HOW MUCH IS YOUR CUSTOMER WORTH?

A GAMMA-POISSON MODEL TO ASSESS CUSTOMER PROFITABILITY

P.B. (Seethu) Seetharaman[1]

Qin Zhang[2]

April 25, 2006

[1] Associate Professor of Marketing, Jesse H. Jones School of Management, Rice University, MS-531, 6100 S Main, Houston, TX 77252, e-mail: seethu@rice.edu, Ph: (713) 348-6342, Fax: (713) 348-6296. Authors are listed in alphabetical order and contributed equally to the paper.

[2] Assistant Professor of Marketing, The School of Management, University of Texas at Dallas, Richardson, TX 75083, e-mail: zhangq@utdallas.edu

HOW MUCH IS YOUR CUSTOMER WORTH?
A GAMMA -POISSON MODEL TO ASSESS CUSTOMER PROFITABILITY

Abstract

Assessing the profitability of a customer is very important to a firm from the standpoint of planning retention (and even acquisition) initiatives. Customer profitability is also the defining concept of *customer relationship management* (CRM). In this paper we propose a framework to quantify the expected profitability of a customer. This is done by first postulating that a customer's probability of responding to a marketing solicitation follows a discrete hazard process (specifically, the customer's inter-purchase times are distributed *Poisson*), whose parameters are allowed to vary across customers on account of both *observed heterogeneity* and *unobserved heterogeneity* (specifically, we assume the heterogeneity distribution to be *Gamma*). This yields an unconditional model of customers' inter-purchase times, with the "memory" property, that we call the *Gamma-Poisson Model* (not to be confused with the NBD model, which compounds "memoryless" *Exponential* inter-purchase times with *Gamma* heterogeneity). We then show how to use the estimated parameters to score the firm's customers in terms of their expected profitability to the firm at any point of time. These profitability scores are customer-specific and generated by (1) Bayesian updating of the estimated population-level parameters using the observed customer purchases, as well as (2) the time elapsed since the customer's last purchase with the firm. Our proposed procedure enables the firm to track the time-varying profitability of its customers, and also facilitates the elimination of unprofitable customers as and when appropriate. Lastly, we discuss possible ways of dealing with optimal customer prospecting based on our analyses.

We illustrate our proposed technique using individual-level purchasing data from 529 customers of a catalog firm. First, we estimate the parameters of our proposed *Gamma-Poisson Model*. As an empirical benchmark, we also estimate the *Beta-Binomial Model*, which can be interpreted as a discrete-time analogue of the widely used *NBD Model*. Our proposed model shows superior empirical performance – in terms of both (1) in-sample calibration fit, and (2) holdout validation fit – over the benchmark. We then demonstrate, using a variety of segmentation analyses, the substantive benefits of employing our proposed *Gamma-Poisson Model* in terms of how one could effectively segment customers based on their predicted profitability scores. We believe that our proposed technique will greatly assist marketing efforts of firms that possess customer databases.

Key-words: *Data-based Marketing, Database Marketing, Customer Retention, Customer Profitability, Customer Relationship Management (CRM), Gamma-Poisson, Beta-Binomial, NBD, Customer Acquisition.*

1. Introduction

Assessing customer profitability is integral to the functioning of many firms, especially from the standpoint of customer relationship management (CRM). Understanding the equity of individual customers enables firms to focus their marketing efforts in a more targeted manner, and hence boost the company's bottom-line (Rust, Lemon and Zeithaml 2004). For example, consider catalog marketers such as L.L. Bean, Land's End etc. that by-pass the retailer and sell directly to the end-user. By quantifying the expected profitability of their customers, these firms can understand whether it is worthwhile to incur the costs of periodically contacting these customers with direct-mail solicitations. Suppose it costs Land's End $1 to send a catalog mailing to John Doe each fortnight, plus another $1 for related costs such as support-staff time etc. Unless John Doe has a reasonable chance of buying from the Land's End catalog in the short- to medium-term, this mailing effort will not be worthwhile. Therefore, the challenge for any marketer is to identify and separate the "gold" customers from the "silver" customers and the "lead" customers, so that they can cut down wasteful marketing expenditures incurred by contacting "lead" customers, and instead channel these expenditures at better serving their "gold" and "silver" customers over time. This principle of "weeding out the losers and hanging on to the winners" is the cornerstone of database marketing (Reichheld 1996). In responding to these concerns of marketing practitioners, there has been renewed academic interest over the past few years in modeling customer profitability (see, for example, Gupta, Lehmann and Stuart 2004, Venkatesan and Kumar 2004, Reinartz, Thomas and Kumar 2004, Fader, Hardie and Lee 2005a).

For the purpose of assessing the profitability of a customer, customer-scoring models are used. These models give each customer a score in terms of their probability of responding to a marketing solicitation, and only if the customer's score is high enough to yield expected revenues that are sufficient to offset the firm's costs of contacting the customer, does the firm continue its dialogue with the customer. For a prospect, i.e., a customer with whom the firm has never interacted in the past, the score is typically a function of the prospect's demographic profile, and is constructed based on what the firm has understood from demographically similar customers in the past, using statistical techniques such as discriminant analysis, logistic regression etc. For example, past data may reveal that high-income prospects are more likely to respond to the marketing offer than low-income prospects. For a customer, i.e., a prospect who has "converted" to buying from the firm, the score not only depends on his demographic profile but also is a function of his history of responses to previous solicitations from the company. For example, a customer who buys frequently would have a higher score than another who buys infrequently, even if both have identical demographic profiles.

Our goal in this paper is to develop a customer profitability model to score existing customers of a firm. Our proposed customer profitability model is built on a Gamma-Poisson model of customer buying behavior. In this Gamma-Poisson model (not to be confused with the NBD model of Morrison and Schmittlein (1981), as will be explained later), the customer's inter-purchase times are assumed to be distributed Poisson, and the parameter of this Poisson distribution is assumed to be heterogeneous across customers, and is assumed to be distributed Gamma. We employ a discrete-time distribution such as the Poisson (as opposed to a continuous-time distribution such as, say, the Generalized Gamma, as in Allenby, Leone and Jen (1999) and Venkatesan and Kumar (2004), or Gompertz, as in Ter Hofstede and Wedel 1998) since customers' purchases typically happen in discrete time in response to marketing solicitations made by the firm at regular discrete time intervals. The appeal of our modeling framework is three-fold: (1) The Poisson distribution has been shown in recent research to be a suitable distribution for customers' inter-purchase times (see Boatwright, Borle and Kadane 2003), (2) The Gamma distribution presents a very flexible way of modeling heterogeneity across customers (e.g., Morrison and Schmittlein 1981), and (3) The Gamma distribution is conjugate to the Poisson distribution, which enables us to derive closed-form expressions for the sample likelihood function (which substantially aids practical computation) and, more importantly, the customer profitability metrics derived by us have an analytical closed form. The attractive feature of the Poisson model of inter-purchase times lies in its ability to handle customers' *purchase cycles*, unlike an alternative discrete-time distribution of inter-purchase times, the Geometric distribution (as employed in the Beta-Binomial models of Buchanan and Morrison 1988, Rao and Steckel 1995, Fader, Hardie and Lee 2005a, b), that is based on the assumption of a "memoryless" hazard, which implies that the customer's conditional probability of buying from the firm is the same regardless of the time elapsed since the customer's previous purchase from the firm. For example, suppose a customer has just purchased apparel from the Land's End catalog this month. On account of inventory effects, this customer may not be likely to purchase from the catalog again immediately after the purchase, i.e., his current-period profit may not be positive, but he will start desiring for apparel and return to Land's End to make his next purchase again "after a few months." Not recognizing such purchase cycles, that are heterogeneous across customers, may lead the firm that uses the current-period profitability model to prematurely terminate some profitable customers. Assessing the long-term profitability of a customer, appropriately allowing for a flexible distribution of customers' inter-purchase times (as we do in this paper), on the other hand, could indeed reveal that the customer will be profitable in the future. In the case of the "memoryless" Geometric hazard (as in the Beta-Binomial model of Buchanan and Morrison 1988, Rao and Steckel 1995), the firm's problem of looking at multi-period profits inappropriately reduces to the problem of looking at the current-period profit only (as we will show later), which leads to the problems discussed above. Alternatively, it may be incumbent on the firm to keep sending catalogs regularly to the customer to gradually re-build his interest in buying from the firm and finally "trigger" his next purchase from the firm. In either case, the firm would like to understand the nature of customers' inter-purchase times.

The rest of the paper is organized as follows. In section 2, we develop our proposed model of customer profitability, for which one of the required inputs is a *Gamma-Poisson* model of customer purchasing behavior, which we explicitly derive in this section. We discuss how to estimate the parameters of the proposed *Gamma-Poisson* model using a customer database, and also present the *Beta-Binomial* model as a benchmark model. In section 3, we demonstrate the empirical application of the proposed model using a real-world customer database. We also illustrate the substantive benefits of using our proposed framework using a variety of managerially useful measures of customer segmentation. Section 4 concludes.

2. *Model Development*

We consider a firm that sends marketing solicitations to customers in its mailing list at regular intervals (say, every month). The substantive question of interest is the following: "*Given the firm's regular solicitation schedule, as well as the existence of the firm's customer database (which has recorded past buying behavior of customers), how does one estimate the profitability of each customer at a given point of time?*" We answer this question in four sub-sections: In sub-section 2.1, we present a canonical model of customer profitability; In sub-section 2.2, we derive a parametric model of customer buying behavior, i.e., our proposed *Gamma-Poisson* model, which is a required input in the customer profitability model discussed in sub-section 2.1; In sub-section 2.3, we derive the parametric model of customer profitability that results from plugging in the Gamma-Poisson model in the canonical customer profitability model; In sub-section 2.4, we discuss how the firm can implement our proposed approach in order to score its customers on expected customer profitability.

In sub-section 2.5, we discuss the estimation. Sub-section 2.6 presents the benchmark Beta-Binomial model, while sub-section 2.7 extends our proposed customer profitability framework to handle customer acquisition.

2.1 Canonical Customer Profitability Model

Our customer profitability model is built on the premise that different customers require a different number of cumulative marketing solicitations from the firm before their purchasing interest is "triggered" in the product category. Alternatively, different customers have different *purchasing cycles* within the product category.[3] For example, some customers may make a purchase in response to every marketing solicitation, while others may make a purchase after every three solicitations etc. In such a case, it is important for the firm to take into account such heterogeneity in inter-purchase times across customers, and appropriately compute each customer's expected profitability.

