for interdependent choices in multiple categories (e.g., Russell and Petersen 2000; Moon and Russell 2008). The model captures cross-category dependence by incorporating the choice outcomes in other categories into the choice decision in the focal category. The model also allows us to derive the closed form expression for the unconditional joint distribution of a customer’s category choices, which greatly facilitates our objective of predicting future shopping basket composition.6

We specify the probability that customer $i$ purchases category $k$ at the $j$th transaction conditional on her choices of other categories, using the following logit form:

$$
\Pr(C_{ijk} = 1 | C_{ijk}', k' \neq k) = \left[ 1 + \exp \left\{ - \left( \pi_{ijk} + \sum_{k' \neq k} \theta_{kk'} C_{ijk'} + \sum_{k' \neq k} \psi_{kk'} C_{i,j-1,k'} \right) \right\} \right]^{-1}, \tag{6}
$$

where $\pi_{ijk}$ captures the time-varying category-specific effect for category $k$ at customer $i$’s $j$th transaction and $\sum_{k' \neq k} \theta_{kk'} C_{ijk'}$ reflects the effects of the choices of other categories on the choice decision in category $k$. Hence, $\theta_{kk'}$ measures the degree of interdependence between category $k$ and $k'$. A positive $\theta_{kk'}$ implies that the purchase of category $k'$ increases the probability of purchasing category $k$ at the same transaction. A negative $\theta_{kk'}$ reflects a negative association between the choices of categories $k$ and $k'$. Finally, $\sum_{k' \neq k} \psi_{kk'} C_{i,j-1,k'}$ captures the effects of the choices of other categories in the prior transaction on the choice of category $k$ in the current transaction. A positive (negative) $\psi_{kk'}$ implies that the purchase of category $k'$ in the prior transaction increases (decreases) the probability of purchasing category $k$ in the current transaction.

To consider the effect of time-varying covariates on the category choice probability, we further specify the category-specific intercept $\pi_{ijk}$ as:

$$
\pi_{ijk} = \beta_{0ik} + Z'_{ijk} \beta_k, \tag{7}
$$
where $\beta_{0ik}$ is a customer-specific baseline purchase tendency for category $k$, $Z_{ijk}$ is a vector of time-varying covariates, and $\beta_k$ is a vector of the corresponding parameters for category $k$.

Because the conditional probabilities in Eq. (6) are mutually dependent across categories, we cannot use the equation to predict a customer’s choice decisions across multiple categories. By assuming $\theta_{kk'} = \theta_{k'k}$ and applying the theorem by Besag (1974), however, it can be shown that the unconditional joint distribution of $C_{ij}$ is given by:

$$
P \left( C_{ij} = C_{ij}^*, \sum_{k=1}^{K} C_{ijk}^* \geq 1 \right) = \frac{\exp(\pi_{ij}'C_{ij}^* + \frac{1}{2}C_{ij}^*\Theta C_{ij}^* + C_{ij-1}'\Psi C_{ij-1})}{\sum_{C_{ij} \neq \{0,0,...,0\}} \exp(\pi_{ij}'C_{ij} + \frac{1}{2}C_{ij}\Theta C_{ij} + C_{ij-1}'\Psi C_{ij-1})},
$$

(8)

where

$$
\pi_{ij} = \begin{bmatrix} 
\pi_{ij1} \\
\pi_{ij2} \\
\vdots \\
\pi_{ijK} 
\end{bmatrix}, \quad \Theta = \begin{bmatrix} 
0 & \theta_{12} & \cdots & \theta_{1K} \\
\theta_{12} & 0 & \cdots & \theta_{2K} \\
\vdots & \vdots & \ddots & \vdots \\
\theta_{1K} & \theta_{2K} & \cdots & 0 
\end{bmatrix}
$$

and

$$
\Psi = \begin{bmatrix} 
0 & \psi_{12} & \cdots & \psi_{1K} \\
\psi_{21} & 0 & \cdots & \psi_{2K} \\
\vdots & \vdots & \ddots & \vdots \\
\psi_{K1} & \psi_{K2} & \cdots & 0 
\end{bmatrix}
$$

and the summation in the denominator is across the possible shopping baskets constructed from $K$ categories excluding the “no-purchase” case. The derivation of Eq. (8) is provided in the Appendix.

While the proposed multi-category choice model in Eq. (8) allows us to flexibly capture the interdependence across category choices within a transaction and their temporal dependence across transactions, one potential problem is that the number of parameters in both matrices $\Theta$ and $\Psi$ quadratically increases as the number of categories increases. To address this issue, we parameterize $\theta_{kk'}$ and $\psi_{kk'}$ as:

$$
\theta_{kk'} = \zeta_{0k} + \zeta_{0k'} + \zeta_1 \|P_k + P_{k'}\| \quad \text{and} \quad \psi_{kk'} = \xi_{0k} + \xi_{1k'} + \xi_2 \|P_k + P_{k'}\|,
$$

(9)

where $\zeta_{0k}$, $\zeta_{0k'}$ and $\xi_{1k'}$ are category-specific parameters, $\|P_k + P_{k'}\|$ is the Euclidean norm of the summed vector, $P_k + P_{k'}$, and $\zeta_1$ and $\xi_2$ are coefficients for $\|P_k + P_{k'}\|$.

Our parameterization of $\theta_{kk'}$ and $\psi_{kk'}$ can be thought of as a decomposition into main effects associated with categories $k$ and $k'$, and the interaction between categories $k$ and $k'$ which is ac-
counted for by the Euclidean norm $\|P_k + P_{k'}\|$. $\theta_{kk'}$ is comprised of the fixed effects of categories $k$ and $k'$ ($\zeta_0k$ and $\zeta_0k'$) and the effect of their spatial relationship in the latent space employed in the timing model ($\zeta_1k \|P_k + P_{k'}\|$). Similarly, $\psi_{kk'}$ is decomposed into the fixed effects of categories $k$ and $k'$ ($\zeta_0k$ and $\zeta_1k'$) and the effect of their spatial relationship in the latent space used in the timing model ($\zeta_2k \|P_k + P_{k'}\|$). Note that the specification of the category-fixed effects for $\psi_{kk'}$ accompanies $K$ more parameters than the specification for $\theta_{kk'}$, because $\psi_{kk'}$ is asymmetric with respect to $k$ and $k'$ and $\theta_{kk'}$ is symmetric with respect to $k$ and $k'$. We set $\xi_{01}$ to zero for identification which serves as a baseline across different categories. No identification condition is needed for $\theta_{kk'}$ given its symmetry. The Euclidean norm, $\|P_k + P_{k'}\|$, is included to allow for the possibility that the relationship between categories $k$ and $k'$ in the multi-category choice process is correlated with the relationship between the categories in driving interarrival times.\footnote{In parameterizing $\theta_{kk'}$ and $\psi_{kk'}$, one may consider employing new latent spaces intended to capture the interactive effects of categories $k$ and $k'$ on the parameters, instead of the one used in the timing model. We have chosen not to pursue this direction because it significantly increases the model complexity and the proposed specification in Eq. (9) allows us to track customers’ purchase patterns across categories accurately.}

