

Dynamically Managing a Profitable Email Marketing Program

Xi (Alan) Zhang

V Kumar

Koray Cosguner

Forthcoming at Journal of Marketing Research

doi: <http://dx.doi.org/10.1509/jmr.16.0210>

Xi (Alan) Zhang is an Assistant Professor in Marketing at College of Business and Innovation, University of Toledo. Toledo, OH 43606; Tel: 419.530.5516; Email: xi.zhang2@utoledo.edu.

V. Kumar is the Regents' Professor, Richard and Susan Lenny Distinguished Chair & Professor of Marketing, Executive Director of Center for Excellence in Brand & Customer Management, and Director of the Ph.D. Program in Marketing at J. Mack Robinson College of Business, Georgia State University. Atlanta, GA 30326; Chang Jiang Scholar, HUST, China, and Lee Kong Chian Fellow, Singapore Management University, Singapore. Tel: 404.413.7590; Email: vk@gsu.edu.

Koray Cosguner is an Assistant Professor in Marketing at J. Mack Robinson College of Business, Georgia State University. Atlanta, GA 30303; Tel: 404.413.7697; Email: kcosguner@gsu.edu.

All authors contributed equally.

Dynamically Managing a Profitable Email Marketing Program

ABSTRACT

Although email-marketing is highly profitable and widely used by marketers, it has received limited attention in marketing literature. Extant research either focuses on customers' email responses or studies the "average" effect of emails on purchases. In this paper, we use data from a U.S. home improvement retailer to study customers' email open and purchase behaviors by using a unified Hidden Markov and Copula framework. Contrary to conventional wisdom, we find that email active customers are not necessarily active in purchases, and vice versa. Furthermore, we find that the number of emails sent by the retailer has a non-linear effect on both the retailer's short- and long-term profitability. Through a counterfactual study, we provide a decision support system to guide the retailer to make optimal email contact decisions. This study shows that sending the right number of emails is vital for long-term profitability. For example, sending 4 (10) emails instead of the optimal number of 7 emails can cause the retailer to lose 32% (16%) of its lifetime profit per customer.

Keywords: Email marketing, Hidden Markov models, Dynamic programming, Copula models

Introduction

Email marketing is a widely-used marketing tool by most B2B and B2C companies. A survey study by Ascend2 (2015) shows that 82% of B2B and B2C companies use email technology as part of their marketing strategies. In addition to its widespread usage, email marketing can be quite profitable. According to a recent study published by the Direct Marketing Association (UK), the average revenue-based return on investment (ROI) of email marketing increased to £38 for each £1 spent compared to the average ROI of £24.93 reported in 2013 (DMA 2015). Based on the same study, roughly 1 in every 5 companies report an ROI of more than £70 which is three times higher than the figures from 2013. Without a doubt, firms aim to generate higher ROI levels from email technology by launching effective email marketing programs. Although the firm's goal is evident, firms may find it difficult to manage effective email programs for multiple reasons.

First, customers who are active in opening a firm's emails may not be active in purchasing from the firm or vice versa. For example, some heavy buyers may not open the firm's emails frequently since they already know the firm and its offerings quite well. This implies that the firm may not need to target such customers aggressively through email marketing since this type of buyer may pay little to no attention to emails received. On the other hand, some light buyers might be quite responsive in opening the firm's emails since they seek to get information through emails. Even though this type of customer opens emails more frequently, from the profit perspective, targeting such customers might be suboptimal for the firm since the chances of converting such customers to repeat-buyers are likely low.

Second, sending the right number of emails is critical for the firm's profitability, especially since most customers tend to complain about the large number of emails sent by firms. A survey study by BlueHornet (2013), a marketing firm focusing on email solutions, finds that email

frequency might overtake content irrelevance as an important reason that drives customers away from email marketing. Although sending the right number of emails is paramount, finding that magic number is very challenging for a firm since not only do its customers have different intrinsic preferences to emails (i.e., email open¹ frequencies differ across customers), but customer preferences might change dynamically over time (i.e., within same customer, email open frequencies might differ over time). In other words, the right number of emails to send may differ across customers and over time. Consistent with this argument, a study by Return Path (2015), an industry expert company on email optimization, suggests that email frequency optimization should depend on the engagement level of customers. However, the study does not provide an actionable tool for firms to adopt to launch effective email marketing programs.

Third, managers tend to treat each email marketing campaign as an independent solicitation process and fail to consider its long-term impact on the email open and purchase behaviors. In addition to immediate impact, an email might also affect the future profitability of the firm by shifting the customer-firm relationship level (Luo and Kumar 2013). For example, customers opening emails might become more interested with the firm and its offerings, and ultimately, they might start to purchase more frequently. In this case, ignoring the long-term effects of emails might cause the firm to make suboptimal email marketing decisions.

Due to all these challenges, firms need guidance in the form of a decision support system (DSS, hereafter) to decide the right number of emails to send to different customers over time. In this study, our objective is to provide such a DSS to firms through 1) linking customers' email open behaviors with their purchase behaviors, 2) capturing heterogeneity in email open and purchase behaviors across customers and over time, 3) incorporating the long-term effects of email

¹ For this study, we define "email open" as the action of opening an email message.

marketing on firm profitability.

Methodologically, we propose a unified framework combining hidden Markov (HMM) (Netzer, Lattin and Srinivasan 2008) and Copula models (Danaher and Smith 2011). In the hidden Markov component of our model, we capture the dynamics in customers' purchase and email behaviors by allowing the latent relationship states to govern both behaviors. As later discussed, modeling the evolution of purchase and email states jointly is important in the context of email marketing since customers' responsiveness in opening emails might also affect their relationship with the firm. In this regard, our study differs from the other studies in the HMM domain which only focus on the customer's purchase behavior and ignores the fact that other non-purchase activities may also affect the customer-firm relationship (except the study by Schweidel et al. 2014) level. In addition, customers' email open and purchase behaviors might be correlated due to time-invariant factors like customer unobserved heterogeneity and time-variant factors like the unobserved customer-firm relationship states. In the Copula component of our model, we use a bivariate Frank Copula to capture that correlation. We model the email open behavior using a binomial distribution (BD, hereafter) and the purchase behavior using a zero-inflated negative binomial distribution (ZINBD, hereafter). In this regard, the Copula component of our model also differs from most Copula applications in marketing where Copulas are used to correlate variables coming from continuous distribution families (except the study by Stephen and Galak (2012), in which double Poisson variables are first converted into continuous ones and treated as continuous in the estimation).

Substantively, we seek to answer the following questions:

- (1) How the latent Markov states govern customers' purchase and email open behaviors,
- (2) Whether there is any correlation between customers' email open and purchase behaviors,

(3) How email contacts (i.e., number of emails sent by a firm) affect customers' purchase behavior both in the short and long run, and

(4) What the optimal number of emails is to send to different customers over time to maximize the long-term firm profit?

The empirical study identifies three customer-firm relationship states governing customers' and email open and purchase behaviors: Low Open – Medium Purchase, High Open – Low Purchase and High Open – High Purchase states. In other words, email active customers may not be active in purchases; and email inactive customers may still be relatively active in their purchase behaviors. In addition, we find that the firm's email contact policy has non-linear short- and long-term effects on customers' purchase and email open behaviors as well as on the retailer's profitability. Furthermore, from our Copula component, we identify a positive correlation between the two customer behaviors. We next derive a dynamic email marketing resource allocation policy using the estimates from our proposed hidden Markov and Copula model of customer purchase and email open behaviors. This way, we provide a DSS to the firm (about the optimal number of emails to send) to maximize its lifetime profit. Based on our DSS, we find that the optimal number of emails to send not only differs across customers, but also differs within the same customer over time significantly. In addition, we find that sending the right number of emails is very critical from the firm's profit perspective. For instance, sending 4 (10) rather than the optimal number of 7 emails per month causes the retailer to lose 32% (16%) of its lifetime profit.

To reiterate our contribution, to the best of our knowledge, this is the first empirical study that jointly models customers' email open and purchase behaviors over time. Methodologically, our paper is the first study that combines the HMM and Copula approaches together in a unified framework. Therefore, this research provides not only important implications for firms to

understand customers' behavioral attitudes towards email marketing, but also an easy-to-implement modeling framework to marketing researchers in the field. The proposed framework is flexible and marketing researchers in the field can use it to study other non-purchase customer behaviors (including but not limited to customers' responses to other electronic and potentially interactive communications) that are possibly correlated with purchase. The framework can incorporate, across customers and over time, unobserved behavioral heterogeneity. Finally, through our DSS, we provide substantive managerial guidelines to firms to implement an effective email-marketing program. Similarly, our suggested DSS is also versatile and firms can use it to manage marketing decisions such as direct mails, telephone and sales person contacts.

In the following sections, we first review the literature on email marketing and customer relationship dynamics in email marketing. Second, we describe our data and present descriptive statistics. Third, we discuss our modeling framework. Fourth, we present the empirical results. Fifth, we derive the optimal email marketing contact policy for the firm. Sixth, we discuss the robustness of our proposed model. Finally, we summarize our findings, and conclude with caveats and directions for future research.

Literature Review

Email Marketing

There are several reasons for email marketing's popularity. First, emails enable marketers to send messages to their customers at very low cost. Chittenden and Rettie (2003) demonstrate that the total cost for acquisition and retention campaigns of email to be \$26,500 per 5,000 customers, as compared to that of direct mail at \$69,600 per 5,000 customers. Second, email messaging requires less preparation and execution time. Industry practice shows that an email marketing campaign targeting 50,000 customers needs only 6 hours to prepare and run; while a

similar size direct mail campaign needs 17 days before it can reach a targeted customer's mailbox. Third, emails typically generate faster response and create opportunity for interactive communication with customers. For example, a customer can respond to a email the moment he/she receives it by clicking the hyperlinks that direct him/her to the sender firm's website via his/her computer or mobile device.

In this study, we focus on permission-based email marketing. Permission-based email marketing requires marketers to seek the customers' permission before sending them email messages (Godin 1999; Kumar, Zhang and Luo 2014). This type of email marketing intends to maintain a repeat purchase relationship with customers, rather than getting customers to buy one time only. In line with this idea, previous research shows that email marketing has a positive effect on customer loyalty. Tezinde et al. (2002) discover that email advertisements are useful to customers by inducing them to visit the physical store. Merisavo and Raulas (2004) find that email marketing can enhance customer attitudinal loyalty towards the brand. Their study shows that customers tend to recommend the email messages to their friends if they find the messages interesting and useful.

Although the overall influence of email marketing is positive, we argue that researchers and practitioners should examine the customers' response to email-based marketing messages based on two perspectives. First, customers open and read an email simply to keep track of the firm's products and offerings. This behavior does not necessarily indicate that they are actively looking for information to assist their purchase decisions. Bonfrer and Dreze (2009) study a series of email marketing campaigns and propose a bivariate hazard model to predict when customers open or click an email. Kumar, Zhang and Luo (2014) look at the total number of emails that are opened and clicked and investigate their impact on the time the customers subscribe to the email

program. Second and more importantly, customers make purchases because of email marketing messages. Sahni, Zou and Chintagunta (2016) analyze 70 randomized field-experiments and find that email promotions not only increase customers' average purchase spending during the promotion window, but also carry over to a week later after the promotion expires. Kumar, Zhang and Luo (2014) find that the average email open-rate has a positive effect on average purchase spending while the effect of the average email click-rate is not significant.

Although there are studies that consider each of these two perspectives separately—email open and purchase, it is surprising that there is no study investigating both behaviors together. Bonfrer and Dreze (2009) do not consider the possible link between email open/click rate and purchase behavior due to the limitation of their data. Kumar, Zhang and Luo (2014) capture the “average” effect of the customer's response to emails on purchase, but they do not consider the dynamics and heterogeneity in both email open and purchase behaviors. Sahni, Zou and Chintagunta (2016) conduct a post-hoc analysis of the experiments to show the aggregate-level effects of emails on customer expenditure. They do not quantify how email open behavior affects customer purchase nor do they consider the dynamic effects.