---

[3] Whether the catalogs endogenously "trigger" customer purchases or customers have exogenous purchase cycles is a matter that cannot be empirically resolved using our dataset. A field experiment is necessary to resolve this issue.

Suppose that customer n ($n=1, \ldots, N$) has already received j marketing solicitations from the firm since the customer's previous purchase from the firm. The expected profitability of the customer's next purchase from the firm[4] is given by

$$\pi_{nj} = \sum_{k_n=1}^{\infty} [P_n - k_n c_n] \mathrm{Pr}_{n, j+k_n}, \qquad (1)$$

where P_n is the expected gross profit associated with the product (assumed to be constant across time), c_n is the cost of contacting the customer (also assumed to be constant across time), k_n stands for the number of additional solicitations required to elicit the next purchase from customer n (which is uncertain from the standpoint of the firm and, therefore, is treated as the realization of a *random variable* that takes positive integer values), and $\mathrm{Pr}_{n, j+k_n}$ stands for the associated probability that customer n's next purchase will occur following k_n additional solicitations.

Making a suitable assumption about customer buying behavior within the product category will yield an appropriate distributional assumption for the random variable whose realization is k_n. In other words, $\mathrm{Pr}_{n, j+k_n}$ will take the probability masses that are associated with some parametric discrete distribution that is believed to adequately capture customer buying behavior. We discuss our modeling assumption about such behavior next.

2.2 Gamma-Poisson (GP) Model of Customer Buying Behavior

As mentioned earlier, customers differ from one another in terms of the number of marketing solicitations it will take to induce them to purchase from the firm. Suppose X_n stands for the number of solicitations (since a purchase from the firm) that it will take to induce a purchase from customer n. This is a random variable from the

---

[4] In Appendix A, we show that the expected profitability of the customer's next purchase is a sufficient statistic for the expected lifetime profitability of the customer.

standpoint of the firm. We assume that X_n follows a shifted Poisson distribution as shown below.

$$\Pr(X_n - 1 = x \mid \lambda_n) = \frac{e^{-\lambda_n} \lambda_n^{\ x}}{x!}, (x = 0, 1, 2, \ldots) \tag{2}$$

where $\lambda_n > 0$ stands for the parameter of the Poisson process governing customer n's purchasing behavior.[5] Further, we assume that the Poisson parameter λ_n follows a Gamma distribution across customers, as shown below.

$$g_G(\lambda \mid \alpha, \beta) = \frac{\beta^{\alpha}}{\Gamma(\alpha)} \lambda^{\alpha-1} e^{-\beta\lambda}, \tag{3}$$

where $\alpha, \beta > 0$ are parameters of the Gamma distribution representing the extent of heterogeneity in λ_n across customers in the population.[6] We justify our parametric assumptions next.

The Poisson distribution is a uni-modal distribution of inter-purchase times which, unlike the Exponential distribution (which has been used extensively in previous marketing studies, e.g., the NBD Model of Morrison and Schmittlein 1981, Gupta and Morrison 1991), or its discrete-time counterpart, the Geometric distribution (e.g., the Beta-Binomial, or BB, Model of Buchanan and Morrison 1988, Rao and Steckel 1995, Fader, Hardie and Lee 2005a,b), is not characterized by a flat ("memoryless") hazard

---

[5] It is useful to note that we assume the customer's *inter-purchase times* to be distributed Poisson, unlike many previous studies (e.g., Morrison and Schmittlein 1981) where the customer's *total number of purchases within a given time interval* is assumed to be distributed Poisson, which implies that the customer's inter-purchase times are distributed *Exponential*.

[6] Again, it is useful to note that our Gamma-Poisson model, which has not been previously estimated (as far as we know) in the purchase timing literature, is not equivalent to the NBD model (e.g., Morrison and Schmittlein 1981) which results from compounding Poisson *purchase rates* (not Poisson inter-purchase times) with Gamma heterogeneity. The NBD model assumes Exponential inter-purchase times.

function. The Poisson distribution has been shown in recent studies to be a suitable distribution for customers' inter-purchase times (see, for example, Boatwright, Borle and Kadane 2003, Borle, Boatwright, Kadane, Nunes and Schmueli 2005). The Gamma distribution of heterogeneity admits a variety of flexible shapes in terms of how the Poisson parameter is distributed across customers. In this sense, it is a versatile distribution of heterogeneity while at the same time maintaining an analytical closed form for its density function. Additionally, the Gamma distribution is conjugate to the Poisson distribution, which means, as we will show next in Step 3, that the researcher and practitioner can obtain customer-specific profitability metrics that also have an analytical closed form and are easy to compute.

We can additionally incorporate the effects of *observed heterogeneity* across customers by allowing the parameters of the Gamma distribution – α and β -- to be functions of customer demographics, as shown below (see Gupta and Chintagunta 1994 for a similar specification of observed heterogeneity in a latent class framework, where the probability masses associated with a semi-parametric distribution of heterogeneity are allowed to be functions of customer demographics).

$$\begin{aligned} \alpha_n &= \alpha_o + Z_n \alpha_1, \\ \beta_n &= \beta_o + Z_n \beta_1, \end{aligned} \tag{4}$$

where Z_n is a vector of demographic variables characterizing customer *n*, while α_1 and β_1 are corresponding vectors of parameters, α_o and β_o are scalar parameters.

2.3 Parametric Customer Profitability Model

Given the Gamma-Poisson specification of customer buying behavior discussed in section 2.2, we can plug in the formula for $\Pr_{n,j+k_n}$ that is implied by it within equation (1) to obtain a parametric version of the customer profitability model discussed in section 2.1. At this point, it is useful to recognize that we will obtain different expressions for $\Pr_{n,j+k_n}$, depending on whether customer n makes a purchase in response to the j^{th} solicitation or not. This is because the hazard function associated with Poisson inter-purchase times is not flat, which means that the customer's (conditional) probability of buying from the firm is a function of the time elapsed since the customer's previous purchase from the firm.[7] The expressions for $\Pr_{n,j+k_n}$ under the two cases are summarized below.

1. If customer n makes a purchase in response to the j^{th} solicitation, the customer's probabilities of making the next purchase in response to the $(j+1)^{\text{th}}$, $(j+2)^{\text{th}}$ etc. solicitations are given by

$$\Pr_{n,j+k_n} = \frac{e^{-E(\lambda_n)} E(\lambda_n)^{k_n-1}}{(k_n-1)!}, (k_n = 1, 2, \ldots) \tag{5}$$

where $E(\lambda_n)$ stands for the expected (from the firm's standpoint) value of customer n's Poisson parameter (which will be derived later in this section).

2. If customer n does not respond to the j^{th} solicitation, the customer's probabilities of making the next purchase in response to the $(j+1)^{\text{th}}$, $(j+2)^{\text{th}}$ etc. solicitations are given by

$$\Pr_{n,j+k_n} = \frac{\dfrac{e^{-E(\lambda_n)} E(\lambda_n)^{(k_n-1+j-j_o)}}{[k_n-1+j-j_o]!}}{1-\displaystyle\sum_{u=1}^{j-j_o} \frac{e^{-\hat{\lambda}} E(\lambda_n)^{(u-1)}}{(u-1)!}}, (k_n = 1, 2, \ldots) \tag{6}$$

where j_0 stands for the last solicitation to which the customer responded.

Correspondingly, by plugging in equations (5) and (6) within equation (1), we obtain the following expressions for customer profitability under the two cases.

---

[7] If inter-purchase times are assumed to be Exponential (as in the NBD model) or Geometric (as in the BB model), however, the customer's conditional purchasing probability will be the same regardless of the time elapsed since the customer's previous purchase. We believe such an assumption to be unrealistic. To test if our belief is validated by the data, we benchmark our specification against the BB specification later.

1. If the customer makes a purchase in response to the jth solicitation, the customer's profitability is

$$\pi_{nj} = P_n - c_n \sum_{k_n=1}^{\infty} k_n \Pr_{n,j+k} = P_n - c_n(E(\lambda_n)+1). \tag{7}$$

2. If the customer does not respond to the jth solicitation, the customer's profitability is

$$\begin{aligned}\pi_{nj} &= P_n - c_n \sum_{k_n=1}^{\infty} k_n \frac{\dfrac{e^{-E(\lambda_n)} E(\lambda_n)^{(k_n-1+j-j_o)}}{[k_n-1+j-j_o]!}}{1-\displaystyle\sum_{u=1}^{j-j_o} \frac{e^{-E(\lambda_n)} E(\lambda_n)^{(u-1)}}{(u-1)!}} \\ &= P_n - \frac{c_n(E(\lambda_n)-y_n)}{1-y_n} + c_n(j-j_0-1),\end{aligned} \tag{8}$$

where

$$y_n = \sum_{u=1}^{j-j_0} \frac{e^{-E(\lambda_n)} E(\lambda_n)^{(u-1)}}{(u-1)!}. \tag{9}$$[8]

Next, we derive the customer-specific $E(\lambda_n)$ given in equations (7) and (8). We present this derivation separately for four specific cases, representing different levels of information that the firm may have on customer n.[9]

Case 1: No information on a Customer

If the firm has neither demographic information nor historical purchasing behavior of a particular customer n, $E(\lambda_n)$ is taken by the firm to be equal to the mean of the Gamma distribution in equation (3). In other words,

$$E(\lambda_n) = \frac{\alpha}{\beta}. \tag{10}$$

Case 2: Demographic Information Only

---

[8] See Appendix B for the derivation of this expression.

[9] These cases are in the spirit of the cases employed by Rossi, McCulloch and Allenby (1996) and Iyengar, Ansari and Gupta (2003), to discuss how different levels of information lead to different levels of predictive ability for the firm.

If the firm knows the demographic profile, but not the historical purchasing behavior, of a particular customer n, it can modify equation (10) using equation (4), as shown below.

$$E(\lambda_n) = \frac{\alpha_o + Z_n \alpha_1}{\beta_o + Z_n \beta_1}. \tag{11}$$

In this case, therefore, the customer's demographic profile Z_n is allowed to influence the expected inter-purchase time of the customer.