4.4 Customer Heterogeneity and Correlation Structure

To incorporate customer heterogeneity into our model, we specify the customer-specific model parameters as follows. For the customer-specific baseline arrival rate in Eq. (2), we assume that $\lambda_i$ follows a lognormal distribution to ensure that $\lambda_i$ is positive. Second, we reparameterize the dropout probability $r_i$ in Eq. (5) as $r_i = \frac{\exp(\omega_i)}{1+\exp(\omega_i)}$ to ensure $r_i \in [0, 1]$, and assume that $\omega_i$ follows a normal distribution. Third, we assume that the customer-specific intercepts $\beta_{0i1}, \beta_{0i2}, \ldots, \beta_{0iK}$ in Eq. (7) follow a normal distribution. Taken together, we assume

$$
\begin{bmatrix}
\log \lambda_i \\
\omega_i \\
\beta_{0i1} \\
\vdots \\
\beta_{0iK}
\end{bmatrix}
\sim \text{MVN}
\begin{bmatrix}
\mu_\lambda \\
\mu_\omega \\
\mu_{\beta_1} \\
\vdots \\
\mu_{\beta_K}
\end{bmatrix}, \Sigma
$$

(10)
to account for the correlation of the model parameters. Incorporating the covariance structure allows us to capture the interdependence of timing (arrival and attrition) and category choice decisions within a customer.

Similar to the matrices $\Theta$ and $\Psi$, the number of variance-covariance parameters in the matrix $\Sigma$ quadratically increases as the number of categories considered increases. When this causes significant computational inefficiency in model estimation, one can consider the reparameterization of the covariance between $\beta_{0ik}$ and $\beta_{0ik'}$ using the category-specific parameters, in a similar manner used in Eq. (9). Alternatively, one may employ a block diagonal covariance structure in the matrix. For example, one may assume covariance parameters between $\beta_{0ik}$ and $\beta_{0ik'}$ are zero at expense of not incorporating the correlation across category choices within individuals.

### 4.5 Computational Approach

We adopt a Bayesian approach and use the Markov chain Monte Carlo (MCMC) methods to estimate the proposed model. The model is estimated using the Gibbs sampler (Gelfand and Smith 1990) with the Metropolis-Hastings steps (Hastings 1970) in the publicly available software WinBUGS. The samples obtained from the MCMC algorithm are then used to compute summary measures of the parameter estimates. The results reflect the output of MCMC draws for 30,000 iterations, with a burn-in period of 30,000 iterations. To complete the Bayesian specification of the model, we assume noninformative conjugate priors to the parameters. For aggregate-level parameters and mean parameters, we use a diffuse normal density prior. For the variance-covariance matrix, we assume that the inverse of the matrix follows a Wishart distribution.

### 5 An Empirical Application

In this section, we provide an empirical demonstration of the proposed model, using data described in section 3. We use the first 30-month period of the data for model calibration and the
remaining six-month period for model validation. The calibration data consist of 2,621 customers who made repeated transactions during the calibration period.

5.1 Covariates

We begin by describing vectors of covariates, $X_{ij}$ and $Z_{ijk}$ included in Eqs. (2) and (7), respectively. $X_{ij}$ contains the lagged interarrival time (i.e., customer $i$’s interarrival time between her $(j-1)$th and $j$th transactions), denoted by $\text{LIT}_{ij}$, to take into account the dependence of the current interarrival time on the past observation of the outcome variable. $Z_{ijk}$ contains (1) the lagged choice of category $k$ (i.e., $C_{i,j-1,k}$) to consider the temporal dependence of choices within a category, (2) the elapsed time between customer $i$’s $j$th transaction and her last purchase of category $k$ before the transaction ($\text{ET}_{ijk}$) to consider the relationship between the frequency of purchasing category $k$ and the category choice probability in a given transaction and (3) the quadratic term of $\text{ET}_{ijk}$ to capture the possible nonlinear time effect of the prior category purchase on the current purchase decision. For example, the probability of purchasing the category at a given transaction may increases as the time since the last purchase of the category increases. On the other hand, a very long elapsed time may indicate that the customer has stopped purchasing the category, resulting in a small purchase probability. By definition, $\text{ET}_{ijk} = \text{LIT}_{ij}$ if customer $i$ purchased category $k$ at her $(j-1)$th transaction and $\text{ET}_{ijk} > \text{LIT}_{ij}$ otherwise. If other time-varying covariates were available and believed to affect the interarrival or multi-category choice processes, they could be incorporated into $X_{ij}$ and $Z_{ijk}$ (e.g., Schweidel and Knox 2013).

Incorporating the aforementioned covariates into the model allows us to parsimoniously capture the association of customers’ purchase timing and category choice decisions across transactions, and the correlation of repeatedly observed outcome measures. Because all covariates are constructed based on the outcome variables we model in section 4, the values of the covariates can be simulated beyond the calibration period to predict the customers’ future purchase patterns.
5.2 Model Comparison

We fit a series of models to assess the importance of accounting for the interplay between shopping basket composition and interarrival times, the cross-category effect on category choices within a transaction, and the temporal dependence across category choices. The alternative models we examine therefore mainly vary with respect to the following three aspects: (1) whether or not (and how) to incorporate the effect of shopping basket composition on the arrival rate, (2) whether or not to account for the interdependence across category choices within a transaction, and (3) whether or not to consider the temporal dependence across category choices.

The first benchmark model (Model 1) is constructed using the BG/NBD model proposed by Fader, Hardie, and Lee (2005). We employ four independent BG/NBD models for the prediction of purchase patterns of the four categories. Model 1 thus completely ignores all of the three aforementioned sources of correlation as well as the effects of covariates included in $X_{ij}$ in Eq. (2) and $Z_{ijk}$ in Eq. (7). Model 2 is the simplest version of our model which also fails to consider the three modeling aspects and the effects of the covariates. This model is formulated by setting $f(C_{ij}) = 1$ and $X_{ij} = 0$ in Eq. (2), $\theta_{kk'} = 0$ and $\psi_{kk'} = 0$ in Eq. (6), and $Z_{ijk} = 0$ in Eq. (7). Note that this model differs from Model 1 in that the model components are still related at the margin due to the correlation that may exist among the underlying model parameters, as specified in Eq. (10). Model 3 extends Model 2 by simply including the covariate vectors $X_{ij}$ and $Z_{ijk}$ in Eqs. (2) and (7), respectively.

Model 4 extends Model 3 by considering the temporal dependence across category choices through $\psi_{kk'}$’s, but still ignores the impact of shopping basket choices on interarrival times ($f(C_{ij}) = 1$) and the cross-category effects within a transaction ($\theta_{kk'} = 0$). Model 5 extends Model 4 by adopting our multi-category choice model to account for the interdependence across category choices through $\theta_{kk'}$’s, but fails to consider the association of shopping basket composition with interarrival times ($f(C_{ij}) = 1$). In contrast to Model 5, Model 6 incorporates the effect of shopping basket
choices on interarrival times by employing the proposed latent space approach, but ignores both
the cross-category effects within a transaction ($\theta_{kk'} = 0$) and the temporal dependence across
category choices ($\psi_{kk'} = 0$).