Customer Relationship in Email Marketing

Previous research shows that the customers' relationship with the firm evolves over their lifetime. Netzer et al. (2008) propose a hidden Markov model to study the transitions of customers among their latent relationship states with the firm. Following Netzer et al. (2008), hidden Markov models were widely used by marketing researchers across a wide range of marketing settings (Montoya et al., 2010; Kumar et al., 2011; Li et al., 2011; Luo and Kumar, 2013). This stream of research primarily examines the evolution of customer-firm relationships by only looking at the purchase behavior, whereas we argue that other customer activities along with the purchase, such

as customers' email open behavior, can also reveal useful information. Conceptually, we claim that studying customers' email open and purchase behaviors jointly is necessary for multiple reasons. First, there is a possible correlation between the two behaviors. For example, customers who are active in purchasing from a firm may also actively search and examine information about the firm and its offerings by opening and reading emails. In contrast, customers who are inactive in purchasing from the firm may not be interested in reading emails related to the firm either. Second, the email open behavior might carry relevant information related to the level of customer-firm relationship. In other words, for a customer, being active or passive in opening emails might also affect her relationship level with the firm, and more importantly (through this effect on the relationship level) affect her purchase behavior. For example, assume we have two customers at the same relationship level with the firm and each make the same purchases in the current period. If this was the case, the purchase only model would predict the same level for the customer-firm relationship for the next period. However, if one of the customers is also active in opening emails, due to her email activity, his/her customer-firm relationship level might be different compared to the other customer. A purchase only model cannot capture this difference in the relationship level of the two customers. Due to this reasoning, we characterize the customer-firm relationship by jointly utilizing customer-level email open and purchase information.

To the best of our knowledge, Schweidel et al. (2014) is the only study that models the customer-firm relationship by utilizing both customer-level purchase and non-purchase behaviors. They study the dependence of two customer behaviors, i.e., digital purchase and digital posting, and the evolution of the associated customer-firm relationship. One of the important findings is there is a correlation between the latent attribution processes of the two activities, meaning a customer's involvement in one activity is informative of another. Methodologically, we differ from

Schweidel et al. (2014) since 1) in that we use a Hidden Markov model rather than a latent changepoint framework; 2) we use a Copula model to capture the correlation between purchase and email open behaviors rather than a multivariate choice framework. More importantly, our research differs from Schweidel et al. (2014) in that we additionally provide a DSS that allows firms to allocate resources by utilizing information from two or more customer-based activities. Due to this, ours is the first empirical study modeling customer relationship over multiple customer decisions and which provides a DSS to help the firm maximize its lifetime profit. We summarize the characteristics of the aforementioned studies along with our study in Table 1.

Insert Table 1

To provide an effective DSS to guide a firm to make profitable email marketing decisions, one needs a good predictive model (that will be an input to the DSS) of customer purchase and email open behaviors. A good predictive model must accommodate the following three critical factors. First, across customers, the intensity of purchase and email open behaviors might be different. For example, some customers might regularly check emails (possibly to keep themselves updated with new product offerings or recent promotional offers) but may not be active in purchasing. On the other hand, some customers might be very active in purchases, but may not open emails frequently (they possibly know the firm and its offerings quite well). Due to these inherent differences in behavior across customers, capturing the heterogeneity in customers' email open and purchase behaviors is critical in predicting their behaviors.

Second, customers' email responsiveness and purchase activeness may evolve over time since customers have varying interests and needs at different times. For example, at the initial period of joining the email program, customers could be extremely active in opening emails to know more about the firm but their purchase activities may not catch up with their enthusiasm

with the emails. After spending some time and effort getting to know the firm, customers may become familiar and favorable to the firm. Having that said, they might reduce the amount of attention they pay to the emails but unsurprisingly increase their purchase activities. Therefore, it is also critical for firms to understand such dynamics in both behaviors to predict the customer's email open and purchase levels. Finally, there is a possible correlation between customers' email open and purchase behaviors, and firms need to understand such dependence between the behaviors to predict them accurately. Consequently, a good predictive model (of customers' purchase and email behaviors) must incorporate 1) customer heterogeneity, 2) customer-relationship dynamics; and 3) correlation among customers' behaviors.

Our paper contributes to the customer relationship literature, by jointly studying customers' email open and purchase behaviors in the same modeling framework. Methodologically, we propose a unified Hidden Markov and Copula framework that can capture not only the correlation between customer email open and purchase behaviors but also the evolution of the customer-firm relationship states controlling these two behaviors. Our model is flexible that purchase activeness does not necessarily align with email responsiveness. In other words, customers who are active in purchases may not be active in opening emails, or vice versa. In addition, we account for unobserved customer heterogeneity in both customer activities through our random coefficient specification. After estimation of the proposed model, the estimated model becomes an input to our dynamic programming model, in which we develop a DSS to guide the firm to make profitable email marketing decisions. In the following sections, we first describe our dataset, and next, we discuss our modeling framework.

Data Description

Our database comprises information from a home improvement retailer in the United States.

The retailer sells products and services from multiple categories, such as Kitchen, Plumbing, Electrical, Flooring, Paint, and Outdoors. The dataset consists of information on the purchase transactions made by the customers², the number of emails sent by the firm to these customers, and the customer's email open histories. The average inter-purchase time is approximately 6 months but typically ranges from a couple of weeks to a few months. Because of the variety and complexity of the products, the large number of categories, and the lack of transactional data at the category level, we study the overall purchase behavior instead of the purchase behavior at the category level.

To form a sample of data for model estimation, we randomly select a cohort of 200 customers who opted-in to the retailer's email program in February 2007. These customers have remained in the email program throughout the observation window and have received emails from the firm continuously³. Thus, we have data comprising customers' email open and purchase activities over a period of 39 months. Although we do not have information on the content of the emails, according to the management team of the focal firm, the type of content contained in these emails varies significantly. The emails could be informative of new products or sales events, persuasive featuring key benefits of the products, or a combination of the two types. In any given email, the firm generally features products from multiple categories such as Kitchen, Lumber and Flooring.

In Table 2, we report the descriptive statistics and the correlation matrix of the customer (purchase and email open) and firm (email contacts) level behaviors. On average, the retailer sent 6.90 emails to its customers per month. The customers open 1.64 emails and make 0.69 purchases

² Although customers have the option to buy online and offline, 95% of the observed purchase transactions are offline.

³ We observe that less than 1% of the customers in our database opted-out from the firm's email marketing program. We only include the customers who have stayed in the email program during the observation window.

on average per month. Note that we only count unique email opens because most the emails are only opened once, if opened. The correlation matrix shows that the number of purchases has a weak correlation with the number of emails sent and the number of emails opened.

Insert Table 2

To demonstrate the complexity of customers' purchase and email open behavior, we randomly select three customers from our database, and plot their purchase count and email open frequency over the observation period of 39 months (see Figure 1). Customer 1 is not active in either purchase or opening emails. Customer 1 only makes purchases in three months and opens emails in two months. The time elapsed between the first and second purchase is 21 months. Customer 1 ceases to purchase or open emails after month 26. In comparison, Customer 2 has a more active purchase behavior over the course of 39 months. However, Customer 2 is not equally active in opening emails as he/she only opens emails in two months—month 10 and 11. In addition, we find that Customer 2 decreases his/her purchase behavior after month 19, demonstrating a time-variant purchase behavior. Customer 3 is moderately active in both purchase and opening emails. We observe that the average inter-purchase time of Customer 3 is approximately 4 months.

Insert Figure 1 here

Figure 1 shows that customers' purchase and email open behaviors are both heterogeneous and time-variant, and that purchase and email open behaviors are not perfectly aligned with each other. The observed pattern indicates that the two behaviors may possibly correlate with each other as well. In the subsequent section of our modeling framework, we discuss how we model the heterogeneity, dynamics and the correlation in purchase and email open behaviors.

In addition, to understand the process of customer purchase and email open frequency, we plot the distributions of both behaviors (see Figure 2). The distribution of purchase count shows

that a discrete distribution such as a Poisson process may be able to capture the data generating process. However, the mean (0.68) and the variance (2.64) of the purchase count variable suggest over-dispersion, which violates the underlying assumption of Poisson distribution. Furthermore, we observe an excess of zero purchases (71%) that can affect the estimation of the Poisson model. To account for both the over-dispersion and excess of zeros, we use the zero-inflated negative binomial distribution (ZINBD) to model the purchase count variable.

The distribution of email open count also suggests a discrete distribution. The (unique) email open count is the number of unique emails that are opened conditional on the total number of emails received in each month. We observe that the maximum number of the emails sent (per customer and per month) in our data set is 20 emails. Therefore, the number of unique emails available to a customer to open cannot exceed 20. To capture this process, we use the typical binomial distribution (BD) which captures the number of success (email open) in a sequence of events (receiving emails).

Insert Figure 2 here

Modeling Framework

In line with previous research, we use a hidden Markov model to identify customer-firm relationship states and the transitions among these states. An HMM describes a Markov process using discrete latent states. HMM is a stochastic model used to capture the transition between these latent states and translate these states to observed behaviors. HMMs are widely used in the marketing literature to study customer-firm relationships (e.g., Netzer et al. 2008; Montoya et al. 2010; Kumar et al. 2011; Luo and Kumar 2013). In the context of email marketing, we use a HMM

to study two observed customer behaviors jointly: purchase and email open⁴. In our setting, the latent states of the Markov process translate into different levels of customer-firm relationships that yield different purchase and email open activeness for customers. In addition, since there is a possible correlation between customers' purchase and email open behaviors, we capture the correlation between the two behaviors through a Frank Copula function (see Figure 3 for the graphical illustration of our proposed hidden Markov and Copula model of customer purchase and email open behaviors).

Insert Figure 3 here

In the following section, we first discuss the primitives of our HMM. Second, we specify our conditional (on latent states of the Markov process) email open and purchase models. Third, we illustrate how we capture the correlation between state-dependent email open and purchase behaviors through our Frank Copula component. Fourth, we discuss the model estimation. Finally, we discuss the model identification.

Overview of the Model

Let O_{it} be the number of emails customer i opens at time⁵ t . Let Y_{it} be the number of purchases customer i makes at time t . For customer i , we model the sequence of observations $\{(Y_{i1} = y_{i1}, O_{i1} = o_{i1}), \dots, (Y_{it} = y_{it}, O_{it} = o_{it})\}$ using a HMM characterized by (1) the initial state distribution $(\boldsymbol{\psi}_i)$, (2) a sequence of transition probabilities (\mathbf{Q}_{it}) , and (3) a vector of

⁴ In addition to our earlier discussion (in the Customer Relationship in Email Marketing part of our literature review section) related to the conceptual need of joint modeling, we provide two pieces of empirical evidence in our Web Appendix showing that the joint model is indeed an empirical requirement. First, we conduct a simulation study in which the joint model (bivariate email open and purchase) is compared with the purchase only model (univariate purchase). We find that the estimated (from a simulated dataset generated from the bivariate model) bivariate model recovers the assumed parameters more closely and efficiently compared to the estimated (from the same simulated dataset) univariate model. Second, we show that based on the actual data fit, the estimated (from the observed empirical purchase data) univariate model predicts the purchase behavior by 16% worse than the estimated (from the observed empirical purchase and email open data) bivariate model. See our Web Appendix for further discussion.

⁵ The unit of time is a month in our study. Therefore, we use time t and month t interchangeably throughout the paper.

probabilities that relate the latent states to the observed purchase and email open behaviors (\mathbf{H}_{it}).

The Initial State Distribution

At any given month t , let s denote the level of customer i 's relationship. Let ψ_{is} be the probability that customer i is initially in relationship state s , where $\psi_{is} \geq 0$ and $\sum_{s=1}^{NS} \psi_{is} = 1$, where NS is the number of the latent Markov states. In this study, we assume that all customers start at the lowest relationship state in the first month. Therefore, $\psi_i' = [\psi_{i1}, \psi_{i2}, \dots, \psi_{i,NS}] = [1, 0, \dots, 0]$.