Case 3: Purchasing Behavior Only

If the firm knows the historical purchasing behavior, but not the demographic profile, of a particular customer n, it can employ a Bayesian updating procedure that exploits the conjugacy of the Gamma distribution of heterogeneity with respect to the Poisson distribution of the customer's inter-purchase times. Suppose T_n stands for the number of historical purchases made by the customer in the past, and M_{Tn} stands for the total number of solicitations required by the firm to induce these T_n purchases. The mean of the "prior" Gamma distribution given in equation (3) is Bayesian updated to obtain the following mean of a customer-specific "posterior" Gamma distribution.

$$E(\lambda_n) = \frac{\alpha + M_{T_n}}{\beta + T_n}, \tag{12}$$

Case 4: Full Information

If the firm knows both the demographic profile, as well as the historical purchasing behavior, of a particular customer n, this results in the following mean of the customer-specific "posterior" Gamma distribution.

$$E(\lambda_n) = \frac{\alpha_o + Z_n \alpha_1 + M_{T_n}}{\beta_o + Z_n \beta + T_n}. \tag{13}$$

This completes the exposition of our proposed customer profitability model. Next, we discuss how the firm can implement the proposed customer profitability framework to score its customers in terms of expected customer profitability.

2.4 Implementation Algorithm for the Firm

From a practical implementation standpoint, the firm can proceed as follows:

Step 1: Estimate the parameters -- $\alpha_0, \alpha_1, \beta_0, \beta_1$ -- of the Gamma distribution given in equations (3) and (4) using historical purchasing data from a "training sample" of customers in the firm's customer database (we will discuss this estimation procedure in the next sub-section 2.5).

Step 2: Depending on the level of information that the firm has on a particular "holdout" customer *n*, whose profitability the firm wants to assess, use one among equations (10)-(13) to estimate the expected value of customer *n*'s inter-purchase time, i.e., $E(\lambda_n)$.

Step 3: Depending on whether or not customer *n*, whose profitability the firm wants to assess, made a purchase in response to the most recent solicitation (*j*) from the firm, use equation (7) or (8) to assess the current expected profitability of the customer, i.e., π_{nj}.

2.5 Estimation

The firm must use an appropriate subset of customers represented in its customer database ("training sample") to estimate model parameters, i.e., α_o, β_o, α_1, β_1, and then predict a profitability score for the remaining customers in its customer database ("holdout sample") using one among equations (10)-(13). Here, we will discuss how to estimate model parameters using the training sample.

Suppose the customer database is over the calendar period (t_1, t_2).[10] Let us also suppose that new customer acquisition activity has ceased so that no new customers are acquired by the firm during this period. In that case, the customer database is left-truncated in that the first purchase made by a customer during the study period, in general, is at some point following calendar time t_1. The customer database is also right-

[10] For example, the firm may retain the "most recent" 5-years of data and expunge earlier data.

truncated in that the most recent purchase of a customer, in general, is at some point prior to calendar time t_2. Within the study period, we observe each purchase made by each customer in the training sample. Recognizing these properties of the training sample, we set up the likelihood function for customer *n* as shown below.[11]

$$L_n = \int_0^{\infty} \Pr(X_n > l_n \mid \lambda_n) \left[\prod_{t=2}^{T_n} \Pr(X_n = x_t \mid \lambda_n) \right] \Pr(X_n > r_n \mid \lambda_n) g_G(\lambda_n \mid \alpha_n, \beta_n) d\lambda_n \qquad (14)$$

where l_n is the number of solicitations received by the customer from t_1 until the time of their first purchase during the study period, r_n stands for the number of solicitations received by the customer between their last purchase during the study period and t_2, and $g_G(.)$ is the density of the Gamma distribution given in equations (3) and (4). This can be simplified to[12]

$$L_n = W_n \left[\frac{1}{(\beta_n + T_n - 1)^{\alpha_n + \sum_{t=2}^{T_n} x_t}} - \frac{Y_R}{(\beta_n + T_n)^{\alpha_n + \sum_{t=2}^{T_n} x_t}} - \frac{Y_L}{(\beta_n + T_n)^{\alpha_n + \sum_{t=2}^{T_n} x_t}} + \frac{Y_{RL}}{(\beta_n + T_n + 1)^{\alpha_n + \sum_{t=2}^{T_n} x_t}} \right], \qquad (15)$$

where

$$W_n = \frac{\left(\alpha_n (\alpha_n + 1) \ldots \left(\alpha_n + \sum_{t=2}^{T_n} x_t - 1 \right) \right) \beta_n^{\alpha_n}}{(x_2!) \ldots (x_{T_n}!)}. \qquad (16)$$

and

$$Y_L = 1 + \frac{\alpha_n + \sum_{t=2}^{T_n} x_t}{\beta_n + T_n} + \frac{\left(\alpha_n + \sum_{t=2}^{T_n} x_t \right) \left(\alpha_n + \sum_{t=2}^{T_n} x_t + 1 \right)}{(\beta_n + T_n)^2 \, 2!} + \ldots + \frac{\left(\alpha_n + \sum_{t=2}^{T_n} x_t \right) \left(\alpha_n + \sum_{t=2}^{T_n} x_t + 1 \right) \ldots \left(\alpha_n + \sum_{t=2}^{T_n} x_t \right) \left(\alpha_n + \sum_{t=2}^{T_n} x_t + l_n - 2 \right)}{(\beta_n + T_n)^{(l_n - 1)} (l_n - 1)!} \qquad (17)$$

[11] Since the first purchase is left-truncated, as discussed above, it is subsumed within the left truncation likelihood. The likelihood for the observed inter-purchase times starts from the second purchase onwards.

[12] See Appendix C for the derivation.

and

$$Y_R = 1 + \frac{\alpha_n + \sum_{t=2}^{T_n} x_t}{\beta_n + T_n} + \frac{\left(\alpha_n + \sum_{t=2}^{T_n} x_t\right)\left(\alpha_n + \sum_{t=2}^{T_n} x_t + 1\right)}{\left(\beta_n + T_n\right)^2 2!} + \ldots + \frac{\left(\alpha_n + \sum_{t=2}^{T_n} x_t\right)\left(\alpha_n + \sum_{t=2}^{T_n} x_t + 1\right)\ldots\left(\alpha_n + \sum_{t=2}^{T_n} x_t\right)\left(\alpha_n + \sum_{t=2}^{T_n} x_t + r_n - 2\right)}{\left(\beta_n + T_n\right)^{(r_n - 1)}(r_n - 1)!} \tag{18}$$

and

$$Y_{RL} = \sum_{j=0}^{l_n - 1}\sum_{k=0}^{r_n - 1}\left[\frac{\left(\alpha_n + \sum_{t=2}^{T_n} x_t\right)\left(\alpha_n + \sum_{t=2}^{T_n} x_t + 1\right)\ldots\left(\alpha_n + \sum_{t=2}^{T_n} x_t + j + k - 1\right)}{\left(\beta_n + T_n + 1\right)^{j+k}\; j!k!}\right] \tag{19}$$

After pooling data across all customers in the training sample, the following sample likelihood function can be maximized to obtain estimates of model parameters, i.e., α_o, β_o, α_1, β_1.

$$L = \prod_{n=1}^{N} L_n \tag{20}$$

2.6 Benchmark Model of Customer Buying Behavior: Beta-Binomial (BB)

Under this model (Buchanan and Morrison 1988, Rao and Steckel 1995, Fader, Hardie and Lee 2005a, b), X_n follows a Geometric distribution as shown below.

$$\Pr(X_n = x \mid p_n) = p_n(1 - p_n)^{x-1}, (x = 1, 2, \ldots) \tag{21}$$

where $p_n > 0$ stands for the parameter of the Binomial process governing customer n's probability of purchasing in response to a solicitation. Further, the Binomial parameter p_n is assumed to follow a Beta distribution across customers, as shown below.

$$g_B(p \mid a, b) = \frac{1}{B(a,b)} p^{a-1}(1-p)^{b-1}, \tag{22}$$

where B(a,b) is the Beta function given by

$$B(a,b) = \int_0^1 u^{a-1}(1-u)^{b-1} du, \tag{23}$$

The Geometric distribution is the discrete-time counterpart to the Exponential distribution and has the restrictive property that the customer's probability of re-purchasing from the firm is the highest immediately after purchasing from the firm and progressively falls thereafter. Said differently, the hazard function associated with the Geometric distribution is flat, which implies that the customer's conditional probability of purchasing from the firm is the same regardless of the time elapsed since the customer's previous purchase from the firm. Such "memoryless" distributions have been shown to be inappropriate for purchase timing data (see Seetharaman and Chintagunta 2003). However, the Beta distribution of heterogeneity admits a variety of flexible shapes in terms of how the Binomial parameter is distributed across customers. Additionally, the Beta distribution is conjugate to the Binomial distribution, which means that this benchmark also presents an analytically alternative model specification to the researcher and practitioner.

The effects of *observed heterogeneity* across customers can be incorporated by allowing the parameters of the Beta distribution – *a* and *b* -- to be functions of customer demographics, as shown below.

$$\begin{aligned} a_n &= a_o + Z_n a_1, \\ b_n &= b_o + Z_n b_1, \end{aligned} \tag{24}$$

where Z_n is a vector of demographic variables characterizing customer *n*, while a_1 and b_1 are corresponding vectors of parameters, a_0 and b_0 are scalar parameters.

The expected profitability of customer *n* can be computed by plugging equation (21) within equation (1), which yields [13]

$$\pi_{nj} = \sum_{x=1}^{\infty} [P_n - xc_n](1-p_n)^{x-1} p_n = \frac{P_n . p_n - c_n}{p_n} \tag{25}$$

Therefore, under the BB specification, in order to see whether a customer is profitable in the long term, it is sufficient to check whether the expected profit in the current period (i.e., $P_n . p_n - c_n$) is positive or not. The reason for this is that the hazard characterizing the

[13] See Appendix D for the derivation of this expression.

customer's Binomial purchasing behavior, i.e. p_n, is constant over time. However, under the Poisson purchasing behavior implied by our proposed GP model, this equivalence no longer holds (since the Poisson hazard depends on the time elapsed since the previous purchase, as is made clear by equations (5) and (6)). Therefore, one may conclude from a single-period perspective of the BB model, for example, that a customer is unprofitable while a long-term computation of the GP model may show the customer to be profitable. This is a significant substantive benefit of our proposed GP model over the benchmark BB model.

The BB counterparts to the four specific cases of information levels represented in equations (10)-(13) (for the GP model) are given below.