Model 7 allows for the cross-category effects within a transaction as well as the temporal de-
pendence across category choices. However, when it comes to the effect of category choices on
the arrival rate, the model only considers the main effects of purchased categories and does not
account for their interplay by assuming that $f(C_{ij})$ is given as in Eq. (3). Model 8 is the most
generalized version constructed using our modeling framework. The model explicitly considers
the effect of shopping basket composition on the arrival rate by including all two-, three- and
four-way interaction terms of $C_{ijk}$’s as well as their main terms within the exponential term in Eq.
(3). The model also explicitly considers both the cross-category effects within a transaction and
the temporal dependence across category choices by employing $\theta_{kk'}$ and $\psi_{kk'}$ in Eq. (6) without
the parameterization specified in Eq. (9). As a result, the model is more heavily parameterized
than our proposed model and the number of parameters exponentially increases as the number of
categories increases. Finally, Model 9 is our proposed model in which the latent space approach is
used in specifying $f(C_{ij})$ and the parameterization of $\theta_{kk'}$ and $\psi_{kk'}$ as shown in Eq. (9).

To provide a comparison of the performance of the benchmark models, we compute the log
marginal density of the models. A larger value of the log marginal density indicates a better
model fit. We also compute the mean absolute error (MAE = $|\text{Predicted} - \text{Observed}|$) with respect
to the number of purchases of each shopping basket by individual customers, averaged across
shopping baskets and customers. For the computation of the MAE of Model 1, which consists of
four independent category purchase processes, we assume that categories are purchased together
in the same shopping basket if the purchases of the categories are predicted to occur in the same

---

8Because Models 4,5 and 7 do not employ the latent space approach in specifying the arrival rate, the models also
do not use the Euclidean norm in reparameterizing $\theta_{kk'}$’s and/or $\theta_{kk'}$’s in Eq. (9).
9We were unable to compute the mean absolute percentage error (MAPE = $|\text{Predicted} - \text{Observed}|/\text{Observed}$)
because the observed numbers of purchases of some shopping baskets were zero for some customers.
Table 3 reports the model comparison results. We find that Model 2 provides a better fit compared to Model 1 according to the all three fit measures. This suggests that, for the prediction of the purchase patterns of shopping baskets, our approach of modeling the firm-level arrival process and category choice decisions conditional on a transaction occurring outperforms applying independent BG/NBD models to individual categories, even when we do not explicitly consider any of the associations in our proposed model. We find that Model 3 provides a better fit compared to Model 2, suggesting that the covariates considered in our study are useful in predicting customers’ purchase patterns across multiple categories. The better fit of Model 4 over Model 3 implies that the importance of incorporating the temporal dependence across category choices. Comparing the performance of Model 5 to Model 4 suggests that customers’ choice decisions are correlated across product categories.

We find that Model 6 provides a better fit compared to Model 3, suggesting an association between the product categories comprising the shopping basket and the arrival rate. The better fit of Model 9 over Model 5 also implies the association of the two model components. In examining the performance of Models 6 and 9 with different dimensional latent spaces, we find that the model fit improves as we include additional dimensions in the space. However, we find negligible differences in model fit between the model with a three-dimensional space and the model with a four-dimensional space, implying the flexibility gained from the additional fourth dimension is marginal. Comparing the performance of Model 7 to Model 5 suggests that considering the main effects of the categories on the arrival rate is helpful in predicting customers’ purchase patterns across categories. Furthermore, the better performance of Model 9 with higher-dimensional spaces, compared to Model 7, indicates that incorporating the effect of the overall shopping basket
composition allows additional benefits beyond simply considering the main effects of the categories.

We find that the fit of Model 8 is slightly better than Model 9 in the calibration period but there is no difference between their performances in the validation period, suggesting that Model 9 successfully approximates the correlations within and across model components with a smaller number of parameters.\footnote{This may not be surprising because a good in-sample fit of more complicated models does not necessarily lead to a good holdout fit, often due to an overfitting problem (Malthouse 1999).} Given that Model 9 has comparable forecasting ability to Model 8 with fewer parameters and that the latent space modeling approach more easily scales to accommodate more categories, we proceed to present detailed findings under this model specification.

In sum, our results suggest the importance of accounting for the association of shopping basket composition with the arrival rate, the cross-category effects within a category, and the temporal dependence across category choices. Compared to the benchmark model that omits all three sources of correlation (Model 3), the MAE for Model 9 is 15\% lower in the calibration period and 22\% lower in the validation period. As such, we focus the remainder of our discussion on the results from our proposed model with a three-dimensional latent space because the parsimony with a smaller number of model parameters outweighs the benefit of the minor improvement in the model fit with a four-dimensional space.\footnote{We also tested and compared the forecasting ability of the proposed models with a three- and four-dimensional space, respectively, and found no significance difference in the performance between the models. Results are available from the authors on request.}

### 5.3 Model Performance

To validate the model, we estimate the following measures, averaged across MCMC iterations: (1) the number of transactions, (2) the mean interarrival times after purchasing a specific shopping basket, and (3) the number of purchases of a shopping basket over the data period.

Using the estimates of the model parameters, we predict the customers’ transaction patterns at
the firm level over time. Fig. 3 compares the predicted posterior means of the cumulative number of transactions to the corresponding observed values, aggregated across the customers. We find that the observed values are contained in the 95% posterior intervals of their respective predicted values throughout the data period. At the end of the calibration (validation) period, the model predicts the observed number at a 3.0% (2.6%) error rate, indicating that the model can accurately track the customers’ cumulative transactions.

Insert Fig. 3 about here

We next compare the observed and predicted values of the mean interarrival times after the customers purchase each combination of categories in Fig. 4. The results show that all of the observed values are contained in the 95% posterior intervals of their respective predicted values except the case of the shopping basket \(\{1,0,0,1\}\), which is among the least purchased shopping baskets. The average of the percentage errors weighted by the number of observations for shopping baskets is 8.9%. As expected, the prediction error tends to be smaller for more frequently observed shopping baskets and larger for less frequently observed ones (see Table 2 for the frequency of each shopping basket). The dotted lines in Fig. 4 represent the 95% posterior intervals of the predicted mean interarrival times under Model 5. Without considering the association of the interarrival process with shopping basket composition, the predicted interarrival times would not vary across shopping baskets. As a result, on average the benchmark model underpredicts the interarrival times after purchasing smaller shopping baskets and overpredicts the interarrival times after purchasing larger shopping baskets. Moreover, the observed values for more than half of the possible shopping basket compositions are not contained in the corresponding 95% posterior intervals. We find that the percentage error under Model 5 (Model 7) is 11.0% (10.8%), an increase in the percentage error of 24% (21%). This suggests that incorporating the effect of shopping basket composition on the arrival process improves firm-level predictions of transactional
Our proposed model also allows one to forecast purchases of customers’ shopping baskets. Marketers can use Eq. (8) to compute the probability that a customer purchases a specific shopping basket at a given transaction conditional on her past purchase records, which can aid firms in their promotional strategies. For example, to drive store traffic, marketers may incorporate those product categories in which a customer is most likely to purchase in targeted communications. Having identified those product categories that customers are most likely to purchase together may also aid retailers with in-store promotional activity.