The Markov Chain Transition Matrix

In our proposed HMM framework, we allow customers to transit to any relationship state $s = 1, 2, \dots, NS$. Following Kumar et al. (2011), we use a multinomial logit specification to formulate this transition process. We define the transition matrix as

$$\mathbf{Q}_{i,t-1 \rightarrow t} = \begin{array}{c} \text{State at } t-1 \\ \begin{array}{c} 1 \\ 2 \\ \vdots \\ NS \end{array} \end{array} \begin{array}{c} \text{State at } t \\ \begin{array}{cccc} 1 & 2 & \dots & NS \\ \hline q_{it,1 \rightarrow 1} & q_{it,1 \rightarrow 2} & \dots & q_{it,1 \rightarrow NS} \\ q_{it,2 \rightarrow 1} & q_{it,2 \rightarrow 2} & \dots & q_{it,2 \rightarrow NS} \\ \vdots & \vdots & \ddots & \vdots \\ q_{it,NS \rightarrow 1} & q_{it,NS \rightarrow 2} & \dots & q_{it,NS \rightarrow NS} \end{array} \end{array} \quad (1)$$

$q_{it,s \rightarrow s'}$ is the probability that customer i moves from state s at time $t - 1$ to state s' at time t , where $0 \leq q_{it,s \rightarrow s'} \leq 1, \forall s, s'$, and $\sum_{s'} q_{it,s \rightarrow s'} = 1$. We specify the indirect transition utility of customer i for transitioning from the relationship state s at period $t - 1$ to state s' at time t ($U_{it,s \rightarrow s'}$) as follows:

$$U_{it,s \rightarrow s'} = V_{it,s \rightarrow s'} + e_{it,s \rightarrow s'} \quad (2)$$

where $e_{it,s \rightarrow s'}$ and $V_{it,s \rightarrow s'}$ denote the stochastic and deterministic components of the above

indirect utility, respectively. We assume that $e_{it,s \rightarrow s'}$ for $i=1, \dots, N$, $t=1, \dots, T$, $s, s'=1, \dots, NS$ are distributed *i.i.d.* Gumbel with location 0 and scale 1. We operationalize the deterministic component, $V_{it,s \rightarrow s'}$, as follows:

$$V_{it,s \rightarrow s'} = \alpha_{s \rightarrow s'} + \gamma_{s \rightarrow s'} X_{i,t-1} \quad (3)$$

where $\alpha_{s \rightarrow s'}$ is the intrinsic utility of transitioning from relationship state s to s' at time t . $X_{i,t-1}$ contains the following variables: $I[O_{i,t-1} > 0]$, $I[Y_{i,t-1} > 0]$, $EM_{i,t-1}$, $EM_{i,t-1}^2$; where $I[A]$ is the indicator function that takes the value of 1 when event A occurs and the value of 0 otherwise, $O_{i,t-1}$ is the lagged email open count, $Y_{i,t-1}$ is the lagged purchase count, $EM_{i,t-1}$ is the number of emails sent by the firm at time $t-1$. $\gamma_{s \rightarrow s'} = [\gamma_{1,s \rightarrow s'}, \gamma_{2,s \rightarrow s'}, \gamma_{3,s \rightarrow s'}, \gamma_{4,s \rightarrow s'}]$ is the vector of corresponding response coefficients. We normalize the deterministic utility for customer i at time t to transition to the lowest state ($V_{it,s \rightarrow 1}$) to be zero for the identification purpose, i.e., $U_{it,s \rightarrow 1} = e_{it,s \rightarrow 1}$, where $e_{it,s \rightarrow 1}$ for $i=1, \dots, N$, $t=1, \dots, T$, $s=1, \dots, NS$ are distributed *i.i.d.* Gumbel with location 0 and scale 1. Therefore, the transition probability for customer i , transitioning from state s to s' at time t ($q_{it,s \rightarrow s'}$) becomes the well-known multinomial Logit share function as seen below.

$$q_{it,s \rightarrow s'} = \frac{\exp(V_{it,s \rightarrow s'})}{1 + \sum_{k=2}^{NS} \exp(V_{it,s \rightarrow k})} \quad (4)$$

Next, we discuss our conditional email count (CEOM, hereafter) and purchase count models (CPM, hereafter).

Conditional Email Open Count Model (CEOM)

We assume that the number of emails customer i (in state s) opens at time t follows a BD with parameters EM_{it} and $p_{it|s}$, given by

$$P(O_{it} = o_{it}|s) = \binom{EM_{it}}{o_{it}} p_{it|s}^{o_{it}} (1 - p_{it|s})^{EM_{it}-o_{it}} \quad (5)$$

where $p_{it|s}$ is the conditional probability that customer i (in state s) opens an email in month t . We model $p_{it|s}$ as a function of customers' past email open behavior as follows.

$$p_{it|s} = \frac{\exp(\alpha_{io|s} + \beta_{io} \ln[LO_{it}])}{1 + \exp(\alpha_{io|s} + \beta_{io} \ln[LO_{it}])} \quad (6)$$

where $\alpha_{io|s}$ and β_{io} capture the intrinsic utility of opening an email and the effect of duration dependence, given state s , respectively. We capture the duration dependence through time from the last email open variable, i.e., LO_{it} . We use the natural logarithm of LO_{it} to capture the diminishing effect. For identification purpose, we impose the following restrictions: 1) $\alpha_{io|s+1} = \alpha_{io|s} + \exp(\Delta\alpha_{io|s+1})$, where $\Delta\alpha_{io|s+1}$ is a parameter to estimate from the data, 2) β_{io} is state-invariant. These two restrictions guarantee that customers in a higher relationship state, all else being equal, have a higher probability to open emails than those in a lower state. In addition, we allow both $\alpha_{io|s}$ and β_{io} to be customer-specific to control for the unobserved customer heterogeneity. We assume that $\alpha_{io|s}$ and β_{io} are distributed with the following normal distributions.⁶

$$\begin{aligned} \alpha_{io|s} &= \alpha_{o|s} + \nabla\alpha_{io} \\ \beta_{io} &= \beta_o + \nabla\beta_{io} \end{aligned} \quad (7)$$

where $\nabla\alpha_{io} \sim N(0, \sigma_{\alpha_o}^2)$ and $\nabla\beta_{io} \sim N(0, \sigma_{\beta_o}^2)$.

Conditional Purchase Count Model (CPM)

⁶ We acknowledge that the random coefficient specification we implemented here might not fully capture all the cross-sectional heterogeneity that might exist. Using customer-level fixed effects might better tease out the unobserved customer-level heterogeneity. However, we believe that the use of random coefficient specification is a reasonable compromise due to its high parsimony over the customer fixed effect specification.

Conditional on being in the relationship state s at time t , we assume that the number of purchases that customer i makes follows a ZINBD with parameters $\varphi_{it|s}$, $\lambda_{it|s}$ and r . For each observation y_{it} , ZINBD assumes that there are two data generating processes (based on whether the outcome is equal to zero or is greater than zero) which are defined as:

$$P(Y_{it} = y_{it}|s) = \begin{cases} \varphi_{it|s} + (1 - \varphi_{it|s}) \left(1 + \frac{\lambda_{it|s}}{r}\right)^{-r} & \text{if } y_{it} = 0 \\ (1 - \varphi_{it|s}) \frac{\Gamma(y_{it} + r)}{y_{it}! \Gamma(r)} \left(1 + \frac{\lambda_{it|s}}{r}\right)^{-r} \left(1 + \frac{r}{\lambda_{it|s}}\right)^{-y_{it}} & \text{if } y_{it} > 0 \end{cases} \quad (8)$$

where r is a dispersion parameter that is assumed not to depend on covariates. The conditional mean and variance of the ZINBD are given as $E(y_{it}|s) = (1 - \varphi_{it|s})\lambda_{it|s}$ and $Var(y_{it}|s) = (1 - \varphi_{it|s})\lambda_{it|s}(1 + \varphi_{it|s}\lambda_{it|s} + \lambda_{it|s}/r)$, respectively. $\varphi_{it|s}$ and $\lambda_{it|s}$ capture the conditional zero-inflated probability and conditional expected purchase count for customer i (in state s) at time t , respectively.

To account for the excess of no purchases, we model the conditional zero-inflation component of the ZINBD ($\varphi_{it|s}$) as

$$\varphi_{it|s} = \frac{1}{1 + \exp(\delta_{0i|s} + \delta_{1i|s} \ln[LY_{it}])} \quad (9)$$

where $\delta_{0i|s}$ and $\delta_{1i|s}$ capture the intrinsic utility of making a purchase and the effect of duration dependence, given state s , respectively. We capture the duration dependence through time from the last purchase variable, i.e., LY_{it} . We use the natural logarithm of LY_{it} to capture the diminishing effect. To account for the unobserved heterogeneity, we allow $\delta_{ki|s}$ (for $k = 0,1$) to be customer-specific. We assume that $\delta_{ki|s}$ (for $k = 0,1$) are normally distributed across customers as follows.

$$\delta_{ki|s} = \delta_{k|s} + \nabla\delta_{ki}, k = 0,1 \quad (10)$$

where $\nabla\delta_{ki} \sim N(0, \sigma_{\delta_k}^2)$ for $k = 0,1$.

We model the second component of the conditional purchase model, i.e., conditional expected purchase count ($\lambda_{it|s}$), as a function of the number of emails sent by the firm (EM_{it}), that is given by

$$\lambda_{it|s} = \exp(\alpha_{ip|s} + \beta_{1,ip|s}EM_{it} + \beta_{2,ip|s}EM_{it}^2) \quad (11)$$

Conditional on state s , $\alpha_{ip|s}$ is the intrinsic propensity to make purchases, $\beta_{1,ip|s}$ and $\beta_{2,ip|s}$ are the corresponding response parameters. Note that, unlike the CEOM, we do not impose any identification restrictions on the parameters of the CPM. In other words, we do not make any restrictions such as $\alpha_{ip|1} \leq \alpha_{ip|2} \leq \dots \leq \alpha_{ip|NS}$ and/or $\delta_{0i|1} \leq \delta_{0i|2} \leq \dots \leq \delta_{0i|NS}$. This restriction free specification (in the CPM) allows our model to be flexible such that, for example, customers in a more active email open state may be less likely to be active in purchases, or vice versa. In other words, instead of pre-imposing restrictions such that email active customers must be active on purchases, we let the data to speak about whether this is really the case.

To account for the unobserved heterogeneity, we allow $\alpha_{ip|s}$, $\beta_{1,ip|s}$ and $\beta_{2,ip|s}$ to be customer-specific. We assume that $\alpha_{ip|s}$ and $\beta_{k,ip|s}$ (for $k = 1,2$) are normally distributed across customers as follows:

$$\begin{aligned} \alpha_{ip|s} &= \alpha_{p|s} + \Delta\alpha_{ip} \\ \beta_{k,ip|s} &= \beta_{k,p|s} + \Delta\beta_{k,ip}, k = 1,2 \end{aligned} \quad (12)$$

where $\Delta\alpha_{ip} \sim N(0, \sigma_{\alpha_p}^2)$ and $\Delta\beta_{k,ip} \sim N(0, \sigma_{\beta_{k,p}}^2)$ for $k=1,2$.

The Correlation between Purchase and Email Open Behavior

In each month, both the number of emails opened and purchases made by a customer might

possibly indicate the customer's interest to and level of interactions with the firm. Thus, we argue that there is a possible correlation between purchase and email open count distributions. Note that both the purchase count Y_{it} and the email open count O_{it} follow a discrete distribution. It is not easy to find a bivariate distribution that can capture the correlation between the ZINBD and the BD.

Danaher and Smith (2011) pioneer the use of Copula approach in marketing to link two marginal distributions that are not from the same family. Copula models are used in various studies to study multi-dimensional marketing problems (e.g., Park and Gupta 2012; Stephen and Galak 2012; Schweidel and Knox 2013; Kumar, Zhang and Luo 2014; Glady, Lemmens, and Croux 2015). However, most studies focus on using Copula models to correlate variables coming from continuous distribution families. Stephen and Galak (2012) model multivariate count variables using a double Poisson model, but they convert these discrete variables into continuous ones first and use the Gaussian Copula to correlate these converted variables. In the context of this study, we are dealing with two discrete variables that are characterized by two distinct distributions—ZINBD and BD. It is uncertain whether Stephen and Galak (2012)'s approach will work with such distinct discrete distributions. In addition, for efficiency, modeling the correlation between the two discrete variables directly rather than converting them to continuous ones first and treating them as continuous is more desirable.