Case 1: No information on a Customer

$$E(p_n) = \frac{a}{a+b}. \tag{26}$$

Case 2: Demographic Information Only

$$E(p_n) = \frac{a_o + Z_n a_1}{a_o + Z_n a_1 + b_o + Z_n b_1}. \tag{27}$$

Case 3: Purchasing Behavior Only

$$E(p_n) = \frac{a + T_n}{a + b + M_{T_n}}, \tag{28}$$

Case 4: Full Information

$$E(p_n) = \frac{a_o + Z_n a_1 + T_n}{a_o + Z_n a_1 + b_o + Z_n b_1 + M_{T_n}}, \tag{29}$$

Under the BB model, the likelihood function for customer n is as shown below.

$$L_n = \int_0^1 p_n^{T_n} (1 - p_n)^{l_n + M_{T_n - 1} + r_n - T_n} g_B(p_n \mid a_n, b_n) dp_n \tag{30}$$

where $g_B(.)$ is the density of the Beta distribution given in equations (19)-(21), and M_{Tn-1} stands for the number of mailings sent by the firm between the consumer's first and last purchase. As mentioned earlier, l_n stands for the number of mailings sent prior to the first purchase, while r_n stands for the number of mailings sent after the last purchase.

This can be simplified to[14]

$$L_n = \frac{\Gamma(a_n + b_n)\Gamma(a_n + T_n)\Gamma\left(b_n + l_n + M_{T_n - 1} + r_n - T_n\right)}{\Gamma(a_n)\Gamma(b_n)\Gamma\left(a_n + b_n + l_n + M_{T_n - 1} + r_n\right)} \tag{31}$$

After pooling data across all customers in the training sample, the sample likelihood function in equation (20) can be maximized to obtain estimates of model parameters, i.e., a_o, b_o, a_1, b_1.

2.7 Customer Acquisition Model

Suppose the firm has obtained a prospect mailing list (acquired from a list broker such as Acxiom). The firm can decide whether or not to include prospect *n* in their mailing list by computing the expected profitability of the prospect's first purchase, as given below.

$$\pi_n = \sum_{x=0}^{\infty} [P_n - (x+1)c_n] \Pr(X_n - 1 = x), \tag{32}$$

where X_n stands for the number of solicitations required to elicit a first response from prospect *n*, and $\Pr(X_n - 1 = x)$ stands for the associated Poisson probability (as in equation 2). This implies

$$\pi_n = P_n - c_n \sum_{x=0}^{\infty} [x+1] \Pr(X_n - 1 = x) = P_n - c_n \sum_{x=0}^{\infty} [x+1] \frac{e^{-\lambda_n} E(\lambda_n)^x}{x!} = P_n - c_n (E(\lambda_n) + 1), \tag{33}$$

which says that the profitability of a prospect is the gross profit that he contributes to the company (*P*) minus the total cost to the firm of the expected number of solicitations sent to the prospect. The prospect-specific $E(\lambda_n)$ can estimated using equations (10) or (11) depending on whether or not the purchased list contains the prospect's demographic profile. If this profitability is positive, the prospect should be sent the direct-mail solicitation, otherwise he should not.

What if the firm in question has no customer database (in other words, does not

---

[14] See Appendix E for the derivation.

know the values of α and β in order to use equation 10)? The only way for such a firm to identify good prospects is to engage in the following procedure: use a random sample of N prospects from the purchased mailing list, send T solicitations to each prospect, and record the purchase times (if any) of each prospect. The parameters α and β can be estimated by maximizing the following likelihood function.

$$L_n = \int_0^{\infty} \left(\prod_{t=1}^{T_n} \Pr(X_n = x_t \mid \lambda_n) \right) \Pr(X_n > T - \sum_{t=1}^{T_n} x_t \mid \lambda_n) g_G(\lambda_n \mid \alpha_n, \beta_n) d\lambda_n \qquad (34)$$

It is possible that some prospects will never respond to a solicitation. Therefore, the firm may use a maximum number of test-mailings, say T, and maximize the following likelihood function that includes a right-censored component representing non-responders.

$$L_n = \int_0^{\infty} \Pr(X_n > T \mid \lambda_n) g_G(\lambda_n \mid \alpha_n, \beta_n) d\lambda_n \qquad (35)$$

Maximizing a sample likelihood function, constructed using equations (34) and (35), yields estimates of α and β which when plugged into equation (10) can be used to score prospects based on their expected customer profitability. If there is a population of N' prospects represented in the purchased mailing list, the firm can rank order them in terms of decreasing customer profitability, and then contact customers in that rank-ordered sequence until a pre-specified number of prospects (based on budget considerations) is reached.

3. Data and Empirical Analyses

3.1 Data

In order to demonstrate a real-world application of our proposed model, we use a customer database of a company (whose identity cannot be revealed for purposes of confidentiality) that publishes safety, wellness, and health promotion education materials for both the healthcare industry and the workplace. The company provides booklets, brochures, videos, workbooks, and electronic products to hospitals, healthcare providers, care management organizations, government agencies, and employers. The major sales channel is through catalogs. Each catalog includes a key code that a customer could use when placing the order. The key code helps the company track the performance of the catalog that is sent to the customer.

The dataset contains information on the ordering of brochures by 529 healthcare professionals who made two or more orders during the period from July 2002 to August 2003. The dataset also contains information on catalogs that were sent to those customers during this period, including both the number and types of those catalogs, both of which may vary across customers. The average number of catalogs a customer receives over the period is 7.8, and the average number of catalogs to which each customer responds is 2.2. The average quantity ordered per response is 33 brochures and the average gross profit when a customer responds to the catalog is $58.60, where the gross profit is defined as the revenues that are generated from the transaction minus the costs of producing the products sold, which is about $0.10 per brochure. The average mailing cost per catalog is about $1.80. We randomly select 424 customers as the calibration sample and the remaining 105 customers as the holdout sample. The dataset also includes two demographic characteristics of customers:

1. Customer Size: *Private Practice* customers with fewer than six doctors, versus *Group Practice* customers with six or more doctors.

2. Customer profile: *Primary doctors* (i.e., doctors of family practice or internal medicine) versus S*pecialists* versus *Others*.

We code Customer Size using an indicator variable that takes the value 1 for a *Private Practice* customer, and the value 0 otherwise. We code Customer Profile using two indicator variables, the first of which takes the value 1 for a *Primary Doctor*, and the value 0 otherwise, the second of which takes the value 1 for a *Specialist*, and the value 0 otherwise.

3.2 Empirical Results

First, we estimate the proposed Gamma-Poisson model, as well as the benchmark Beta-Binomial model, using the calibration sample ("training sample") of 424 customers and applying the Maximum Likelihood technique (see equations (14)-(20) for the GP model, and equations (30)-(31) and equation (20) for the BB model). We then compute a validation log-likelihood measure for the observed purchases in the holdout sample, based on the estimates from the calibration sample. The fit results are reported in Table 1. Based on fit criteria, we conclude that our proposed GP model is superior to the benchmark BB model in terms of both explaining the observed purchases in the calibration sample (calibration log-likelihood values of -1490.82 versus -1967.63) and predicting observed purchases in the holdout sample (validation log-likelihood values of -369.81 versus -489.02). These fit and validation results lend strong empirical support to our proposed GP model over the benchmark BB model. It is useful to note here that we estimate the *Standard Gamma* distribution (that restricts the scale parameter of the Gamma distribution to be equal to 1, i.e., $\beta = 1$, and only estimates the shape parameter, i.e., α)[15] in the empirical application. Since the standard Gamma has one fewer parameter compared to the general Gamma distribution, and therefore also compared to the Beta distribution, we are pleased to note that the implemented version of our proposed GP model convincingly outperforms the benchmark BB model despite having one fewer parameter to estimate! In other words, if one adjusts the maximized log-likelihood for the number of parameters in the model (using criteria such as AIC, CAIC, BIC etc.), our model will be found to outperform the benchmark BB model by an even wider margin.

---

[15] We imposed this restriction since there were convergence problems when we tried to freely estimate the scale parameter β. Based on a manual grid search on the space of β, we restricted it to be equal to 1 since this yielded a larger value of the log-likelihood compared to other values of β.

(Insert Table 1 here)

Next, we report the estimates yielded by the GP model in Table 2. Our estimated value of α_0 is found to be 3.81. This implies that the baseline (i.e., ignoring customer demographics) average estimated inter-purchase cycle across the 424 customers in the calibration sample is 4.81 mailings. In contrast, the benchmark BB model implies a baseline purchasing probability for the customer of 0.27, which implies a baseline average inter-purchase cycle of 2.66 mailings. In other words, the benchmark BB model is over-optimistic about customer inter-purchase cycles than our proposed GP model. Further, we find in Table 2 that none of the parameters in the GP model that are associated with customer demographics (i.e., α_1) is statistically significant. This implies that the estimated heterogeneity across customers in our dataset cannot be explained using customer demographics, i.e., there is no *observed heterogeneity* across customers although there is significant *unobserved heterogeneity*. This finding about the lack of usefulness of customer demographic variables in explaining across-customer heterogeneity is consistent with findings obtained using models of inter-purchase times in the packaged goods context (see, for example, Seetharaman and Chintagunta 2003).

(Insert Table 2 here)

Next, we score our 105 holdout customers based on their expected profitability using the procedure detailed in section 2.3. In our empirical application, since we observe both demographic profiles and historical purchasing behavior of all 105 customers in the holdout sample, we use equation (13) to compute each customer's average inter-purchase cycle (that is a required input to calculate the customer's profitability). We score each holdout customer at the time of the last mailing from the firm to the customer. We generate the cumulative response curve for the firm by rank-ordering the 105 holdout customers in decreasing order of their expected profitability, and then plot the expected cumulative profits as a function of the number of customers contacted (where we assume that customers are contacted in decreasing order of expected profitability). We generate these plots for rank-ordering implied by three models: (1) the GP model, (2) the BB model, and (3) a naïve model that ranks customers at random.[16] We report the difference between the three plots in Figure 1. Specifically, Figure 1 plots the difference in expected cumulative profits that will accrue to the firm from using our proposed GP model, rather than using the benchmark BB model or the naïve model, when ranking its customers. Figure 1(a) reports the plot for the calibration sample, while Figure 1(b) reports it for the holdout sample. First, it is useful to note that these plots intersect the x-axis both at the beginning and at the end. In other words, as long as the firm either (a) does not contact any customer at all, or (b) decides to contact all customers in the sample, the GP model offers no benefit over either the BB model or the naïve model, which is not surprising. What is important about the plots, however, is that the difference in cumulative profits between not only the GP and naive model, but also the GP and BB model, is *always positive* as long as one is not at the end-points of the curve. In other words, regardless of the nature of the "cut-off" used by the firm, as long as the firm wants to exploit the rank-ordering of customers and contact only a subset (top J customers, for example) of its customer database, the GP model yields higher profits than both the BB model and the naïve model. The difference between the GP and BB model in absolute profit terms is most pronounced if the firm decides to contact the top 380 (70) customers in the calibration (holdout) sample. The resulting difference in profits is about \$425 (\$90) on a total profit of \$20,000 (\$3600), which translates to a percentage difference of 2.1% (2.5%). The percentage differences between the GP and BB

[16] Under the BB model, although we rank-order customers based on its implied customer profitability metrics, we must still calculate cumulative profits for the y-axis based on what is implied by the GP model, since the GP model is assumed to be the "true" model as it fits and validates customer purchases better, as shown in Table 1. Under the naïve model, of course, there are no implied profits, so it is easier to understand that the cumulative profits are calculated based on the GP model.

models are found to vary from 2-5% depending on what fraction of the customer database is chosen for mailing. To the extent that the BB model itself is clearly a modeling improvement over the naïve scoring models (such as RFM) that are typically used in practice, the fact that our GP model produces an incremental profit lift of 2-5% over the BB model speaks well of the practical efficacy of our proposed model. In relation to the naïve model, our proposed GP model wins even more convincingly, as is made clear by Figure 1. In percentage terms, the difference between the profits implied by the GP model and the profits implied by the naïve model is found to be as high as 51%!