To ensure that the model captures the customers’ purchase patterns of shopping baskets, we predict the cumulative number of purchases of each shopping basket over the data period. Fig. 5 compares the predicted posterior means of the cumulative number of transactions to the corresponding observed values, aggregated across the customers. We find that for all 15 different shopping baskets, the observed values are contained in the 95% posterior intervals of their respective predicted values throughout the data period. The tracking plots suggest that our model tracks the customers’ purchases of shopping baskets fairly well. At the end of the calibration (validation) period, the model predicts the cumulative purchases at a 7.2% (8.6%) error rate, averaged across shopping baskets. In comparison, the prediction errors increase to 10.2% (11.9%) under Model 4, corresponding to an increase in the error by 42% (38%). We find that Model 4 overestimates the purchases of shopping baskets containing only one category and underestimates the purchases of shopping baskets consisting of multiple categories as a result of ignoring the cross-category dependence detected under our proposed model. By incorporating the model component, Model 5 has lower errors compared to Model 4 (8.1% in calibration period and 9.3% in validation period). The results therefore highlight the importance of considering interdependence across category
choices in evaluating customers’ likelihood of purchasing a specific combination of categories and thus forecasting cross-category sales over a future time period.

Insert Fig. 5 about here

Lastly, by aggregating the customers’ purchase patterns of shopping baskets, we predict the cumulative category purchases over the data period. Fig. 6 shows the comparison of the distributions of the observed and predicted values across the customers for each category. The model fit results clearly support the ability of our model to capture the purchases of individual categories.

Insert Fig. 6 about here

5.4 Parameter Inferences

We describe inferences based on the posterior distributions of our model parameters. We begin with the results for the model of customer arrivals and attrition. From the estimates of $\mu_\lambda$, $\mu_\omega$, $\Sigma_{\lambda\lambda}$, and $\Sigma_{\omega\omega}$ in Table 4a, we find that the mean baseline arrival rate is $0.10 = e^{\mu_\lambda + \Sigma_{\lambda\lambda}/2}$, and on average customers defect with a probability of 0.01 after their each transaction.12 Directionally, we find the estimate of the coefficient for LIT$_{ij}$ is negative, suggesting that a long lagged interarrival time reduces the purchase frequency, but this effect does not significantly differ from 0.

Insert Table 4a about here

In Table 4b, we report the estimated coordinates of the four categories in the three-dimensional latent space. As specified in Eq. (4), the coordinates jointly determine the Euclidean norm for a chosen shopping basket which governs the subsequent interarrival times (see Fig. 4 for the predicted mean interarrival times after purchasing each shopping basket). Fig. 7 visualizes the position of the categories in the latent space, showing a set of two-dimensional maps to portray

\[12\text{Note that if } X \sim \text{N}(\mu, \sigma^2), Y = e^X \text{ follows a log-normal distribution whose mean is given by } e^{\mu + \sigma^2/2}.\]
the three-dimensional latent space. From the first two panels of the figure, we see that basics and makeup are relatively close to each other with respect to their first coordinates, as are creams and serums. As such, the first dimension of the latent space may reflect characteristics of the categories which contrast basics and makeup to creams and serums in women’s use of cosmetics, such as “essential versus optional” or “basic versus functional.” From the third panel of the figure, we find that the second (third) dimension allows for differentiation between creams (serums) and the other product categories.

Insert Table 4b and Fig. 7 about here

A next set of inferences arises from the multi-category choice model. Table 5a reports the parameter estimates that capture the category-specific effects specified in Eq. (7). Based on the estimates of \( \mu_k \)'s, we find that customers’ average baseline purchase tendencies are highest for creams and lowest for makeups. The parameter estimates for \( C_{i,j-1,k} \)'s indicate that the purchase of basics, creams and serums in the prior transaction reduces the likelihood of purchasing in those categories at the current transaction. In contrast to these categories, the prior purchase of makeups does not significantly affect the likelihood of purchasing the category. In all categories, the parameter estimates for both \( ET_{ijk} \)'s and \((ET_{ijk})^2\)'s are significant, suggesting an inverse U-shaped relationship between the time since the last purchase of the category and the probability of purchasing the category. This indicates an increasing purchase probability up to some point, after which customers are increasingly less likely to purchase the category. Combined with the results of the timing model, this implies that a customer’s shopping timing and category choice decisions are sequentially associated with each other across transactions, an association that has not previously been discussed in the customer relationship management literature.

Insert Table 5a about here
In Table 5b, we report the estimates of parameters which jointly capture the interdependence across category choices, reflected by $\theta_{kk'}$. We find that the estimates of $\zeta_0$'s for all categories except makeups are significant. The estimate of $\zeta_1$ is not significant, which implies that the cross-category effect in the multi-category choice process is not correlated with the relationship of categories in driving interarrival times. Using the parameter estimates, we compute the values of $\theta_{kk'}$'s according to the specification in Eq. (9) across MCMC iterations, and summarize the results in Table 5c. With the exception of a pair of basics and makeups, the remaining cross-category effects are all positive and significant, which implies that purchasing in one category increases the probability of purchasing in the other category in the same transaction. Thus, the other five pairs of the categories are more likely to be co-purchased than a pair of basics and makeups. Among the five significantly associated pairs, creams and serums (creams and makeups) have the largest (smallest) positive impact on each other. Interestingly, we find that the cross-category effects involving makeups (the third column of Table 5c) tend to be smaller compared to the effects involving the other categories, suggesting a weaker association with the other categories. This may be due to makeups being relatively more distinct from the other three categories.

Insert Tables 5b and 5c about here

Table 5d presents the estimates of parameters which jointly capture the temporal dependence across category choices, reflected by $\psi_{kk'}$. From the estimates, the values of $\psi_{kk'}$'s are computed in Table 5e. The parameter estimates in the first column of the table reveal that the lagged purchase of basics has a significant and negative effect on the purchase of creams. From the estimates in the second (third) column, we find that the lagged purchase of creams (serums) has a significant and negative effect on the purchase of basics and serums (basics and creams). The estimates in the fourth column and row indicate that purchasing makeups on the prior trip adversely affects the probability to purchase other categories while lagged purchases of other categories do not affect
purchases of makeups. These results illustrate the modeling framework’s flexibility in capturing an asymmetric relationship in the cross-category associations across transactions. Such insights can be of value to salesforces and marketers in tailoring how they approach customers, whether with cross-selling offers and customized communications, based on their past purchases.