To tackle the challenge of correlating discrete distributions, we turn to the literature of mathematics and statistics. Based on Sklar's theorem (1959), in a bivariate case as an example, the cumulative distribution functions (cdfs) of any two variables can be connected using a Copula function, and such Copula function is unique if the two variables are continuous. With two continuous variables, the bivariate density can be derived from the partial derivatives of the chosen

Copula function. With two discrete variables, such as the bivariate count data in this study, although we cannot rely on partial derivatives, we can still obtain the bivariate probability mass function (pmf) using finite differences of the chosen Copula function. There are several statistical applications of the idea to model bivariate count data. Lee (1999) develop a bivariate negative binomial distribution to model rugby league scores using Frank Copula. Song (2000) develops a multivariate dispersion model generated from Gaussian Copula. Nikoloulopoulos and Karlis (2010) capture the association between the purchase counts of certain product categories. Following Nikoloulopoulos and Karlis (2010) and others, we use a Copula to construct the bivariate pmf of O_{it} and Y_{it} as follows:

$$\begin{aligned}
h(o_{it}, y_{it}) = & C(F_1(o_{it}), F_2(y_{it})) - C(F_1(o_{it} - 1), F_2(y_{it})) \\
& - C(F_1(o_{it}), F_2(y_{it} - 1)) + C(F_1(o_{it} - 1), F_2(y_{it} - 1))
\end{aligned} \tag{13}$$

where $C(\cdot)$ is the Copula function, $F_1(o_{it})$ and $F_2(y_{it})$ are the distribution function of O_{it} and Y_{it} , respectively. We use Frank Copula (e.g., Frank 1979; Genest 1987) in this context because of its flexibility to capture the full range of correlation. The Frank Copula function is given by

$$C(u_1, u_2; \theta) = -\frac{1}{\theta} \log \left\{ 1 + \frac{(e^{-\theta u_1} - 1)(e^{-\theta u_2} - 1)}{e^{-\theta} - 1} \right\} \tag{14}$$

where u_1 and u_2 are the distribution functions (ZINBD and BD in our case) and θ is the Frank Copula correlation parameter.

Endogeneity Correction

There is a possibility that the firm is targeting customers by sending different number of emails based on their past purchase activities and inherent characteristics (that are unobserved to the researcher). In other words, the email contact variable (i.e., EM) might be endogenous. To solve this potential endogeneity problem in EM , we use a control function approach (Petrin and

Train 2010, the idea of which is like the approach in Villas-Boas and Winer (1999). This involves running a first-stage linear regression of EM on instruments⁷, which include customer-level purchase recency, purchase frequency (to capture customers' past purchase activeness) and customer level fixed effects (to capture every other customer-level characteristics that are possibly observed by the firm (to do targeted emailing), but not the researcher). From the first-step regression, we obtain an R^2 of 48.5%. In addition, we find the coefficients of purchase recency and frequency to be positive and statistically significant. In the second step, we incorporate the residuals from the first-stage regression as additional control variables into the HMM transition (Equation 3) and CPM (Equation 11) components of our proposed customer email open and purchase model.

Model Estimation

There are four sets of parameters to be estimated from our model: 1) parameters of the transition matrix (Equation 3); 2) parameters of the CEOM (Equations (6-7)); 3) parameters of the CPM (Equations 9-12); 4) the correlation parameter of the Frank Copula function (Equation 14). Following Netzer et al. (2008), we write the vector of the bivariate probability mass function as a diagonal matrix \mathbf{H}_{it} . Given the proposed HMM structure, the likelihood function for a sequence of observations $\{(Y_{i1} = y_{i1}, O_{i1} = o_{i1}), \dots, (Y_{it} = y_{it}, O_{it} = o_{it})\}$ can be expressed as

$$\begin{aligned}
 L &= \prod_{i=1}^N P((Y_{i1} = y_{i1}, O_{i1} = o_{i1}), \dots, (Y_{it} = y_{it}, O_{it} = o_{it})) \\
 &= \prod_{i=1}^N \boldsymbol{\psi}_i' \mathbf{H}_{i1} \prod_{t=2}^T \mathbf{Q}_{it} \mathbf{H}_{it}
 \end{aligned} \tag{15}$$

⁷ Since the variation in EM variable is large, we treat it as a continuous variable rather than discrete in our application. In addition, we run a Poisson regression and find that residuals from that count regression and the linear regression are highly correlated ($r=0.97$).

where $\mathbf{1}$ is an $S \times 1$ vector of ones.

Model Identification

Since all model components (transition matrix, CEOM and CPM) always enter the likelihood calculation (see Equation 15) as a bundle, it is not possible to separately identify the coefficients of a specific variable if that variable is used repeatedly across different equations. Consequently, we choose and use a distinct set of variables in different components of our model. The chosen variables represent customers' past activities and the firm's past and current marketing activities; and their choice is in line with the CRM literature (e.g., Netzer et al. 2008; Kumar et al. 2011; Li et al. 2011). Specifically, we model customer i 's transition probability (from period $t-1$ to t) among the states of the Markov process as a function of interactions between customer i and the firm at period $t-1$. We use $O_{i,t-1}$, $Y_{i,t-1}$ to capture the recent interaction of the customer i with the firm, and $EM_{i,t-1}$ to capture the firm's recent interaction with customer i . We model the CEOM component as a function of the customer i 's time from the last email open variable, i.e., LO_{it} . We model the CPM component as a function of customer i 's time from the last purchase variable (LY_{it}) and the firm's email marketing contact policy (EM_{it}). Also, note that we use lagged and contemporaneous levels of customer and firm level behavioral variables in the transition and conditional behavior (CEOM and CPM) components of our model. Since we have sufficient within variable variation (for these behavioral variables) over time, we would be able to separately identify all three components of our model, i.e., the transition, CEOM and CPM.

We identify the parameters of each of the abovementioned component as follows. For the transition component itself, we observe a significant within customer behavior (email open and purchase) variation over time. This implies that customers' relationship levels with the firm vary significantly over time. This is the key for us to empirically identify parameters of the transition

component (i.e., the parameters in Equation 3). Within relationship levels, we also observe customers behaving differently, and this helps us identify the unobserved customer heterogeneity parameters (i.e., the parameters of the heterogeneity distributions in Equations 7, 10 and 12). For the CEOM and CPM components, we observe a significant within time variation across email open and purchase behaviors of customers. For example, at a given time, some customers are active in one, some active in both behaviors, and some passive in both. In addition, within customers, the intensity of the two behaviors significantly vary over time. These variations help us identify the parameters of CEOM and CPM components (i.e., the parameters in Equations 6, 9 and 11). In the next section, we discuss our empirical results.

Empirical Results

We estimate our HMM model using the maximum likelihood estimation (MLE) method. We maximize the simulated sample log-likelihood to estimate our model parameters. We decide on the number of HMM states based on the Bayesian information criterion (BIC). In our application, we compare the performance of HMMs of up to 4 states (see Table 3). We find that the HMM with 3 states provides the best fit to the data as it gives the lowest BIC value. Thus, we choose the 3-state HMM for further analysis.

Insert Table 3 here

In the following section, we first discuss the parameter estimates from our proposed hidden Markov and Copula Model of customer purchase and email open behaviors. Second, we illustrate how the firm's email contact decision would affect the customer-firm relationship states. Specifically, we discuss how the firm's email contact decision affects customers' transitions among different relationship states. Third, we discuss the distribution of customers across the three customer-firm relationship states. Specifically, we illustrate the proportion of customers that

belong to different relationship states over time.

Parameter Estimates

Table 4 reports the parameter estimates for the 3-state HMM. In order to label the states in terms of their purchase and email open behaviors, we calculate the predicted purchase and email open counts (at the average observed levels of customer and firm level behavioral variables such as *LY*, *LO*, *EM* etc.). We find those customers in the relationship states 1, 2 and 3 open 0.215, 2.745 and 2.764 emails (on average) per month, respectively. In terms of the purchase behavior, we find those customers in the relationship states 1, 2 and 3 make 0.338, 0.148 and 1.219 purchases (on average) per month, respectively. Based on that calculation, we label the three latent relationship states, which govern the frequency of customers' email open and purchase in each month as "Low Open – Medium Purchase", "High Open – Low Purchase", and "High Open – High Purchase"⁸. The results suggest that in State 3, the conventional wisdom, i.e., the more customers make purchases, the more they open emails, holds. However, such conventional wisdom does not hold in states 1 and 2. Specifically, email active customers might be inactive in purchases (State 2), and email inactive customers may still be reasonably active in purchases (State 1). This result underlines the importance of our assumption related to not restricting the parameters of the purchase model along with the email open model (i.e., imposing the following restriction: higher the state, higher the email open and higher the purchase levels).

Insert Table 4 here

Table 4 shows that customers in all relationship states show a negative purchase duration dependence. The longer they have not made a purchase, the less frequently they make purchases.

⁸ For ease of presentation, we use purchase only labels, i.e., Low, Medium and High Purchase to refer states 2, 1 and 3, when we have a discussion related to purchases only. We use the full labels Low Open – Medium Purchase, High Open – Low Purchase and High Open – High Purchase to refer states 1, 2 and 3, when we have a discussion related to both purchase and email open behaviors.

The effect is strongest for the Low Purchase customers (-2.235), and it is the weakest for the High Purchase customers (-1.294). Similarly, customers also show a negative duration dependence for the email open behavior (-0.637), i.e., the longer they have not opened an email, the less frequently they open emails. The results suggest that customers making purchases (opening emails) in the previous month increases (decreases) the transition probability to the High Open – High Purchase state. The results also suggest that, across all relationship states, the number of emails sent by the firm initially increases, but then decreases (inverse U-shape) the short-term purchase count. In other words, the email contact has a nonlinear effect on the purchase count. In addition, the peak-point of the inverse U-shaped curve varies across different relationship states, i.e., customers respond in diverse ways to email contacts depending on their relationship state. This implies that there are different optimal numbers of emails to send to maximize the short-term purchase count across the three relationship states.

Results suggest that two out of the five endogeneity correction parameters are significant. In other words, it is important to control for the endogeneity of *EM* to be able to recover the short- and long-term effects correctly of *EM* in customers' purchase behaviors and transitions among the relationship states. We also find significant heterogeneity in both the estimates of intercepts and the response coefficients of the CEOM and CPM components of our proposed model.

Transition Probability Matrix

The transition matrix from the HMM model shows how customers evolve across different relationship states. We calculate the transition probabilities of a “typical” customer using Equations (1-4). We vary the number of emails received in the previous period and check the effect of email contacts on the state transitions (see Table 5).

Insert Table 5 here

When there is no email contact, the customers from Medium Purchase and Low Purchase states do not tend to move, while those from High Purchase state tend to move down to a Medium Purchase state or stay within the High Purchase state. Table 5 shows that email contact has diverse effects on the transition probabilities of the customers who are from different purchase states. For example, ten email contacts per month increases the likelihood that customers from the Low Purchase State move up to the Medium Purchase state (from 4.53% to 21.42%). However, one email contact per month only marginally increases such likelihood to 5.67%. For the customers who are from the Medium Purchase state, email contact increases their probabilities of staying in the same state. In summary, email contact has a positive long-run effect on either making customers be more engaged with the firm if they were not so or making them stay engaged if they have already been so. However, email contact has a negative long-run effect on the customers who have already been in a good relationship with the firm. For the customers who are from the High Purchase state, excessive email contacts move them down to lower purchase states. While the likelihood of them staying in the High Purchase state is 46.62% in the no-email scenario, such likelihood decreases significantly to 25.97% in the 10-emails scenario. This non-linear dependence between the email contacts and the transition probabilities among the purchase states motivate us for our policy simulation (in the next section) in which we determine the optimal number of emails to send to maximize the firm' lifetime profit per customer.

In Figure 4, we plot the average probabilities of customers residing in the three purchase states over time. We calculate the state membership distribution of each customer using the filtering approach (Montgomery et al. 2004; Netzer et al. 2008). Since we assume that each customer starts from the lowest state, we drop the first five periods as initialization periods and plot the states evolution over the rest of the 33 periods. We find that, on average, 74%, 12% and

14% of customers started in the Medium-, Low- and High-Purchase states, respectively. Over the course of 33 months, the majority of the customers remained in the medium purchase state while the rest moved up or down to different states. In addition to understanding the aggregate distribution of customers across the different relationship states (Figure 4), the firm might use the filtering approach to understand the state membership distribution at the customer-level. Then, the firm might rely on this information to make targeted email contact decisions. We illustrate this idea in our policy simulation section next.