(Insert Figure 1 here)

Next, we report the expected customer profitability implied by the GP and BB models for each customer in the sample. This is reported for the calibration sample in Figure 2(a) and for the validation sample in Figure 2(b). Interestingly, the customer profitability plot under the GP model is quite "volatile" (i.e., heterogeneous across customers), while it is "smooth" under the BB model. This means that the BB model severely understates the extent of heterogeneity in customer profitability across customers. Furthermore, and more importantly, for some customers, the BB model significantly overstates their expected profitability (which is consistent with our earlier observation that the implied average inter-purchase time is lower under the estimated BB model than under the estimated GP model). For example, let us consider customer # 347 (91) in the calibration (holdout) sample. For this customer, the BB model predicts a profitability of about \$50 (\$50), while the GP model predicts a profitability of -\$5 (\$20)! This translates to a percentage difference of 1200% (150%)! This critically underscores the importance of using a Poisson process as governing customer buying behavior when purchase cycles characterize customers' purchases. Instead, assuming purchases to be Binomial leads the BB model to be unable to estimate the true extent of heterogeneity in customer purchase behavior, even though the Beta distribution is a flexible distribution of heterogeneity. In fact, the extent of heterogeneity is so severely under-estimated under the BB model that all customers are implied to be almost equally profitable to the firm! The GP model, on the other hand, is able to correctly estimate heterogeneous purchase cycles across customers which, in turn, translate into a truly heterogeneous distribution of customer profitability metrics across customers, as is evident in Figure 2.

(Insert Figure 2 here)

In order to understand the substantive consequences to the firm of using the wrong distribution of customer profitability metrics across customers, as implied by the BB model, as opposed to our proposed GP model, we ask the following managerial question, "What are the profit consequences of falsely including a customer for mailing when, in fact, the customer will be unprofitable?" The profit answer to this question presents a *realized loss* to the firm. We also ask the related question,"What are the profit consequences of falsely excluding a customer from the mailing list when, in fact, the customer will be profitable?" The answer to this question presents an *opportunity cost* (of a lost customer) to the firm. We compute both measures under various assumptions of the variable cost, c, of mailing to a customer. The *realized loss* is computed as the sum of the expected losses across all customers who will be wrongly included by the BB model in the mailing list (when the GP model shows that they are unprofitable). The *opportunity cost* is computed as the sum of the expected lost profits across all customers who will be wrongly excluded by the BB model from the mailing list (when the GP model shows that they are profitable). We report both measures – realized loss and opportunity cost – under various assumptions of c in Figure 3. Interestingly, we see that both the realized loss on wrongly included customers, as well as the opportunity cost of wrongly excluded customers, is highest for an intermediate value of the mailing cost c. This can be understood as follows: for very small values of c, practically every customer will be profitable, so the value of the GP model over the BB model in terms of correctly segmenting customers is likely to be minimal. Similarly, for very large values of *c*, practically every customer will be unprofitable, so the value of the GP model over the BB model in terms of correctly segmenting customers is likely to be minimal. Only for intermediate values of *c* is segmentation beneficial, and the corresponding ability of the GP model to correctly include good customers and exclude bad customers gives it an edge over the BB model.

(Insert Figure 3 here)

6. Conclusions

In this study, we propose a customer profitability model that can be used to segment a firm's customer database in terms of customers' expected profitability. The proposed model is built on a Gamma-Poisson model of customer buying behavior that treats customers' inter-purchase times as a Poisson random variable (unlike the Beta-Binomial model that treats customers' decisions to buy or not in a given period as a Binomial random variable). We show that our proposed GP model outperforms the BB model in terms of both explaining customers' purchases in a calibration sample, and predicting customers' purchases in a holdout sample. When we apply our proposed customer profitability model to score customers in the sample in terms of their expected profitability, we are able to uncover significant heterogeneity across customers. However, the benchmark BB model severely underestimates the extent of heterogeneity in customer profitability metrics across customers. We demonstrate the profit consequences to the firm of using the incorrectly specified BB model rather than our proposed GP model.

There are some important areas of future investigation. First, it would be of interest to test whether alternative distributions – such as Generalized Gamma (see Allenby, Leone and Jen 2005) – better characterize customers' inter-purchase times compared to the Poisson distribution tested in this study.[17] Second, a customer's probability of buying from the firm, as well as the gross profit from the customer to the firm, may systematically change (increase) over time as the customer gains experience with the firm. It may be useful to incorporate such effects within the proposed framework, and apply them to customer databases that span a longer time interval (such as 10-15 years). Third, customers' purchasing decisions may be multi-dimensional, for example manifesting themselves as whether or not to buy from the firm, product choices conditional on purchase (Li, Sun and Wilcox 2005), product expenditure conditional on product choices etc. Incorporating such multi-dimensional decision-making within the proposed framework will be useful. Fourth, it will be of interest to endogenize customer-specific solicitation frequencies for the firm (as in Gonul, Kim and Shi 2000, Gonul and Hofstede 2006). We abstract from all of these issues in this paper for three main reasons: (1) our interest is in demonstrating the importance of modeling customers' purchase cycles, or the cumulative "triggering" effect of marketing solicitations on customers' purchasing behavior, and this can be done using an inter-purchase timing model such as our Gamma-Poisson model, (2) our customer database does not show significant variations in customers' order quantities, expenditures or gross profits over time, and (3) we want to exploit the conjugacy of the Gamma and Poisson distributions to propose an analytically appealing and easy-to-implement modeling approach that firms for whom predicting customers' inter-purchase times is the main concern can effectively use to score their customers in terms of their expected profitability. In conclusion, we believe that our proposed customer profitability framework will greatly assist firms that possess customer databases.

[17] The Generalized Gamma does not have a conjugate distribution that is suitable to model heterogeneity across customers. Therefore, one would not obtain analytical closed-forms for customer profitability as in this study. This will hamper its ease of use in practical applications.

TABLE 1: Fit Results

Fit Criterion	**Fit Value**
In-Sample Log-Likelihood for GP Model	-1490.82
In-Sample Log-Likelihood for BB Model	-1967.63
Holdout Validation Log-Likelihood for GP Model	-369.81
Holdout Validation Log-Likelihood for BB Model	-489.02

TABLE 2: Estimation Results for GP Model (Standard Errors within Parentheses)

Number of customers in Calibration Sample: 424

Parameter	**Estimate**
Constant (α_0)	3.81 (0.35)
$\alpha_{1,\text{Private Practice}}$	0.09 (0.36)
$\alpha_{1,\text{Primary Doctors}}$	0.17 (0.97)
$\alpha_{1,\text{Specialists}}$	-0.11 (0.25)

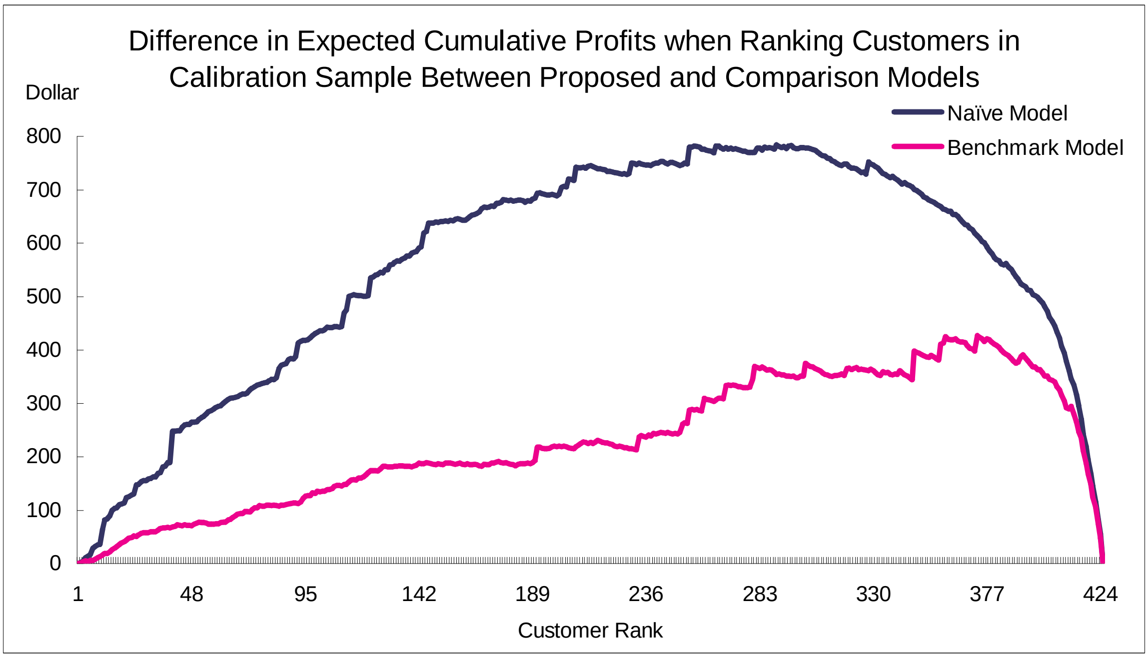

FIGURE 1(a)