**Insert Tables 5d and 5e about here**

Finally, we report the estimated variance-covariance matrix $\Sigma$ in Table 6. The first row of the matrix shows the covariance between $\log \lambda_i$ and the intercepts for the other model components. The negative relationship between $\log \lambda_i$ and $\omega_i$ indicates that customers who purchase more frequently are less likely to become inactive. We also find a negative relationship between $\log \lambda_i$ and $\beta_{0i1} - \beta_{0i3}$, which suggests that the shopping baskets of more (less) frequent shoppers are likely to be smaller (larger) in the number of categories. The result indicates the need to consider the association of the interarrival process and shopping basket composition across customers as well as across transactions within individuals.

**Insert Table 6 about here**

### 6 Conclusion and Future Research

This research develops a model for multi-category customer base analysis in a noncontractual setting. We jointly model customers’ arrival process to the firm, multi-category choice decisions, and latent attrition in an integrated framework, allowing for shopping dynamics due to the interplay of purchase timing and incidence across categories. To capture the association between the arrival rate and shopping basket choice parsimoniously, we propose a novel modeling approach using a latent space of product categories. Our model also accounts for the interdependent choices of multiple categories and the correlation of repeatedly observed outcomes both within and across categories. Applying our model to category-level customer transaction data, we demonstrate the
importance of accounting for the interplay of interarrival times and shopping basket composition, and the cross-category effects within and across transactions. The proposed model offers superior fit and performance in predicting customers’ purchase patterns across categories compared to several benchmark models. The ability of our model to forecast the purchase patterns of individual categories can assist marketers in managing relationships with customers. Our modeling approach goes beyond extant “buy ‘til you die” models that focus on a firm-level or a single type of transactional activity by enabling firms to assess customers’ likely future behavior at the more granular category level. Should firms be interested in making decisions based on customers’ behavior across multiple categories, the category-level results from our framework can be aggregated. Our model also allows marketers to identify which customers are more likely to purchase in a given set of categories on a particular visit, enabling the firm to dynamically tailor its marketing efforts to customers based on their purchase history in the categories.

There are several limitations in the current work that can be addressed in future research. First, given the limited extent of latent customer attrition inferred in our data, we model customer attrition at the firm level and employ the time-invariant attrition process. However, in other empirical contexts where customer attrition is more prevalent, customer attrition may vary across transactions depending on shopping basket composition. In particular, Schweidel, Park and Jamal (2013) provide a multi-activity latent attrition model for customer base analysis. Their empirical findings suggest that the latent attrition processes are related, indicating that conducting one type of activity is informative of whether customers are still “alive” to conduct another type of activity, and consequently impacts inferences of customer value. It would be fruitful to incorporate different sources of attrition into our proposed model and untangle their effects on future behavior. Second, one could extend our modeling framework by modeling customers’ purchase amount decisions across categories. As our model extends the “buy ‘til you die” framework commonly used for customer valuation to a multivariate context, incorporating the amount model would allow one
to quantify the contribution of individual categories to customer lifetime value (CLV) and assess the relationship between multi-category shopping behavior and CLV, which could assist the firm in allocating marketing resources across categories within individual customers.

Third, given the focus of this research and due to the limitation of the data, we have not considered the impact of marketing mix variables on customers’ shopping behavior. In practice, firms often implement several concurrent marketing programs at different levels. Some of them may be store-wide campaigns which affect both purchase timing and choice decisions across all categories (albeit with different degrees of effectiveness), while others focus on specific product categories and influence purchase behavior in not only promoted ones but also related others. To assess the impacts of such different marketing efforts on customers’ purchase behavior and thus the firm’s revenue, the interplay between the interarrival times and shopping basket composition as well as the cross-category dependence, need to be considered. For example, store-wide campaigns may accelerate customers’ shopping frequency. As a result of the increased interarrival time, however, their probabilities of purchasing categories on a given trip may decrease. Similarly, several concurrent category-level promotions may help increase the probability of category choice. However, the resulting large shopping baskets could lead to a longer interarrival time until the customers’ next shopping trip. Failing to consider such interactions could result in erroneous estimates of the impact of the firm’s marketing activities. While we do not have access to such data in this empirical application, our modeling framework provides a general platform for the inclusion of different types of marketing covariates and thus allows one to disentangle their effects on customers’ purchase patterns across categories. Our model can be also modified to bring in the methods proposed by Gupta (1991), who demonstrated a sophisticated way of incorporating time-varying marketing mix variables and customers’ product inventory levels into a multi-event timing model. We hope this study generates further interest and accelerate the progress in this important area of research.
References


Appendix: Derivation of Eq. (8)

Let us denote the probability of the “no-purchase” case derived using Eq. (6) as

\[
\Pr(C_{ij1}^0, C_{ij2}^0, \ldots, C_{ijK}^0) = \Pr(C_{ij1} = 0, C_{ij2} = 0, \ldots, C_{ijK} = 0). \tag{A.1}
\]

Then, from the theorem by Besag (1974), we have

\[
\frac{\Pr(C_{ij1}, C_{ij2}, \ldots, C_{ijK})}{\Pr(C_{ij1}^0, C_{ij2}^0, \ldots, C_{ijK}^0)} = \frac{\Pr(C_{ij1} | C_{ij2}^0, C_{ij3}, \ldots, C_{ijK})}{\Pr(C_{ij1} | C_{ij1}^0, C_{ij3}, \ldots, C_{ijK})} \cdot \frac{\Pr(C_{ij2} | C_{ij1}^0, C_{ij2}^0, C_{ij3}, \ldots, C_{ijK})}{\Pr(C_{ij2} | C_{ij1}^0, C_{ij2}, C_{ij3}, \ldots, C_{ijK})} \cdot \ldots \cdot \frac{\Pr(C_{ijK} | C_{ij1}^0, C_{ij2}^0, \ldots, C_{ijK}^0)}{\Pr(C_{ijK} | C_{ij1}^0, C_{ij2}, \ldots, C_{ijK}^0)}. \tag{A.2}
\]

By substituting Eq. (6) into Eq. (A.2), we have

\[
\Pr(C_{ij1}, C_{ij2}, \ldots, C_{ijK}) = \Pr(C_{ij1}^0, C_{ij2}^0, \ldots, C_{ijK}^0) \cdot \{\exp(\pi_{ij1} + \Theta_i C_{ij} + \Psi_i C_{ij-1})\} C_{ij1} \cdot \{\exp(\pi_{ij2} + \Theta_2 C_{ij}^{(023\ldots K)} + \Psi_2 C_{ij-1})\} C_{ij2} \cdot \ldots \cdot \{\exp(\pi_{ijK} + \Theta_K C_{ij}^{(000\ldots K)} + \Psi_K C_{ij-1})\} C_{ijK}, \tag{A.3}
\]

where \( \Theta_k \) and \( \Psi_k \) are the \( k \)th row vector of matrices \( \Theta \) and \( \Psi \), respectively, and

\[
C_{ij}^{(023\ldots K)} = (0, C_{ij2}, C_{ij3}, \ldots, C_{ijK}) \text{ and } C_{ij}^{(000\ldots K)} = (0, 0, 0, \ldots, C_{ijK}).
\]