Insert Figure 4 here

Optimal Email Marketing

The optimal marketing contact strategy is the focus of several earlier studies (e.g., Khan, Lewis, and Singh 2009; Kumar et al. 2011; Li, Sun and Montgomery 2013). In this study, since we find that sending more emails is not necessarily good for the firm to encourage its customers to make more purchases, determining the optimal number of emails might provide substantive profit gains for the firm. Therefore, our objective in this section is to provide a DSS that managers can rely on to maximize the long-run firm profit by sending the right number of emails to their customers over time.

In this setting, at any given time t , the firm has to make the decision of how many emails to send. Given our estimated customer response function, the email contact decision has both short- and long-term implications on the customer behavior. Short-term effect derives from the direct effect of emails on customer purchase and email open behaviors at time t (see Equations 5 and 11). Long-term effect comes from two sources: (1) customers' purchase and email open behaviors at time t affect the evolution of customer relationship state from time t to $t+1$ (see Equation 3); (2) the email contacts at time t influence the relationship state transition from time t to $t+1$ (see

Equation 3). Due to these long-term effects, determining the optimal number of emails to send by the firm requires us to solve a dynamic programming problem.

From the retailer's perspective, the variable of interest is the number of times a customer purchases per month from the store. Under the assumptions of constant purchase amount per each purchase transaction and fixed gross margin for the retailer⁹, the purchase count is directly translated into the firm's profit (per customer and per transaction). For the firm's dynamic optimization problem, the payoff relevant state variables are: (1) the probabilities that the customer exists in each of the customer-firm relationship states (p_{1t} , p_{2t} , $p_{3t} = 1 - p_{1t} - p_{2t}$), (2) the time since last purchase (LY), and (3) the time since last email open (LO). Therefore, the state vector at time t becomes $S_t = (p_{1t}, p_{2t}, LY_t, LO_t)$. Following Kumar et al. (2011), we assume the timing of the email contact decisions as follows. At the beginning of each month t , the firm predicts the probability that the customer exists in each of the three relationship states, p_{1t} , p_{2t} and $p_{3t} (= 1 - p_{1t} - p_{2t})$. Next, based on the predicted p_{1t} , p_{2t} and p_{3t} , as well as LY_t and LO_t , the firm decides how many emails to send to the customer. We use a multinomial logit share function to capture the state membership probabilities (p_{1t} , p_{2t} and p_{3t}) and relate them to two parameters ω_{1t} and ω_{2t} as follows:

$$\begin{aligned}
 p_{1t} &= \frac{\exp(\omega_{1t})}{1 + \exp(\omega_{1t}) + \exp(\omega_{2t})} \\
 p_{2t} &= \frac{\exp(\omega_{2t})}{1 + \exp(\omega_{1t}) + \exp(\omega_{2t})} \\
 p_{3t} &= 1 - p_{1t} - p_{2t}
 \end{aligned} \tag{16}$$

⁹ For confidentiality reasons, customer-level purchase amount is not observed in our dataset. However, based on our communication with the managers of the firm, we are advised that the average purchase amount for each customer transaction is around \$80-\$110. We are also advised that the average profit margin is around 18-22% across different product categories. Thus, in our profit and lifetime value calculations, we assume an average purchase amount of \$100 per customer, per transaction; and 20% profit margin for each customer purchase occasion. This assumption yields \$20 net profit per customer at each purchase occasion.

Let $S_t = (p_{1t}, p_{2t}, LY_t, LO_t)$ denote the state vector at time t . Time from last purchase and open states, LY_t and LO_t , evolve based on whether the customer makes purchases and opens emails at time t . If the customer makes purchases at time t , the corresponding state LY_{t+1} becomes 1, and if he or she doesn't make any purchases at time t , LY_{t+1} becomes $LY_t + 1$. Similarly, if the customer opens emails at time t , the corresponding state LO_{t+1} becomes 1, and if he or she doesn't open any emails, LO_{t+1} becomes $LO_t + 1$. Since the purchase and email open processes of customers are modeled with ZINBD and BD, the time from last purchase and open states, LY_t and LO_t , evolves in a stochastic manner as follows

$$\begin{aligned}
LY_{t+1} &= \begin{cases} 1, & \text{with } Pr(Y_{it} > 0) \\ LY_t + 1, & \text{with } Pr(Y_{it} = 0) \end{cases} \\
LO_{t+1} &= \begin{cases} 1, & \text{with } Pr(O_{it} > 0) \\ LO_t + 1, & \text{with } Pr(O_{it} = 0) \end{cases}
\end{aligned} \tag{17}$$

The evolution of the first two state variables $p_{1,t+1}$ and $p_{2,t+1}$ conditional on the firm's email contact decision EM_t are given as

$$\begin{aligned}
p_{s,t+1}|EM_t &= Pr(Y_{it} > 0, O_{it} > 0) \sum_{k=1}^s p_{kt} q_{t+1,k \rightarrow s} (1,1, EM_t) \\
&+ Pr(Y_{it} = 0, O_{it} > 0) \sum_{k=1}^s p_{kt} q_{t+1,k \rightarrow s} (0,1, EM_t) \\
&+ Pr(Y_{it} > 0, O_{it} = 0) \sum_{k=1}^s p_{kt} q_{t+1,k \rightarrow s} (1,0, EM_t) \\
&+ Pr(Y_{it} = 0, O_{it} = 0) \sum_{k=1}^s p_{kt} q_{t+1,k \rightarrow s} (0,0, EM_t)
\end{aligned} \tag{18}$$

where $q_{t+1,k \rightarrow s}(I[Y_t > 0], I[O_t > 0], EM_t)$ is the transition function from Equation 4 which is used to calculate the probability of transitioning customers from state k at time t to s at time $t+1$

conditional on email contacts, purchase and email open indicators at time t .

At each time t conditional on the state vector S_t , the objective of the firm is to determine the optimal number of email contacts to maximize the discounted sum of expected future profits. Under some regularity conditions this objective can be written in the following Bellman equation.

$$V(S_t) = \max_{EM_t} \{ \pi(EM_t, S_t) + \rho EV(S_{t+1} | S_t, EM_t) \} \quad (19)$$

where ρ is the discount factor, $\pi(\cdot)$ is the per period profit and the expectation is over all the future states and actions of the firm.

In order to solve this dynamic optimization problem, we discretize the state space with 10 levels for each state dimension p_{1t} , p_{2t} , LY_t and LO_t yielding 10,000 state combinations. We use the value iteration algorithm (Rust 1987) to find the optimal mappings of the firm's email contacts to our chosen state combinations. Due to the discretization of the first two state dimensions, the value functions for the other points in the state space are computed via interpolation (Keane and Wolpin 1994). After we calculate the vector of the optimal mapping of email contacts (through the value iteration algorithm), we fit a multinomial logit model (MNL)¹⁰ to predict optimal email contacts as a flexible function of the chosen state space. Next, we use this fitted MNL policy function to predict the optimal number of emails to send for any state combinations out of the chosen state space.

Based on the optimal mapping of email contacts to states from the value iteration algorithm, we find that the optimal email contact number ranges from 5 to 14. We see a lot of heterogeneity in the ranges of optimal number of emails sent based on customers' relationship states. For instance, if the firm has a belief that with more than 30% probability that the customer is in the

¹⁰ The choice of the MNL functional form for the policy function is because the optimal mapping of the email contacts takes discrete and finite number of values (ranging between 5 and 14 emails to send as discussed in the next paragraph).

High Purchase state, the optimal number of emails to send ranges between 5 and 7 (based on the value of the remaining state combinations). If the firm's belief becomes such that the customer is in the Medium Purchase state with more than 30% probability, then the optimal number of emails to send ranges between 6 and 10. If the firm has a strong belief (more than 90% probability) that the customer is in Low Purchase state, the optimal number of emails to send ranges between 12 to 14 emails. Whereas, if the opposite is the case, i.e., the firm has a strong belief that the customer is not in the Low Purchase state, the optimal number of emails to send ranges between 5 and 7.

This heterogeneity in the optimal number of emails to send suggests that the firm should target different customers (by sending different number of emails) based on their observed behaviors (*LY* and *LO*) and its belief about their relationship states. To illustrate how the firm might use our DSS to target its customers through email marketing, we randomly pick two customers from our database. We then calculate their HMM state membership probabilities over time through the filtering approach, and use our optimal policy function to set the number of email contacts (based on their observed *LY* and *LO* levels in the data). We plot the optimal number of email contacts over time for these two customers in Figure 5. As can be seen in Figure 5, the optimal number of emails to send varies significantly not only across these two customers, but also within the same customer over time. This exercise suggests that our proposed DSS can easily be used as an actionable management tool by the firm to dynamically manage its email contact decisions in a profitable manner.

Insert Figure 5 here

Next, we use our optimal policy function to simulate the firm's email contact decisions along with the customer purchase and email open responses over a long-time horizon for three representative customers in Low, Medium and High Purchase states, respectively. This forward-

simulation exercise helps us understand the CLV (customer lifetime value) of customers differ across three purchase states. Based on our assumptions (see Footnote 9), we find that CLVs of customers in Low, Medium and High Purchase states as \$1,333, \$1,411, and \$1,465, respectively. This exercise also reveals that, in the steady-state, the optimal number of emails to send to each customer becomes 7 emails per month. In addition, the discounted sum of lifetime purchase and email open counts become 73 and 118 per customer, respectively.

Lastly, we conduct a what-if simulation study to measure how much profit the firm would leave on the table if it deviates from the recovered optimal email policy function (see Figure 6). We use the steady-state distribution as the starting state combinations and test the alternative scenarios in which the firm deviates from sending the optimal number of 7 emails. We test the scenarios where the firm sends 4, 5, 6, 8, 9, and 10 emails instead of 7. Figure 6 shows that sending sub-optimal number of emails might cause the firm to lose significant amount of profit. For instance, sending 4 (10) emails instead of the optimal level of 7 emails causes the firm to lose 32% (16%) of its lifetime profit per customer. This result suggests that sending the right number of emails is highly critical for the profitability of the firm's email marketing program.

Insert Figure 6 here

Robustness Checks

To check whether our model and findings are robust to alternative model specifications, we conduct four sets of robustness checks:

- 1) Using alternative distributions to model the conditional purchase counts,
- 2) Using different Copula functions to model the correlation between customers' purchase and email open behaviors,
- 3) Imposing HMM identification restriction on the CPM, and

4) Including holiday dummies as additional control variables in the CPM model.

As our first robustness check, we estimate three alternative model specifications where the conditional purchase count is assumed to be distributed as Poisson (PD), Zero-Inflated Poisson (ZINPD), and Negative Binomial (NBD), respectively. We find that the estimates from these three alternative specifications look qualitatively similar to our CPM with ZINBD specification. However, based on the BIC criteria, we determine that our proposed ZINBD specification outperforms the alternative models ($BIC_{ZINBD} = 30,658.11$, $BIC_{PD} = 33,440.23$, $BIC_{ZINPD} = 32,201.11$, $BIC_{NBD} = 30,994.36$). Therefore, we assume the conditional purchase count is distributed by ZINBD.

As our second robustness check, we estimate two alternative model specifications. Specifically, we estimate the alternative models with Clayton (1978) and Gumbel (1960) Copula functions, which have been widely used in the statistics literature. Although, all three specifications yield qualitatively very similar estimates, based on the BIC criteria, we find that the Frank Copula ($BIC = 30,658.11$) slightly outperforms Clayton ($BIC = 30,676.25$) and Gumbel ($BIC = 30,682.59$) Copula specifications. Therefore, we use Frank Copula as our proposed Copula specification, and use the Frank Copula specification in the rest of our analysis.