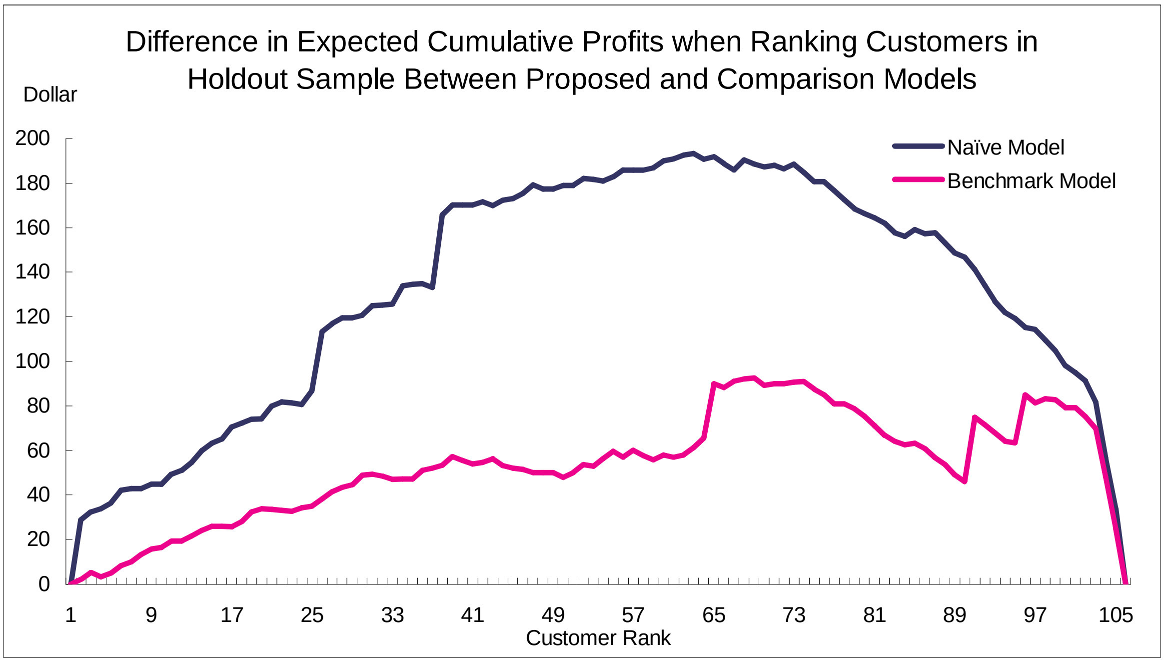

FIGURE 1(b)

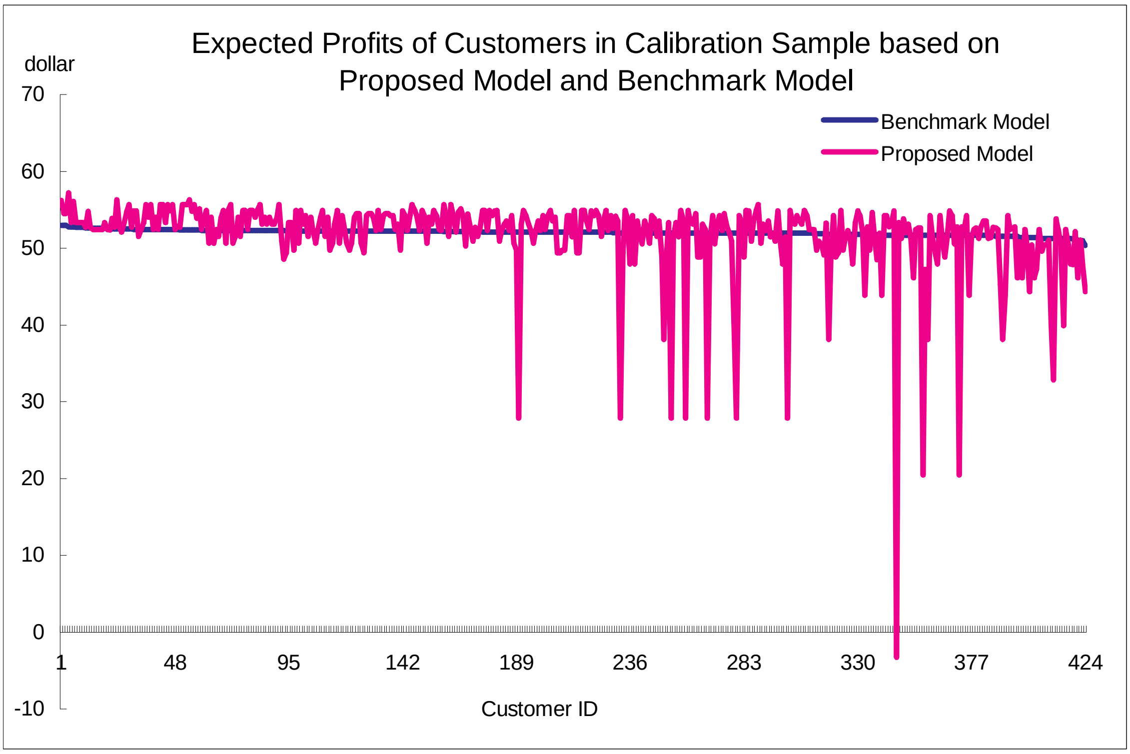

FIGURE 2(a)

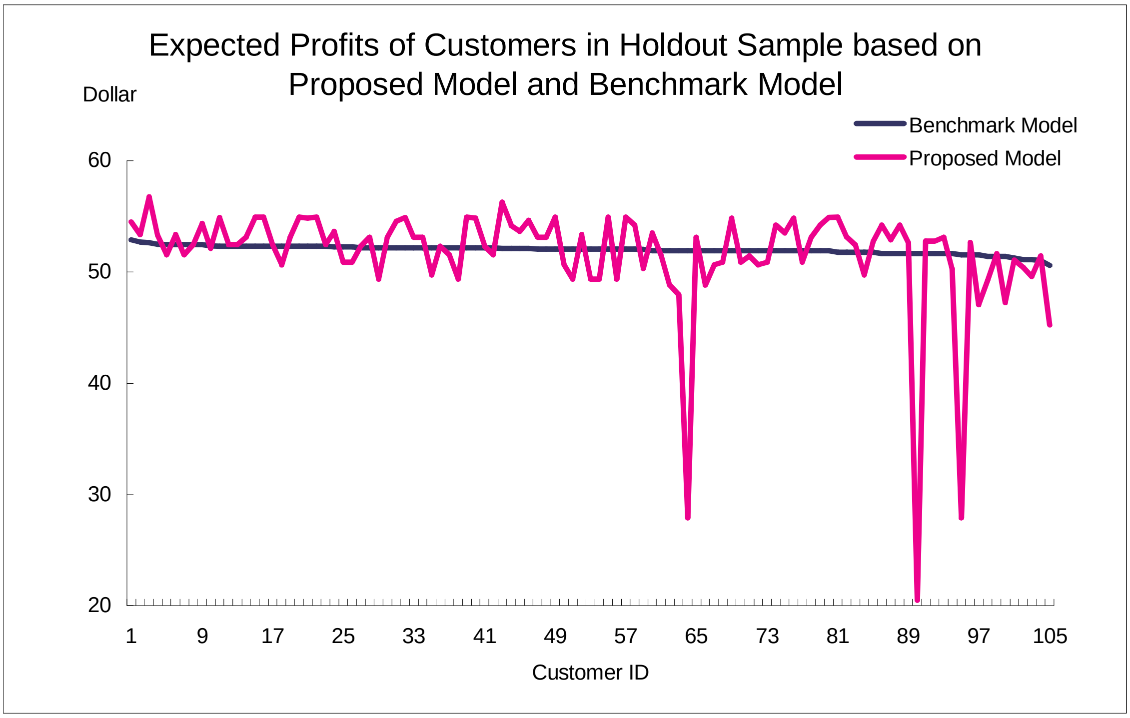

FIGURE 2(b)

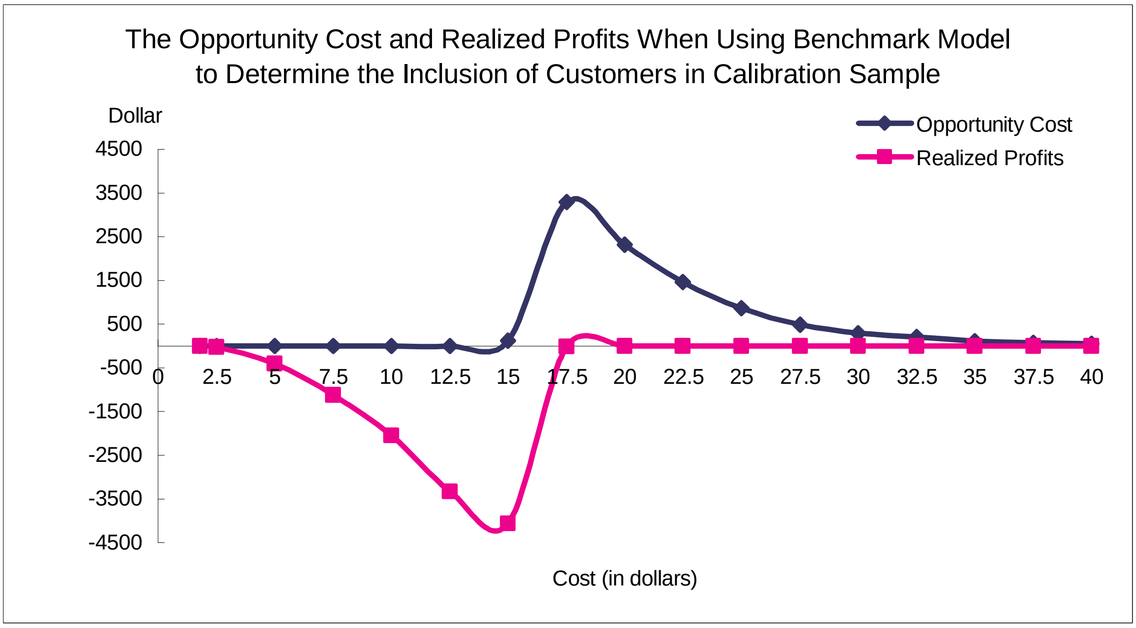

FIGURE 3(a)