Eq. (A.3) can be written as:

\[
\Pr(C_{ij1}, C_{ij2}, \ldots, C_{ijK}) = \Pr(C_{ij1}^0, C_{ij2}^0, \ldots, C_{ijK}^0) \cdot \exp\{C_{ij1}(\pi_{ij1} + \Theta_i C_{ij} + \Psi_i C_{ij-1})\} \cdot \exp\{C_{ij2}(\pi_{ij2} + \Theta_2 C_{ij}^{(023\ldots K)} + \Psi_2 C_{ij-1})\} \cdot \ldots \cdot \exp\{C_{ijK}(\pi_{ijK} + \Theta_K C_{ij}^{(000\ldots K)} + \Psi_K C_{ij-1})\}. \tag{A.4}
\]

By arranging exponential terms in Eq. (A.4), we have

\[
\Pr(C_{ij1}, C_{ij2}, \ldots, C_{ijK}) = \Pr(C_{ij1}^0, C_{ij2}^0, \ldots, C_{ijK}^0) \cdot \exp(\pi_{ij}/C_{ij} + \frac{1}{2} C_{ij}^* \Theta'C_{ij} + C_{ij-1}' \Psi C_{ij-1}). \tag{A.5}
\]

Because the sum of the joint probability in Eq. (A.5) across all possible combinations of categories should be equal to one and our data do not include the “no-purchase” case, the joint distribution of \( C_{ij} = \{C_{ij1}, C_{ij2}, \ldots, C_{ijK}\}' \) is given by:

\[
P\left(C_{ij} = C_{ij}^*, \sum_{k=1}^{K} C_{ijk}^* \geq 1 \right) = \frac{\exp(\pi_{ij}/C_{ij}^* + \frac{1}{2} C_{ij}^* \Theta'C_{ij}^* + C_{ij-1}' \Psi C_{ij-1})}{\sum_{C_{ij} \neq \{0,0,\ldots,0\}} \exp(\pi_{ij}/C_{ij} + \frac{1}{2} C_{ij} \Theta'C_{ij} + C_{ij-1}' \Psi C_{ij-1})}. \tag{A.6}
\]
<table>
<thead>
<tr>
<th>No. of transactions</th>
<th>Mean</th>
<th>Std. Dev.</th>
</tr>
</thead>
<tbody>
<tr>
<td>No. of categories purchased per transaction</td>
<td>9.06</td>
<td>5.19</td>
</tr>
<tr>
<td>No. of purchases of basics</td>
<td>1.64</td>
<td>0.80</td>
</tr>
<tr>
<td>No. of purchases of creams</td>
<td>3.84</td>
<td>2.97</td>
</tr>
<tr>
<td>No. of purchases of serums</td>
<td>4.62</td>
<td>3.01</td>
</tr>
<tr>
<td>No. of purchases of makeups</td>
<td>3.97</td>
<td>3.03</td>
</tr>
<tr>
<td>No. of different compositions of shopping baskets purchased</td>
<td>2.46</td>
<td>2.71</td>
</tr>
</tbody>
</table>

Table 1: Descriptive Statistics

<table>
<thead>
<tr>
<th>Shopping basket</th>
<th>No. of observations</th>
<th>No. of customers who purchased the shopping basket</th>
<th>Time taken until the next transaction (days)</th>
</tr>
</thead>
<tbody>
<tr>
<td>{1,0,0,0}</td>
<td>3,579</td>
<td>1,457</td>
<td>73.6</td>
</tr>
<tr>
<td>{0,1,0,0}</td>
<td>4,489</td>
<td>1,789</td>
<td>73.6</td>
</tr>
<tr>
<td>{0,0,1,0}</td>
<td>3,383</td>
<td>1,418</td>
<td>70.8</td>
</tr>
<tr>
<td>{0,0,0,1}</td>
<td>2,569</td>
<td>1,123</td>
<td>66.3</td>
</tr>
<tr>
<td>{1,1,0,0}</td>
<td>1,956</td>
<td>1,130</td>
<td>83.6</td>
</tr>
<tr>
<td>{1,0,1,0}</td>
<td>1,506</td>
<td>919</td>
<td>79.4</td>
</tr>
<tr>
<td>{1,0,0,1}</td>
<td>559</td>
<td>387</td>
<td>71.9</td>
</tr>
<tr>
<td>{0,1,1,0}</td>
<td>2,149</td>
<td>1,228</td>
<td>84.4</td>
</tr>
<tr>
<td>{0,1,0,1}</td>
<td>975</td>
<td>661</td>
<td>73.6</td>
</tr>
<tr>
<td>{0,0,1,1}</td>
<td>712</td>
<td>492</td>
<td>71.0</td>
</tr>
<tr>
<td>{1,1,1,0}</td>
<td>1,880</td>
<td>1,151</td>
<td>95.3</td>
</tr>
<tr>
<td>{1,1,0,1}</td>
<td>480</td>
<td>349</td>
<td>80.0</td>
</tr>
<tr>
<td>{1,0,1,1}</td>
<td>430</td>
<td>339</td>
<td>80.2</td>
</tr>
<tr>
<td>{0,1,1,1}</td>
<td>694</td>
<td>477</td>
<td>84.6</td>
</tr>
<tr>
<td>{1,1,1,1}</td>
<td>640</td>
<td>462</td>
<td>87.8</td>
</tr>
</tbody>
</table>