For identification purpose, we could have imposed HMM restrictions on the CPM instead of the CEOM. However, the former restriction requires us to impose $6*(NS-1)$ restrictions, including $NS-1$ intercepts and $NS-1$ response coefficients in the φ_{it} component, and $NS-1$ intercepts and $3*(NS-1)$ response coefficients in the λ_{it} component, while the latter one only requires $2*(NS-1)$ restrictions ($NS-1$ intercepts and $NS-1$ response coefficients in the p_{it} component). Due to this, we prefer to impose restrictions on the CEOM rather than the CPM. As our third robustness check, we estimate the alternative specification in which the identification

restriction is imposed on the conditional purchase model. Based on the BIC criteria, our proposed model (BIC = 30658.11) outperforms that alternative specification (BIC = 31308.9). Therefore, we use our proposed model with HMM restrictions on CEOM as our main model in the rest of the analysis.

Finally, as our last robustness check, we incorporate the holiday dummies into our CPM as additional control variables. Essentially, customers might be more active in their purchase activities during the holidays. Since our data is at the monthly level, we define the months corresponding to Memorial Day, Fourth of July, Labor Day, Thanksgiving and Christmas as holiday months, and dummy coded this holiday variable. From the comparison of our proposed model and the model with holiday dummies, we find that 1) all parameters look qualitatively very similar under both specifications; 2) coefficients for the holiday dummies are statistically insignificant; and 3) based on the BIC criteria our proposed model slightly outperforms the alternative model (30658.11 versus 30661.85). As a result, we use our proposed model without holiday dummies as our main model in the rest of the analysis. We discuss our empirical findings next.

Conclusions, Limitations and Future Research

The email marketing programs are used extensively in various industries to engage with customers. The general industry practice in measuring the effectiveness of an email-marketing program is to examine customer responsiveness to emails such as email open rate. However, we show that considering only the email open rate could be misleading. Our empirical study shows that some of the very high email openers are the least active ones in their purchase behaviors. In addition, our findings show that some email inactive customers are relatively active in their purchase behaviors. Thus, if firms solely focus on the email open rate to allocate its resources, they

might potentially overlook a pool of customers who are inactive in responding to emails but are relatively active in purchases.

To the best of our knowledge, substantively, this is the first empirical study that models the customer's email open and purchase behaviors jointly. In addition, methodologically, this is the first study that combines the hidden Markov and Copula models in a unified framework. In our HMM specification, we model the latent customer-firm relationship states that govern customers' both purchase activeness and email responsiveness. In our Copula specification, we use a Frank Copula to correlate the customer's email open and purchase behaviors.

Note that, the purpose of this study is not diverting firms' attention from email open rate. This study shows that, on average, there is a positive correlation between email open and purchase behaviors. Instead, we recommend firms to look at customers' purchase behaviors in addition to customers' response rates to emails. If the goal is to maximize the long-term profitability, firms should be informed about the optimal level of email contacts they should make to their customers. Along these lines, we calculate the optimal email marketing contact policy by solving the firm's dynamic optimization problem. Eventually, we propose an implementable framework to study an important substantive problem that can save firms millions of dollars.

In our specific application, although we use email open as the non-purchase customer behavior, our framework is quite flexible that it is applicable to other customer-level non-purchase behaviors. In addition, our DSS is designed to guide the firm about how many emails to send, however, other firm-level marketing decisions might also be guided with the proposed structure without significant issues. Furthermore, as more information that might affect the purchase and non-purchase behaviors becomes available, one might incorporate that information through constructing additional control variables entering purchase and non-purchase models. Among such

variables, the ones, which are relevant for the firm's decision in the supply side, might be incorporated into the operationalization of the proposed DSS. This will allow the firm to target different customers at a given time, and the same customers over time differently based on additionally observed customer level characteristics.

One of the limitations of this study is that we do not observe the content of the emails. Therefore, we do not focus on the emailing strategy that intends to match with customers' contemporary needs, such as cross- and up selling. Previous research has investigated this important issue (e.g., Kumar, George and Pancras 2008; Li, Sun and Wilcox 2005). If the email content data is available, future research might not only consider the effect of emailing content on customers' response to emails and purchase behavior, but also provide guidance to the firm related to targeted email marketing based on personalized customization of the email content.

Another limitation of this study is that we do not observe the information from the competitors. Customers' lack of response to emails could simply be due to subscribing to many email programs from different firms. Each email delivered to the customer's inbox is a load of information. Customers who are not capable of processing the information will be overwhelmed and stop responding. If the firm is aware of its customers' inbox activity, incorporating this information into the study is imperative. However, due to the sensitivity of such information the likelihood that the retailer can obtain this information is unlikely. Future studies could consider conducting field experiments to understand how competing emails affect customers' reactions to the firm's emails.

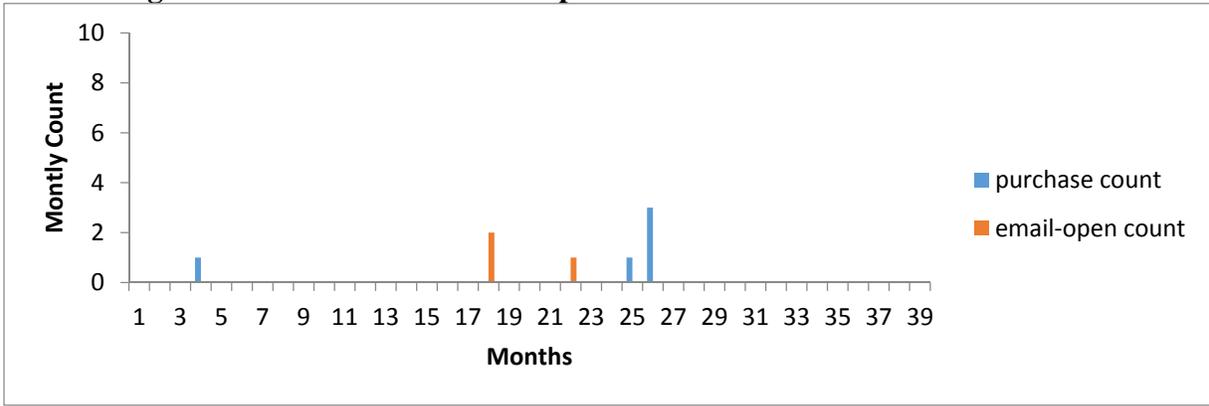
REFERENCES

- Ascend2 (2015), "Marketing Technology Strategy Survey Summary Report," [available at <http://ascend2.com/home/wp-content/uploads/Marketing-Technology-Strategy-Survey-Summary-Report-150813.pdf>].
- BlueHornet (2013), "One More Time: Email Frequency Chief Culprit in Unsubscribes," [available at <http://www.marketingcharts.com/online/one-more-time-email-frequency-chief-culprit-in-unsubscribes-27944/>].
- Bonfrer, André and Xavier Drèze (2009), "Real-Time Evaluation of E-mail Campaign Performance," *Marketing Science*, 28 (2), 251-263.
- Chittenden, Lisa and Ruth Rettie (2003), "An Evaluation of E-mail Marketing and Factors Affecting Response," *Journal of Targeting, Measurement and Analysis for Marketing*, 11 (3), 203-217.
- Clayton, D. G. (1978), "A Model for Association in Bivariate Life Tables and Its Application in Epidemiological Studies of Familial Tendency in Chronic Disease Incidence," *Biometrika*, 65 (1), 141-151.
- Danaher, Peter J. and Michael S. Smith (2011), "Modeling Multivariate Distributions Using Copulas: Applications in Marketing," *Marketing Science*, 30 (1), 4-21.
- Direct Marketing Association (UK) Ltd (2015), "National Client Email 2015," [available at <http://dma.org.uk/uploads/ckeditor/National-client-email-2015.pdf>].
- Frank, M. J. (1979), "On the Simultaneous Associativity of $F(x,y)$ and $x + y - F(x,y)$," *Aequationes Mathematicae*, 19, 194-226.
- Genest, Christian (1987), "Frank's Family of Bivariate Distributions," *Biometrika*, 74 (3), 549-555.
- Glady, Nicolas, Aurélie Lemmens, and Christophe Croux (2015), "Unveiling the Relationship between the Transaction Timing, Spending and Dropout Behavior of Customers," *International Journal of Research in Marketing*, 32 (1), 78-93.
- Godin, Seth (1999), *Permission Marketing: Turning Strangers into Friends and Friends into Consumers*. New York: Simon & Schuster.
- Gumbel, E. J. (1960), "Distributions des Valeurs Extremes en Plusieurs Dimensions," *Publications de l'Institut de Statistique de l'Université de Paris*, 9, 171-173.
- Keane, Michael P. and Kenneth I. Wolpin (1994), "The Solution and Estimation of Discrete Choice Dynamic Programming Models by Simulation and Interpolation: Monte Carlo Evidence," *The Review of Economics and Statistics*, 648-672.

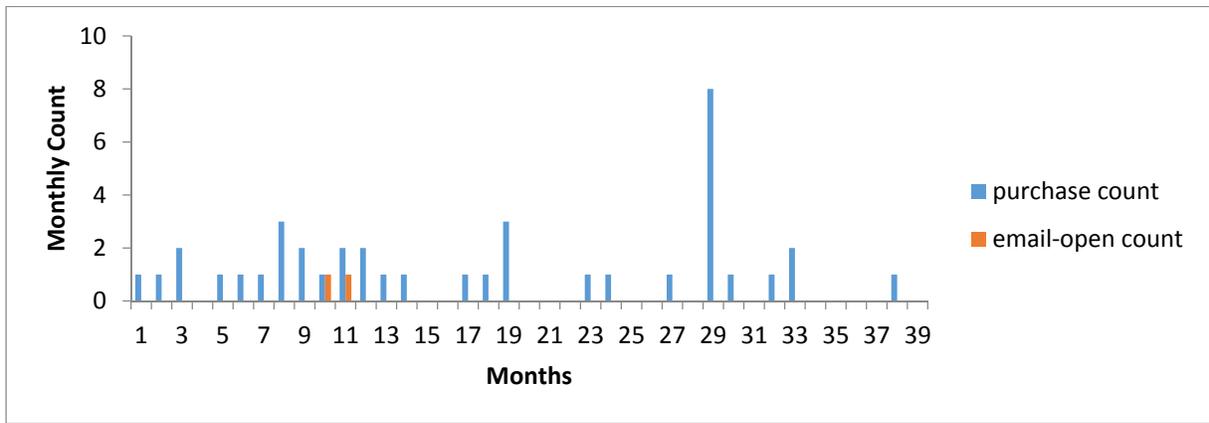
- Khan, Romana, Michael Lewis, and Vishal Singh (2009), "Dynamic Consumer Management and the Value of One-to-One Marketing," *Marketing Science*, 28 (6), 1063-1079.
- Kumar, V., Morris George, and Joseph Pancras (2008), "Cross-buying in Retailing: Drivers and Consequences," *Journal of Retailing*, 84 (1), 15-27.
- Kumar, V., Sriram S., Anita Luo, and Pradeep K. Chintaguta (2011), "Assessing the Effect of Marketing Investments in a Business Marketing Context," *Marketing Science*, 30 (5), 924-940.
- Kumar, V., Xi (Alan) Zhang, and Anita Luo (2014), "Modeling Consumer Opt-In and Opt-Out in a Permission-Based Marketing Context," *Journal of Marketing Research*, 51 (4), 403-419.
- Lee, Alan (1999), "Applications: Modelling Rugby League Data via Bivariate Negative Binomial Regression," *Australian & New Zealand Journal of Statistics*, 41 (2), 141-152.
- Li, Shibo, Baohong Sun, and Alan L. Montgomery (2011), "Cross-Selling the Right Product to the Right Consumer at the Right Time," *Journal of Marketing Research*, 48 (4), 683-770.
- Li, Shibo, Baohong Sun, and Ronald T. Wilcox (2005), "Cross-selling Sequentially Ordered Products: An Application to Consumer Banking Services," *Journal of Marketing Research*, 42 (2), 233-239.
- Luo, Anita and V. Kumar (2013), "Recovering Hidden Buyer-Seller Relationship States to Measure the Return on Marketing Investment in Business-to-Business Markets," *Journal of Marketing Research*, 50 (1), 143-160.
- Merisavo, Marko and Mika Raulas (2004), "The Impact of E-Mail Marketing on Brand Loyalty," *Journal of Product & Brand Management*, 13 (7), 498-505.
- Montgomery, Alan, Shibo Li, Kannan Srinivasan, and John C. Liechty (2004), "Modeling Online Browsing and Path Analysis using Clickstream Data," *Marketing Science*, 23 (4), 579-595.
- Montoya, Ricardo, Oded Netzer, and Kamel Jedidi (2010), "Dynamic Allocation of Pharmaceutical Detailing and Sampling for Long-Term Profitability," *Marketing Science*, 29 (5), 909-924.
- Netzer, Oded, James M. Lattin, and V. Srinivasan (2008), "A Hidden Markov Model of Consumer Relationship Dynamics," *Marketing Science*, 27 (2), 185-204.
- Nikoloulopoulos, Aristidis K. and Dimitris Karlis (2010), "Regression in a Copula Model for Bivariate Count Data," *Journal of Applied Statistics*, 37 (9), 1555-1568.
- Park, Sungho and Sachin Gupta (2012), "Handling Endogenous Regressors by Joint Estimation Using Copulas," *Marketing Science*, 31 (4), 567-586.