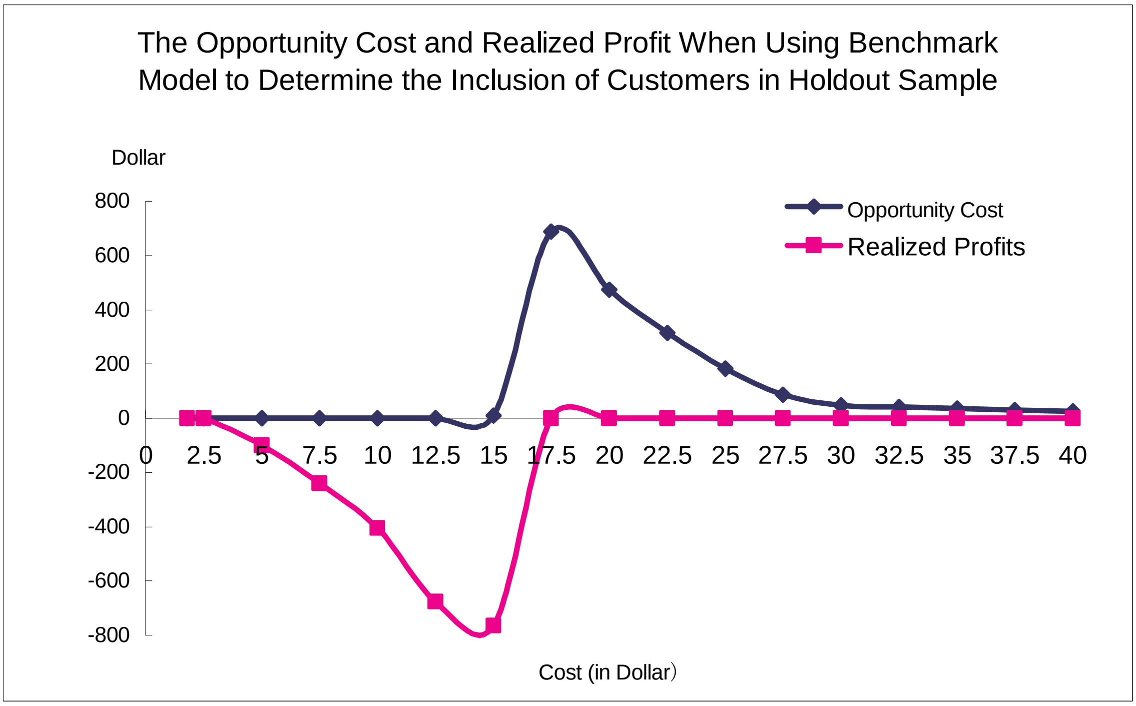

FIGURE 3(b)

APPENDIX A

Suppose that we want to compute customer lifetime profitability over a lifetime horizon of Y periods. Since we know that the expected profitability of customer n's next purchase immediately following a purchase from the firm (see equation 7 of the manuscript) is given by

$$\pi_n = P_n - c_n(E(\lambda_n)+1), \tag{A1}$$

the customer's expected lifetime profitability can be computed as the product of π_n and the number of purchases that the customer is expected to make within Y periods. This yields

$$LT\pi_{nj} = \pi_n \frac{Y}{E(\lambda_n)+1}, \tag{A2}$$

and plugging in for $E(\lambda_n)+1$ from equation (A1) yields

$$LT\pi_{nj} = \pi_n \frac{Yc}{P_n - \pi_n} = \frac{Yc_n}{\frac{P_n}{\pi_n} - 1}, \tag{A3}$$

which shows that the customer's lifetime value is a increasing function of the expected profitability of customer n's next purchase. This proves that π_n is a sufficient statistic for $LT\pi_{nj}$. [Q.E.D]

APPENDIX B

The derivation for equation (8) of the manuscript is given below.

$$\pi_{nj} = P_n - c_n \sum_{k=1}^{\infty} k_n \frac{\dfrac{e^{-E(\lambda_n)} E(\lambda_n)^{(j-j_o+k_n-1)}}{[j-j_o+k_n-1)]!}}{1-\sum_{u=1}^{j-j_o} \dfrac{e^{-E(\lambda_n)} E(\lambda_n)^{(u-1)}}{(u-1)!}}$$

$$= P_n - \frac{c_n}{1-\sum_{u=1}^{j-j_o} \dfrac{e^{-E(\lambda_n)} E(\lambda_n)^{(u-1)}}{(u-1)!}} \sum_{k_n=1}^{\infty} k_n \frac{e^{-E(\lambda_n)} E(\lambda_n)^{(j-j_o+k_n-1)}}{[j-j_o+k_n-1]!}$$

$$= P_n - \frac{c_n}{1-\sum_{u=1}^{j-j_o} \dfrac{e^{-E(\lambda_n)} E(\lambda_n)^{(u-1)}}{(u-1)!}} \left(\sum_{k_n=1}^{\infty} (j-j_0+k_n-1) \frac{e^{-E(\lambda_n)} E(\lambda_n)^{(j-j_o+k_n-1)}}{[j-j_o+k_n-1]!} - \sum_{k_n=1}^{\infty} (j-j_0-1) \frac{e^{-E(\lambda_n)} E(\lambda_n)^{(j-j_o+k_n-1)}}{[j-j_o+k_n-1]!} \right)$$

$$= P_n - \frac{c_n}{1-\sum_{u=1}^{j-j_o} \dfrac{e^{-E(\lambda_n)} E(\lambda_n)^{(u-1)}}{(u-1)!}} \left(\left(E(\lambda_n) - \sum_{u=1}^{j-j_o} \frac{e^{-E(\lambda_n)} E(\lambda_n)^{(u-1)}}{(u-1)!} \right) - (j-j_0-1) \left(1 - \sum_{u=1}^{j-j_o} \frac{e^{-E(\lambda_n)} E(\lambda_n)^{(u-1)}}{(u-1)!} \right) \right)$$

$$= P_n - \frac{c_n \left(E(\lambda_n) - \sum_{u=1}^{j-j_o} \dfrac{e^{-E(\lambda_n)} E(\lambda_n)^{(u-1)}}{(u-1)!} \right)}{1-\sum_{u=1}^{j-j_o} \dfrac{e^{-E(\lambda_n)} E(\lambda_n)^{(u-1)}}{(u-1)!}} + c_n (j-j_0-1),$$

We denote $y_n = \sum_{u=1}^{j-j_o} \frac{e^{-E(\lambda_n)} E(\lambda_n)^{(u-1)}}{(u-1)!}$. Thus, the above equation can be rewritten as below.

$$\pi_{nj} = P_n - \frac{c_n (E(\lambda_n) - y_n)}{1-y_n} + c_n (j-j_0-1) \qquad \text{[Q.E.D]}$$

APPENDIX C

The derivation for equation (15) of the manuscript is given below.

$$
\begin{aligned}
L_n &= \int_0^\infty \Pr(X_n > l_n \mid \lambda_n)\left[\prod_{t=2}^{T_n} \Pr(X_n = x_t \mid \lambda_n)\right]\Pr(X_n > r_n \mid \lambda_n) g_G(\lambda_n \mid \alpha_n, \beta_n) d\lambda_n \\
&= \int_0^\infty [1-\Pr(X_n < l_n \mid \lambda_n)][1-\Pr(X_n < r_n \mid \lambda_n)]\left(\frac{\beta_n^{\alpha_n}}{\Gamma(\alpha_n)(x_2!)\ldots(x_{T_n}!)}\lambda_n^{(\alpha_n+\sum_{t=2}^{T_n} x_t)-1} e^{-(\beta_n+T_n-1)\lambda_n}\right) d\lambda_n \\
&= \frac{\beta_n^{\alpha_n}\Gamma(\alpha_n+\sum_{t=2}^{T_n} x_t)}{\Gamma(\alpha_n)(x_2!)\ldots(x_{T_n}!)(\beta_n+T_n-1)^{(\alpha_n+\sum_{t=2}^{T_n} x_t)}} \\
&\quad * \int_0^\infty [1-\Pr(X_n < l_n \mid \lambda_n)][1-\Pr(X_n < r_n \mid \lambda_n)] g_G(\lambda_n \mid \alpha_n + \sum_{t=2}^{T_n} x_t, \beta_n + T_n - 1) d\lambda_n \\
&= \mathrm{K}(\Phi - \Psi)
\end{aligned}
\tag{C1}
$$

where,

$$
\mathrm{K} = \frac{\beta_n^{\alpha_n}\Gamma(\alpha_n+\sum_{t=2}^{T_n} x_t)}{\Gamma(\alpha_n)(x_2!)\ldots(x_{T_n}!)(\beta_n+T_n-1)^{(\alpha_n+\sum_{t=2}^{T_n} x_t)}} \tag{C2}
$$

$$
\begin{aligned}
\Phi &= \int_0^\infty [1-\Pr(X_n < r_n \mid \lambda_n)] g_G(\lambda_n \mid \alpha_n + \sum_{t=2}^{T_n} x_t, \beta_n + T_n - 1) d\lambda_n \\
&= \int_0^\infty \left(g_G(\lambda_n \mid \alpha_n + \sum_{t=2}^{T_n} x_t, \beta_n + T_n - 1) - g_G(\lambda_n \mid \alpha_n + \sum_{t=2}^{T_n} x_t, \beta_n + T_n - 1) \cdot \sum_{k=o}^{r_n-1} \frac{e^{-\lambda_n}\lambda^k}{k!} \right) d\lambda_n \\
&= \int_0^\infty \left(g_G(\lambda_n \mid \alpha_n + \sum_{t=2}^{T_n} x_t, \beta_n + T_n - 1) - \sum_{k=o}^{r_n-1} \left(\frac{(\beta_n+T_n-1)^{\alpha_n+\sum_{t=2}^{T_n} x_t}}{\Gamma\left(\alpha_n+\sum_{t=2}^{T_n} x_t\right)} \cdot \lambda_n^{\left(\alpha_n+\sum_{t=2}^{T_n} x_t\right)-1} e^{-(\beta_n+T_n-1)\lambda_n} \frac{e^{-\lambda_n}\lambda^k}{k!} \right) \right) d\lambda_n \\
&= \int_0^\infty \left(\begin{aligned} & g_G(\lambda_n \mid \alpha_n + \sum_{t=2}^{T_n} x_t, \beta_n + T_n - 1) \\ & - \sum_{k=o}^{r_n-1} \left(\frac{(\beta_n+T_n-1)^{\alpha_n+\sum_{t=2}^{T_n} x_t}}{\Gamma\left(\alpha_n+\sum_{t=2}^{T_n} x_t\right) k!} \cdot \frac{\Gamma\left(\alpha_n+\sum_{t=2}^{T_n} x_t + k\right)}{(\beta_n+T_n)^{\left(\alpha_n+\sum_{t=2}^{T_n} x_t + k\right)}} \cdot g_G(\lambda_n \mid \alpha_n + \sum_{t=2}^{T_n} x_t + k, \beta_n + T_n) \right) \end{aligned} \right) d\lambda_n
\end{aligned}
$$