Table 2: Shopping Basket and Interarrival Time
### Table 3: Model Comparison

<table>
<thead>
<tr>
<th>Model</th>
<th>Effect of shopping baskets on the arrival rates</th>
<th>Interdependence across category choices</th>
<th>Temporal dependence across category choices</th>
<th>Number of model parameters</th>
<th>Log marginal density</th>
<th>In-sample MAE</th>
<th>Out-of-sample MAE</th>
</tr>
</thead>
<tbody>
<tr>
<td>Model 1</td>
<td>×</td>
<td>×</td>
<td>×</td>
<td>16</td>
<td>-94,161</td>
<td>0.66</td>
<td>0.30</td>
</tr>
<tr>
<td>Model 2</td>
<td>×</td>
<td>×</td>
<td>×</td>
<td>27</td>
<td>-93,946</td>
<td>0.62</td>
<td>0.28</td>
</tr>
<tr>
<td>Model 3</td>
<td>×</td>
<td>×</td>
<td>×</td>
<td>40</td>
<td>-93,853</td>
<td>0.61</td>
<td>0.27</td>
</tr>
<tr>
<td>Model 4</td>
<td>×</td>
<td>×</td>
<td>✓</td>
<td>47</td>
<td>-93,742</td>
<td>0.60</td>
<td>0.26</td>
</tr>
<tr>
<td>Model 5</td>
<td>×</td>
<td>✓</td>
<td>✓</td>
<td>51</td>
<td>-93,178</td>
<td>0.57</td>
<td>0.24</td>
</tr>
<tr>
<td>Model 6</td>
<td>✓</td>
<td>×</td>
<td>×</td>
<td>43</td>
<td>-93,828</td>
<td>0.61</td>
<td>0.27</td>
</tr>
<tr>
<td>1 dim. space</td>
<td></td>
<td></td>
<td></td>
<td>46</td>
<td>-93,562</td>
<td>0.59</td>
<td>0.26</td>
</tr>
<tr>
<td>2 dim. space</td>
<td></td>
<td></td>
<td></td>
<td>48</td>
<td>-93,398</td>
<td>0.58</td>
<td>0.25</td>
</tr>
<tr>
<td>3 dim. space</td>
<td></td>
<td></td>
<td></td>
<td>49</td>
<td>-93,347</td>
<td>0.58</td>
<td>0.25</td>
</tr>
<tr>
<td>4 dim. space</td>
<td></td>
<td></td>
<td></td>
<td>54</td>
<td>-93,171</td>
<td>0.57</td>
<td>0.24</td>
</tr>
<tr>
<td>Model 7</td>
<td>Main effects only</td>
<td>✓</td>
<td>✓</td>
<td>72</td>
<td>-92,413</td>
<td>0.51</td>
<td>0.21</td>
</tr>
<tr>
<td>Model 8</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>54</td>
<td>-93,171</td>
<td>0.57</td>
<td>0.24</td>
</tr>
<tr>
<td>Model 9</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>72</td>
<td>-92,413</td>
<td>0.51</td>
<td>0.21</td>
</tr>
<tr>
<td>1 dim. space</td>
<td></td>
<td></td>
<td></td>
<td>56</td>
<td>-93,094</td>
<td>0.57</td>
<td>0.24</td>
</tr>
<tr>
<td>2 dim. space</td>
<td></td>
<td></td>
<td></td>
<td>59</td>
<td>-92,802</td>
<td>0.54</td>
<td>0.22</td>
</tr>
<tr>
<td>3 dim. space</td>
<td></td>
<td></td>
<td></td>
<td>61</td>
<td>-92,602</td>
<td>0.52</td>
<td>0.21</td>
</tr>
<tr>
<td>4 dim. space</td>
<td></td>
<td></td>
<td></td>
<td>62</td>
<td>-92,543</td>
<td>0.52</td>
<td>0.21</td>
</tr>
</tbody>
</table>

### Table 4a: Parameter Estimates of the Timing and Attrition Models

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Posterior mean</th>
<th>95% Posterior interval</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\mu_\lambda$</td>
<td>-2.47</td>
<td>[-2.51, -2.42]</td>
</tr>
<tr>
<td>$\mu_\omega$</td>
<td>-4.91</td>
<td>[-5.18, -4.65]</td>
</tr>
<tr>
<td>$\Sigma_{\lambda\lambda}$</td>
<td>0.32</td>
<td>[0.24, 0.39]</td>
</tr>
<tr>
<td>$\Sigma_{\omega\omega}$</td>
<td>0.21</td>
<td>[0.06, 0.36]</td>
</tr>
<tr>
<td>$LIT_{ij}$</td>
<td>$-5.18 \times 10^{-4}$</td>
<td>[-2.16 $\times 10^{-3}$, 1.09 $\times 10^{-3}$]</td>
</tr>
</tbody>
</table>
Table 4b: Parameter Estimates of the Timing Model: Coordinates of Categories in a Three-Dimensional Latent Space

<table>
<thead>
<tr>
<th></th>
<th>Coordinates</th>
<th>1st</th>
<th>2nd</th>
<th>3rd</th>
</tr>
</thead>
<tbody>
<tr>
<td>Basics</td>
<td>µ_{β1}</td>
<td>-1.35</td>
<td>[-1.48,-1.24]</td>
<td></td>
</tr>
<tr>
<td></td>
<td>C_{i,j-1,1}</td>
<td>-0.29</td>
<td>[-0.38,-0.19]</td>
<td></td>
</tr>
<tr>
<td></td>
<td>ET_{ij1}</td>
<td>0.03</td>
<td>[0.02, 0.04]</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(ET_{ij1})^{2}</td>
<td>-2.35×10^{-4}</td>
<td>[-3.00×10^{-4},-1.72×10^{-4}]</td>
<td></td>
</tr>
<tr>
<td>Creams</td>
<td>µ_{β2}</td>
<td>-0.93</td>
<td>[-1.02,-0.82]</td>
<td></td>
</tr>
<tr>
<td></td>
<td>C_{i,j-1,2}</td>
<td>-0.34</td>
<td>[-0.43,-0.25]</td>
<td></td>
</tr>
<tr>
<td></td>
<td>ET_{ij2}</td>
<td>0.02</td>
<td>[0.01, 0.03]</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(ET_{ij2})^{2}</td>
<td>-2.36×10^{-4}</td>
<td>[-3.07×10^{-4},-1.65×10^{-4}]</td>
<td></td>
</tr>
<tr>
<td>Serums</td>
<td>µ_{β3}</td>
<td>-1.49</td>
<td>[-1.62,-1.37]</td>
<td></td>
</tr>
<tr>
<td></td>
<td>C_{i,j-1,3}</td>
<td>-0.31</td>
<td>[-0.39,-0.22]</td>
<td></td>
</tr>
<tr>
<td></td>
<td>ET_{ij3}</td>
<td>0.02</td>
<td>[0.02, 0.03]</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(ET_{ij3})^{2}</td>
<td>-2.67×10^{-4}</td>
<td>[-3.33×10^{-4},-2.02×10^{-4}]</td>
<td></td>
</tr>
<tr>
<td>Makeups</td>
<td>µ_{β4}</td>
<td>-2.13</td>
<td>[-2.25,-2.00]</td>
<td></td>
</tr>
<tr>
<td></td>
<td>C_{i,j-1,4}</td>
<td>0.11</td>
<td>[-0.01, 0.23]</td>
<td></td>
</tr>
<tr>
<td></td>
<td>ET_{ij4}</td>
<td>0.02</td>
<td>[0.01, 0.02]</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(ET_{ij4})^{2}</td>
<td>-1.71×10^{-4}</td>
<td>[-2.25×10^{-4},-1.16×10^{-4}]</td>
<td></td>
</tr>
</tbody>
</table>

Note: The numbers in brackets denote the 95% posterior interval.