- Petrin, A., & Train, K. (2010). A control function approach to endogeneity in Consumer choice models. *Journal of marketing research*, 47(1), 3-13.
- Return Path (2015), "Frequency Matters: The Keys to Optimizing Email Send Frequency," [available at <http://returnpath.com/wp-content/uploads/2015/06/RP-Frequency-Report-FINAL.pdf>].
- Rust, John (1987), "Optimal Replacement of GMC Bus Engines: An Empirical Model of Harold Zurcher," *Econometrica: Journal of the Econometric Society*, 999-1033.
- Sahni, Navdeep, Dan Zou, and Pradeep K. Chintagunta (2016), "Do Targeted Discount Offers Serve as Advertising? Evidence from 70 Field Experiments," *Management Science*, forthcoming.
- Schweidel, David A., Young-Hoon Park, and Zainab Jamal (2014), "A Multiactivity Latent Attrition Model for Consumer Base Analysis," *Marketing Science*, 33 (2), 273-286.
- Schweidel, David A., and George Knox (2013), "Incorporating Direct Marketing Activity into Latent Attrition Models," *Marketing Science*, 32 (3), 471-487.
- Sklar, A. (1959), "Fonctions de répartition à n dimensions et leurs marges," *Publications de l'Institut de Statistique de L'Université de Paris*, 8, 229-231.
- Stephen, Andrew T., and Jeff Galak (2012), "The Effects of Traditional and Social Earned Media on Sales: A Study of a Microlending Marketplace," *Journal of Marketing Research*, 49 (5), 624-639.
- Tezinde, Tito, Brett Smith, and Jamie Murphy (2002), "Getting Permission: Exploring Factors Affecting Permission Marketing," *Journal of Interactive Marketing*, 16 (4), 28-39.
- Trivedi, Pravin K. and David M. Zimmer (2005), "Copula Modeling: An Introduction for Practitioners," *Foundations Trends Econometrics*, 1 (1), 1-111.
- Villas-Boas, J. M., & Winer, R. S. (1999). Endogeneity in brand choice models. *Management Science*, 45(10), 1324-1338.
- Xue - Kun Song, Peter (2000), "Multivariate Dispersion Models Generated from Gaussian Copula," *Scandinavian Journal of Statistics*, 27 (2): 305-320.

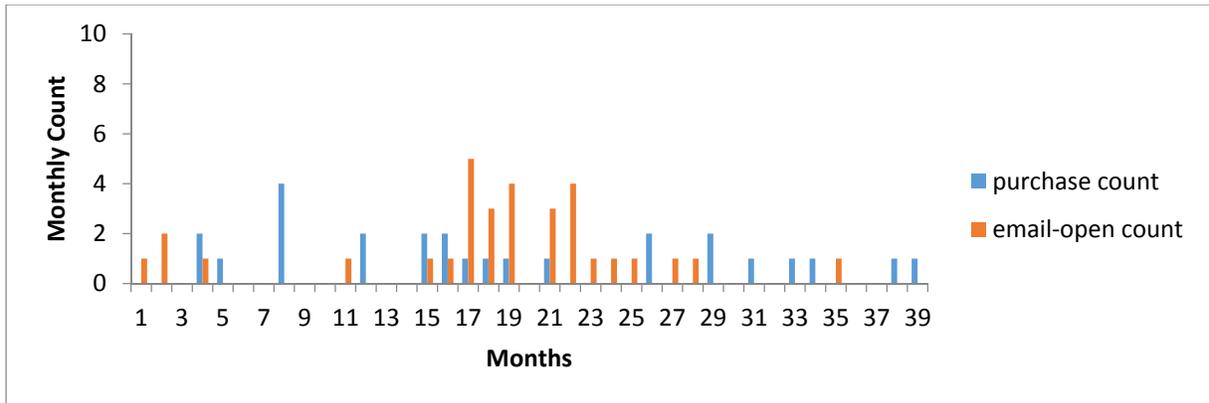
Figure 1: Purchase and Email Open Count of Three Selected Customers



Customer 1



Customer 2



Customer 3

Figure 2: Distributions of Purchase and Email Open Count

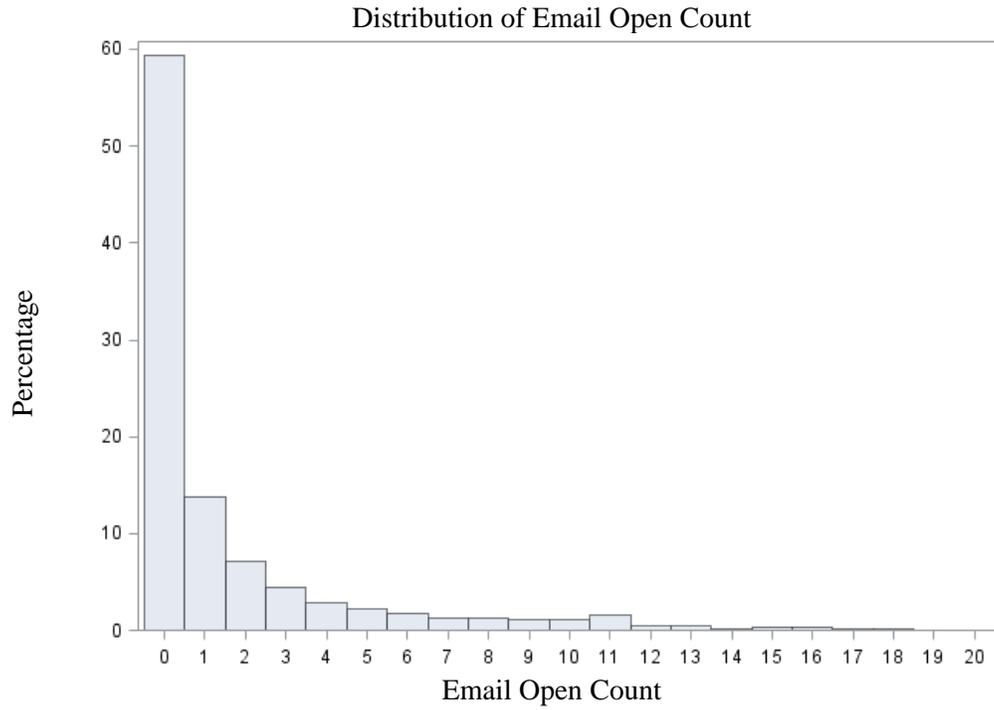
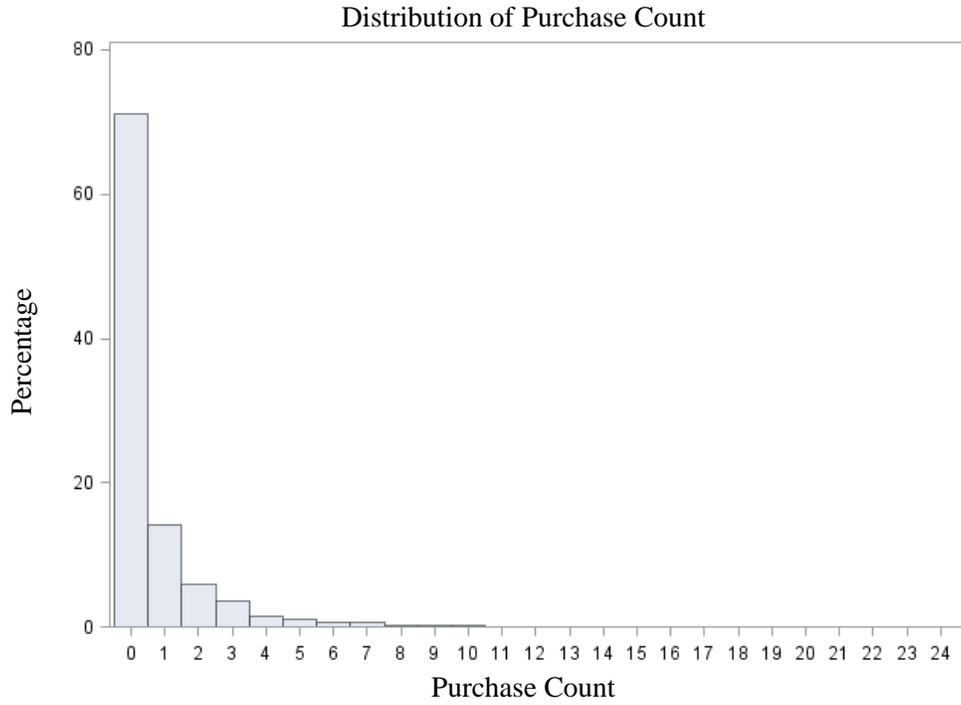


Figure 3: Hidden Markov and Copula Model of Customer Purchase and Email Open Behaviors

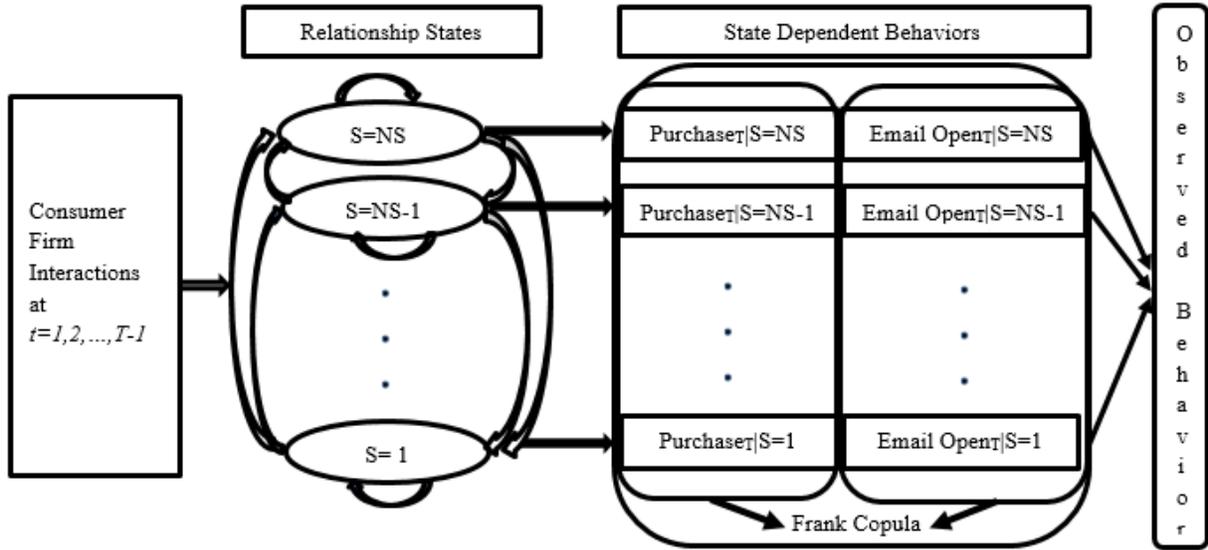
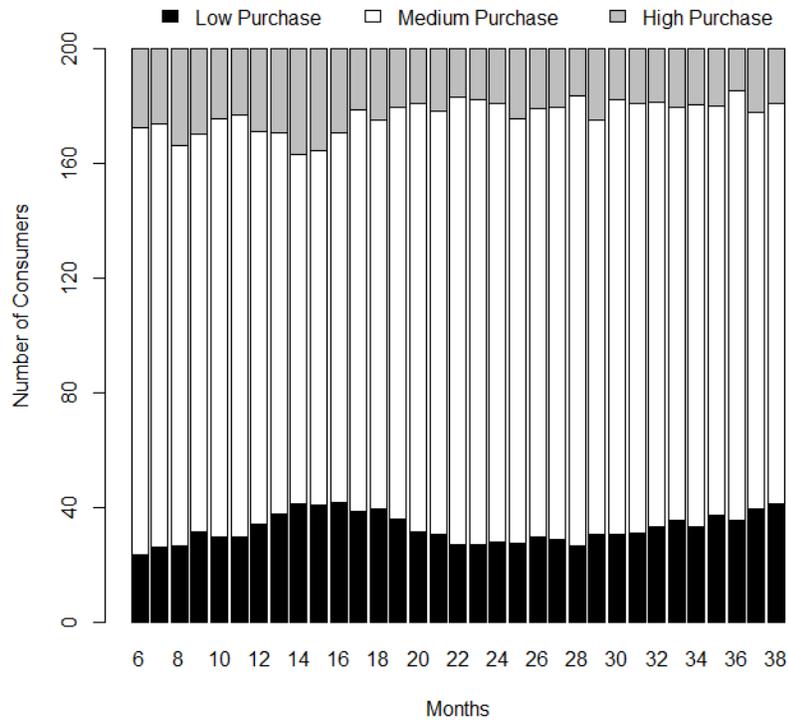


Figure 4: Distribution of Customers' State Membership over Time



Note: On average, Low, Medium and High Purchase customers make 0.148, 0.338 and 1.219 purchases per month, respectively.

Figure 5: Within Customer Optimal Email Contacts Over Time

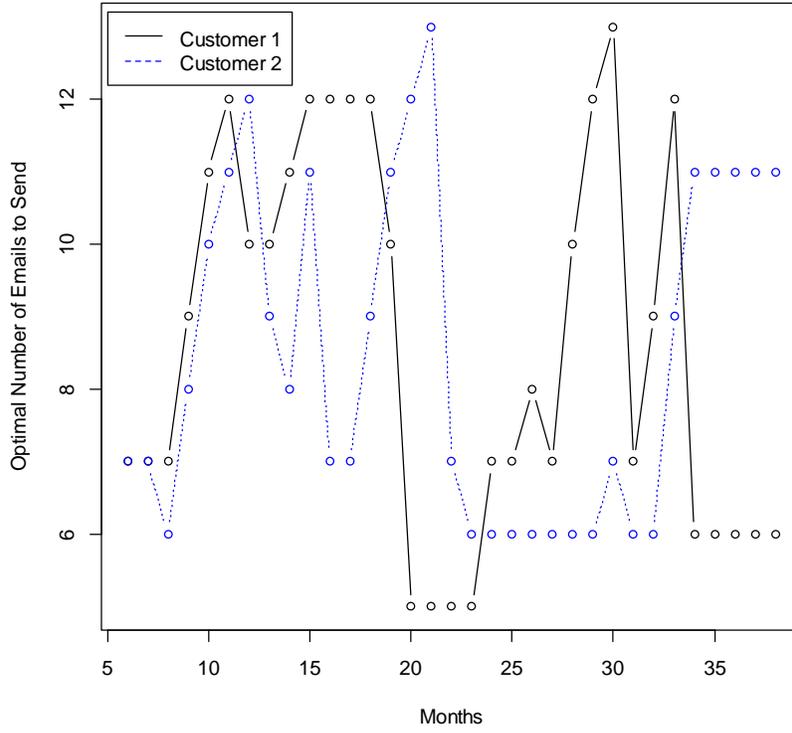


Figure 6: Steady-State CLV versus Number of Email Contacts

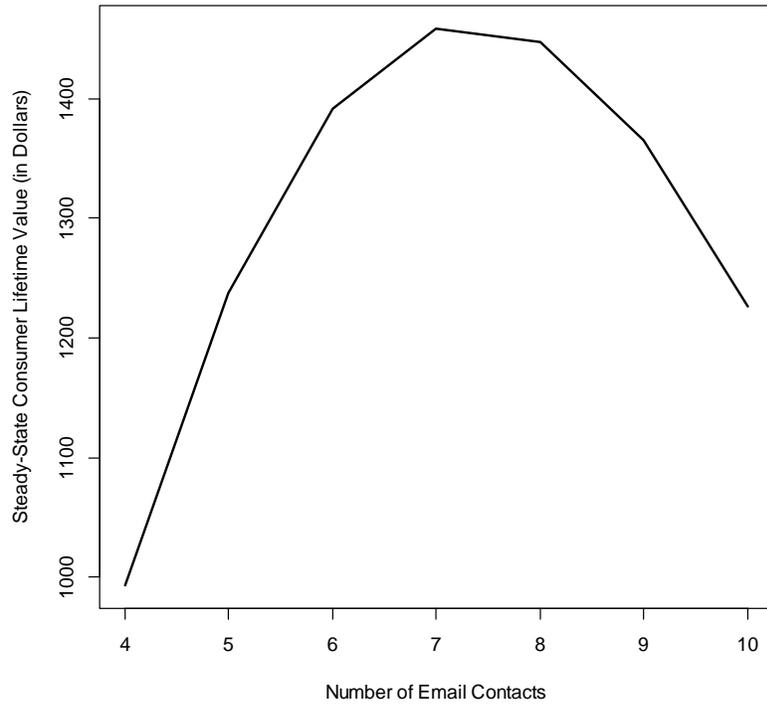


Table 1: Comparison of Existing Studies

Studies	Type of Data	Type of Models		Modeling Customer Relationship over Purchase	Modeling Customer Relationship over Multiple Decisions	Providing a DSS
		Dynamics Component	Marginal Distribution Component			
Netzer et al. (2008)	Donation	HMM	Logit	Yes	No	No
Montoya et al. (2010)	Prescription	HMM	Binomial	Yes	No	Yes
Kumar et al. (2011)	B2B	HMM	Multivariate Tobit	Yes	No	Yes
Li et al. (2011)	Banking	HMM	Multivariate Probit	Yes	No	Yes
Luo and Kumar (2013)	B2B	HMM	Multivariate Tobit	Yes	No	No
Schweidel et al. (2014)	E-commerce	Latent Changeoint	Bivariate Choice	Yes	Yes	No
This Study	Retailing	HMM	Copula of Zero-Inflated Negative Binomial and Binomial Distributions	Yes	Yes	Yes

Table 2: Descriptive Statistics & Correlation Matrix¹¹

	Descriptive Statistics				Correlation Matrix		
	Mean	Std. Dev.	Lower 5%	Upper 95%	Number of Purchases	Number of Emails Sent	Number of Emails Opened
Number of Purchases (per month, per customer)	0.69	1.63	0	4	1		
Number of Emails Sent (per month, per customer)	6.90	4.91	0	15	0.013	1	
Number of Emails Opened (per month, per customer)	1.64	3.16	0	10	0.050	0.425	1
Time from Last Purchase (in months)	6.33	7.16	1	22			
Time from Last Email Open (in months)	5.45	7.23	1	22.05			
Indicator of Open (per month, per customer)	0.40	0.49	0	1			
Indicator of Purchase (per month, per customer)	0.29	0.45	0	1			

Table 3: Selecting the Number of States

HMM States	Log-Likelihood	BIC
1	-17956.41	36064.73
2	-17152.83	34582.67
3	-15119.06	30658.11
4	-15098.16	30777.15

¹¹ The monthly descriptive statistics are calculated from our data consisting N=200 randomly chosen customers (from the retailer's customer database) who are observed for a period of T=39 months.

Table 4: Estimation Results for the 3-State Hidden Markov Model

	Estimates	Standard Error
Transition Matrix		
Intercept for transition (State 1 to 2)	-2.647 ^{***}	0.315
Intercept for transition (State 1 to 3)	-1.432 ^{***}	0.248
Intercept for transition (State 2 to 2)	2.829 ^{***}	0.352
Intercept for transition (State 2 to 3)	-3.815	4.554
Intercept for transition (State 3 to 2)	-2.398 [*]	0.953
Intercept for transition (State 3 to 3)	0.061	0.274
I[Lagged Purchase > 0] on transition to State 2	0.173	0.218
I[Lagged Purchase > 0] on transition to State 3	0.363 [*]	0.150
I[Lagged Open > 0] on transition to State 2	0.422	0.376
I[Lagged Open > 0] on transition to State 3	-0.486 [*]	0.216
Lag Email Sent on transition to State 2	-0.245 ^{**}	0.088
Lag Email Sent on transition to State 3	-0.110	0.064
Lag Email Sent Square on transition to State 2	0.007	0.005
Lag Email Sent Square on transition to State 3	0.001	0.004
Lag Email Sent Residual on transition to State 2	0.144 ^{***}	0.038
Lag Email Sent Residual on transition to State 3	0.003	0.033
Email Open Frequency (Binomial)		
Intercept State 1	-2.367 ^{***}	0.065
Intercept (additional State 2, exp)	1.105 ^{***}	0.016
Intercept (additional State 3, exp)	-4.450 ^{***}	0.969
Time since last open (log)	-0.637 ^{***}	0.067
Variance for the intercept	1.817 ^{***}	0.064
Variance for Time since last open	0.388 ^{***}	0.062
Conditional Purchase Frequency (ZINBD)		
<i>Purchase Count Equation</i>		
Intercept (State 1)	-0.716 ^{***}	0.105
Intercept (State 2)	-2.951 ^{***}	0.312
Intercept (State 3)	-0.413	0.317
Email Sent (State 1)	0.196 ^{***}	0.031
Email Sent (State 2)	0.470 ^{***}	0.011
Email Sent (State 3)	0.321 ^{***}	0.091
Email Sent Square (State 1)	-0.013 ^{***}	0.002
Email Sent Square (State 2)	-0.020 ^{***}	0.001
Email Sent Square (State 3)	-0.027 ^{***}	0.007
Email Sent Residual (State 1)	-0.039 [*]	0.016
Email Sent Residual (State 2)	0.049	0.037
Email Sent Residual (State 3)	-0.002	0.036
Variance for the intercept	1.021 ^{***}	0.049
Variance for Email Sent	0.020 [*]	0.009

Variance for Email Sent Square	0.000	0.001
Dispersion, exp	0.655***	0.071
<i>Excess of Zeros Equation</i>		
Intercept (State 1)	2.135***	0.066
Intercept (State 2)	3.203***	0.710
Intercept (State 3)	3.447***	0.761
Time since last purchase (State 1) (log)	-1.528***	0.073
Time since last purchase (State 2) (log)	-2.235***	0.391
Time since last purchase (State 3) (log)	-1.294***	0.338
Variance for the intercept	2.008***	0.208
Variance for Time since last purchase	0.168	0.091
Correlation (Email Open and Purchase)		
Frank Copula correlation coefficient	0.324***	0.067
Likelihood	-15119.06	
BIC	30658.11	

Legends: * p-value<0.05, ** p-value<0.01, *** p-value<0.001

Labels of the Markov states – State 1: Low Open - Medium Purchase, State 2: High Open - Low Purchase, State 3: High Open - High Purchase

Table 5 Transition Probability Matrix of the HMM

	<i>To Low Purchase</i>	<i>To Medium Purchase</i>	<i>To High Purchase</i>
<i>From Low Purchase</i>			
Without Email Contact	95.38%	4.53%	0.09%
With One Email Contact	94.22%	5.67%	0.10%
With Five Email Contacts	87.93%	11.92%	0.14%
With Ten Email Contacts	78.42%	21.42%	0.16%
<i>From Medium Purchase</i>			
Without Email Contact	6.75%	76.53%	16.72%
With One Email Contact	5.49%	79.02%	15.48%
With Five Email Contacts	2.66%	86.20%	11.14%
With Ten Email Contacts	1.40%	91.27%	7.33%
<i>From High Purchase</i>			
Without Email Contact	5.43%	47.95%	46.62%
With One Email Contact	4.55%	50.99%	44.46%
With Five Email Contacts	2.45%	61.93%	35.62%
With Ten Email Contacts	1.43%	72.60%	25.97%