Table 5a: Parameter Estimates of the Choice Model: Category-Specific Effects
<table>
<thead>
<tr>
<th>Posterior mean</th>
<th>95% Posterior interval</th>
</tr>
</thead>
<tbody>
<tr>
<td>ζ₀₁</td>
<td>0.38</td>
</tr>
<tr>
<td></td>
<td>[0.28, 0.47]</td>
</tr>
<tr>
<td>ζ₀₂</td>
<td>0.41</td>
</tr>
<tr>
<td></td>
<td>[0.35, 0.47]</td>
</tr>
<tr>
<td>ζ₀₃</td>
<td>0.58</td>
</tr>
<tr>
<td></td>
<td>[0.52, 0.65]</td>
</tr>
<tr>
<td>ζ₀₄</td>
<td>-0.05</td>
</tr>
<tr>
<td></td>
<td>[-0.13, 0.03]</td>
</tr>
<tr>
<td>ζ₁</td>
<td>-0.15</td>
</tr>
<tr>
<td></td>
<td>[-0.30, 0.01]</td>
</tr>
</tbody>
</table>

Table 5b: Parameter Estimates of the Choice Model: Cross-Category Effects

<table>
<thead>
<tr>
<th>Basics</th>
<th>Creams</th>
<th>Serums</th>
<th>Makeups</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>0.57</td>
<td>0.73</td>
<td>-0.04</td>
</tr>
<tr>
<td></td>
<td>[0.51, 0.63]</td>
<td>[0.66, 0.79]</td>
<td>[-0.10, 0.03]</td>
</tr>
<tr>
<td>Creams</td>
<td>0.74</td>
<td>0.15</td>
<td></td>
</tr>
<tr>
<td></td>
<td>[0.66, 0.81]</td>
<td>[0.09, 0.21]</td>
<td></td>
</tr>
<tr>
<td>Serums</td>
<td></td>
<td>0.31</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>[0.25, 0.37]</td>
<td></td>
</tr>
</tbody>
</table>

Note: The numbers in brackets denote the 95% posterior interval.

Table 5c: Estimated Θ

<table>
<thead>
<tr>
<th>Posterior mean</th>
<th>95% Posterior interval</th>
</tr>
</thead>
<tbody>
<tr>
<td>ξ₀₁</td>
<td>0.00</td>
</tr>
<tr>
<td></td>
<td>—</td>
</tr>
<tr>
<td>ξ₀₂</td>
<td>0.03</td>
</tr>
<tr>
<td></td>
<td>[-0.06, 0.13]</td>
</tr>
<tr>
<td>ξ₀₃</td>
<td>0.12</td>
</tr>
<tr>
<td></td>
<td>[0.02, 0.23]</td>
</tr>
<tr>
<td>ξ₀₄</td>
<td>0.12</td>
</tr>
<tr>
<td></td>
<td>[0.03, 0.20]</td>
</tr>
<tr>
<td>ξ₁₁</td>
<td>-0.10</td>
</tr>
<tr>
<td></td>
<td>[-0.32, 0.12]</td>
</tr>
<tr>
<td>ξ₁₂</td>
<td>-0.12</td>
</tr>
<tr>
<td></td>
<td>[-0.30, 0.06]</td>
</tr>
<tr>
<td>ξ₁₃</td>
<td>-0.04</td>
</tr>
<tr>
<td></td>
<td>[-0.21, 0.13]</td>
</tr>
<tr>
<td>ξ₁₄</td>
<td>-0.21</td>
</tr>
<tr>
<td></td>
<td>[-0.42, -0.01]</td>
</tr>
<tr>
<td>ξ₂</td>
<td>-0.07</td>
</tr>
<tr>
<td></td>
<td>[-0.18, 0.04]</td>
</tr>
</tbody>
</table>

Table 5d: Parameter Estimates of the Choice Model: Lagged Cross-Category Effects
<table>
<thead>
<tr>
<th>Basics</th>
<th>Creams</th>
<th>Serums</th>
<th>Makeups</th>
</tr>
</thead>
<tbody>
<tr>
<td>Basics</td>
<td>-0.20</td>
<td>-0.12</td>
<td>-0.34</td>
</tr>
<tr>
<td>[0.27,-0.13]</td>
<td>[-0.19,-0.04]</td>
<td>[-0.42,-0.26]</td>
<td></td>
</tr>
<tr>
<td>Creams</td>
<td>-0.15</td>
<td>-0.10</td>
<td>-0.25</td>
</tr>
<tr>
<td>[0.21,-0.08]</td>
<td>[-0.18,-0.02]</td>
<td>[-0.32,-0.18]</td>
<td></td>
</tr>
<tr>
<td>Serums</td>
<td>-0.06</td>
<td>-0.09</td>
<td>-0.17</td>
</tr>
<tr>
<td>[0.13, 0.01]</td>
<td>[-0.17,-0.02]</td>
<td>[-0.25,-0.09]</td>
<td></td>
</tr>
<tr>
<td>Makeups</td>
<td>-0.10</td>
<td>-0.07</td>
<td>0.01</td>
</tr>
<tr>
<td>[0.20, 0.01]</td>
<td>[-0.15, 0.02]</td>
<td>[-0.07, 0.08]</td>
<td></td>
</tr>
</tbody>
</table>

Note: The numbers in brackets denote the 95% posterior interval.

Table 5e: Estimated $\Psi$

<table>
<thead>
<tr>
<th>log $\lambda_i$</th>
<th>$\omega_i$</th>
<th>$\beta_{0i1}$</th>
<th>$\beta_{0i2}$</th>
<th>$\beta_{0i3}$</th>
<th>$\beta_{0i4}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>log $\lambda_i$</td>
<td>0.31</td>
<td>-0.05</td>
<td>-0.14</td>
<td>-0.07</td>
<td>0.04</td>
</tr>
<tr>
<td>$\omega_i$</td>
<td>[0.23, 0.38]</td>
<td>[-0.09,-0.01]</td>
<td>[-0.11,-0.05]</td>
<td>[-0.17,-0.11]</td>
<td>[-0.10,-0.04]</td>
</tr>
<tr>
<td>$\beta_{0i1}$</td>
<td>0.21</td>
<td>0.29</td>
<td>0.24</td>
<td>0.26</td>
<td>0.19</td>
</tr>
<tr>
<td>$\beta_{0i2}$</td>
<td>0.90</td>
<td>0.30</td>
<td>0.12</td>
<td>0.27</td>
<td></td>
</tr>
<tr>
<td>$\beta_{0i3}$</td>
<td>[0.77, 1.04]</td>
<td>[0.21, 0.39]</td>
<td>[0.04, 0.21]</td>
<td>[0.18, 0.38]</td>
<td></td>
</tr>
<tr>
<td>$\beta_{0i4}$</td>
<td>[0.78, 1.04]</td>
<td>[0.15, 0.34]</td>
<td>[0.27, 0.47]</td>
<td></td>
<td></td>
</tr>
<tr>
<td>$\beta_{0i5}$</td>
<td>0.89</td>
<td>0.21</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$\beta_{0i6}$</td>
<td>[0.76, 1.02]</td>
<td>[0.12, 0.30]</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Note: The numbers in brackets denote the 95% posterior interval.

Table 6: Estimated $\Sigma$
Fig. 1: Scatter Plot between the No. of Transactions and the Mean No. of Categories Purchased per Transaction
Fig. 2: Two-Category Representation in a Two-Dimensional Space
Fig. 3: Model Prediction on Cumulative Transactions
Fig. 4: Model Prediction on Mean Interarrival Time after Purchasing a Shopping Basket
Fig. 5: Model Prediction on Cumulative Purchases of a Shopping Basket

Note: (1) In all charts, the x-axis is week and the y-axis is the cumulative purchases of a shopping basket.
(2) The solid and dotted lines represent the observed and predicted values, respectively.

Fig. 5: Model Prediction on Cumulative Purchases of a Shopping Basket
Fig. 6: Model Prediction on Category Purchases
Fig. 7: Estimated Positions of Categories in a Latent Space