$$=1-\sum_{k=o}^{r_n-1}\left(\frac{(\beta_n+T_n-1)^{\alpha_n+\sum_{t=2}^{T_n}x_t}}{\Gamma\left(\alpha_n+\sum_{t=2}^{T_n}x_t\right)k!}\cdot\frac{\Gamma\left(\alpha_n+\sum_{t=2}^{T_n}x_t+k\right)}{(\beta_n+T_n)^{\left(\alpha_n+\sum_{t=2}^{T_n}x_t+k\right)}}\right) \tag{C3}$$

$$\Psi=\int_0^\infty\left[\Pr(X_n<l_n\mid\lambda_n)-\Pr(X_n<l_n\mid\lambda_n)\cdot\Pr(X_n<r_n\mid\lambda_n)\right]g_G(\lambda_n\mid\alpha_n+\sum_{t=2}^{T_n}x_t,\beta_n+T_n-1)d\lambda_n$$

$$=\int_0^\infty\left(\begin{array}{l}\Pr(X_n<l_n\mid\lambda_n)\cdot g_G(\lambda_n\mid\alpha_n+\sum_{t=2}^{T_n}x_t,\beta_n+T_n-1)\\ -\Pr(X_n<l_n\mid\lambda_n)\cdot\Pr(X_n<r_n\mid\lambda_n)g_G(\lambda_n\mid\alpha_n+\sum_{t=2}^{T_n}x_t,\beta_n+T_n-1)\end{array}\right)d\lambda_n$$

$$=\int_0^\infty\left(\begin{array}{l}\sum_{j=0}^{l_n-1}\left(\frac{(\beta_n+T_n-1)^{\alpha_n+\sum_{t=2}^{T_n}x_t}}{\Gamma\left(\alpha_n+\sum_{t=2}^{T_n}x_t\right)}\cdot\lambda_n^{\left(\alpha_n+\sum_{t=2}^{T_n}x_t\right)-1}e^{-(\beta_n+T_n-1)\lambda_n}\frac{e^{-\lambda_n}\lambda^j}{j!}\right)\\ -\sum_{j=0}^{l_n-1}\frac{e^{-\lambda_n}\lambda_n{}^j}{j!}\cdot\sum_{k=0}^{r_n-1}\frac{e^{-\lambda_n}\lambda_n{}^k}{k!}\cdot\frac{(\beta_n+T_n-1)^{\alpha_n+\sum_{t=2}^{T_n}x_t}}{\Gamma\left(\alpha_n+\sum_{t=2}^{T_n}x_t\right)}\cdot\lambda_n^{\left(\alpha_n+\sum_{t=2}^{T_n}x_t\right)-1}e^{-(\beta_n+T_n-1)\lambda_n}\end{array}\right)d\lambda_n$$

$$=\int_0^\infty\left(\begin{array}{l}\sum_{j=0}^{l_n-1}\left(\frac{(\beta_n+T_n-1)^{\alpha_n+\sum_{t=2}^{T_n}x_t}}{\Gamma\left(\alpha_n+\sum_{t=2}^{T_n}x_t\right)j!}\cdot\frac{\Gamma\left(\alpha_n+\sum_{t=2}^{T_n}x_t+j\right)}{(\beta_n+T_n)^{\alpha_n+\sum_{t=2}^{T_n}x_t+j}}g_G(\lambda_n\mid\alpha_n+\sum_{t=2}^{T_n}x_t+j,\beta_n+T_n)\right)\\ -\sum_{j=0}^{l_n-1}\frac{e^{-\lambda_n}\lambda_n{}^j}{j!}\cdot\sum_{k=0}^{r_n-1}\left(\frac{(\beta_n+T_n-1)^{\alpha_n+\sum_{t=2}^{T_n}x_t}}{\Gamma\left(\alpha_n+\sum_{t=2}^{T_n}x_t\right)k!}\cdot\frac{\Gamma\left(\alpha_n+\sum_{t=2}^{T_n}x_t+k\right)}{(\beta_n+T_n)^{\left(\alpha_n+\sum_{t=2}^{T_n}x_t+k\right)}}\cdot g_G(\lambda_n\mid\alpha_n+\sum_{t=2}^{T_n}x_t+k,\beta_n+T_n)\right)\end{array}\right)d\lambda_n$$

$$
=\int_0^\infty \left(\begin{array}{l} \sum_{j=0}^{l_n-1}\left(\frac{(\beta_n+T_n-1)^{\alpha_n+\sum_{t=2}^{T_n}x_t}}{\Gamma\left(\alpha_n+\sum_{t=2}^{T_n}x_t\right)j!}\cdot\frac{\Gamma\left(\alpha_n+\sum_{t=2}^{T_n}x_t+j\right)}{(\beta_n+T_n)^{\alpha_n+\sum_{t=2}^{T_n}x_t+j}}g_G(\lambda_n\,|\,\alpha_n+\sum_{t=2}^{T_n}x_t+j,\beta_n+T_n)\right) \\ -\sum_{j=0}^{l_n-1}\sum_{k=0}^{r_n-1}\left(\frac{(\beta_n+T_n-1)^{\alpha_n+\sum_{t=2}^{T_n}x_t}}{\Gamma\left(\alpha_n+\sum_{t=2}^{T_n}x_t\right)j!k!}\cdot\frac{\Gamma\left(\alpha_n+\sum_{t=2}^{T_n}x_t+j+k\right)}{(\beta_n+T_n+1)^{\left(\alpha_n+\sum_{t=2}^{T_n}x_t+j+k\right)}}\cdot g_G(\lambda_n\,|\,\alpha_n+\sum_{t=2}^{T_n}x_t+j+k,\beta_n+T_n+1)\right) \end{array} \right) d\lambda_n
$$

$$
=\sum_{j=0}^{l_n-1}\left(\frac{(\beta_n+T_n-1)^{\alpha_n+\sum_{t=2}^{T_n}x_t}}{\Gamma\left(\alpha_n+\sum_{t=2}^{T_n}x_t\right)j!}\cdot\frac{\Gamma\left(\alpha_n+\sum_{t=2}^{T_n}x_t+j\right)}{(\beta_n+T_n)^{\alpha_n+\sum_{t=2}^{T_n}x_t+j}}\right)-\sum_{j=0}^{l_n-1}\sum_{k=0}^{r_n-1}\left(\frac{(\beta_n+T_n-1)^{\alpha_n+\sum_{t=2}^{T_n}x_t}}{\Gamma\left(\alpha_n+\sum_{t=2}^{T_n}x_t\right)j!k!}\cdot\frac{\Gamma\left(\alpha_n+\sum_{t=2}^{T_n}x_t+j+k\right)}{(\beta_n+T_n+1)^{\left(\alpha_n+\sum_{t=2}^{T_n}x_t+j+k\right)}}\right) \tag{C4}
$$

Combine (C1) – (C4), L_n can be rewritten as below:

$$
\begin{aligned}
L_n &= \mathrm{K}(\Phi-\Psi) \\
&= \mathrm{K}\Phi-\mathrm{K}\Psi \\
&= \frac{\Gamma\left(\alpha_n+\sum_{t=2}^{T_n}x_t\right)\beta_n^{\alpha_n}}{\Gamma(\alpha_n)(x_2!)\ldots(x_{T_n}!)}\cdot\left(\begin{array}{l}\frac{1}{(\beta_n+T_n-1)^{\alpha_n+\sum_{t=2}^{T_n}x_t}}-\frac{Y_R}{(\beta_n+T_n)^{\alpha_n+\sum_{t=2}^{T_n}x_t}} \\ -\frac{Y_L}{(\beta_n+T_n)^{\alpha_n+\sum_{t=2}^{T_n}x_t}}+\frac{Y_{RL}}{(\beta_n+T_n+1)^{\alpha_n+\sum_{t=2}^{T_n}x_t}}\end{array}\right)
\end{aligned} \tag{C5}
$$

where,

$$
\begin{aligned}
Y_L &= 1+\frac{\alpha_n+\sum_{t=2}^{T_n}x_t}{\beta_n+T_n}+\frac{\left(\alpha_n+\sum_{t=2}^{T_n}x_t\right)\left(\alpha_n+\sum_{t=2}^{T_n}x_t+1\right)}{(\beta_n+T_n)^2\,2!}+\ldots \\
&+\frac{\left(\alpha_n+\sum_{t=2}^{T_n}x_t\right)\left(\alpha_n+\sum_{t=2}^{T_n}x_t+1\right)\ldots\left(\alpha_n+\sum_{t=2}^{T_n}x_t\right)\left(\alpha_n+\sum_{t=2}^{T_n}x_t+l_n-2\right)}{(\beta_n+T_n)^{(l_n-1)}\,(l_n-1)!}
\end{aligned} \tag{C6}
$$

and

$$Y_R = 1 + \frac{\alpha_n + \sum_{t=2}^{T_n} x_t}{\beta_n + T_n} + \frac{\left(\alpha_n + \sum_{t=2}^{T_n} x_t\right)\left(\alpha_n + \sum_{t=2}^{T_n} x_t + 1\right)}{\left(\beta_n + T_n\right)^2 2!} + \ldots$$
$$+ \frac{\left(\alpha_n + \sum_{t=2}^{T_n} x_t\right)\left(\alpha_n + \sum_{t=2}^{T_n} x_t + 1\right)\ldots\left(\alpha_n + \sum_{t=2}^{T_n} x_t\right)\left(\alpha_n + \sum_{t=2}^{T_n} x_t + r_n - 2\right)}{\left(\beta_n + T_n\right)^{(r_n - 1)}\left(r_n - 1\right)!} \quad \text{(C7)}$$

and

$$Y_{RL} = \sum_{j=0}^{l_n - 1} \sum_{k=0}^{r_n - 1} \left[\frac{\left(\alpha_n + \sum_{t=2}^{T_n} x_t\right)\left(\alpha_n + \sum_{t=2}^{T_n} x_t + 1\right)\ldots\left(\alpha_n + \sum_{t=2}^{T_n} x_t + j + k - 1\right)}{\left(\beta_n + T_n + 1\right)^{j+k} j!k!} \right] \quad \text{(C8)}$$

Since for $\sum_{t=1}^{T_n} x_t \geq 1, \Gamma\left(\alpha_n + \sum_{t=2}^{T_n} x_t\right) = \alpha_n\left(\alpha_n + 1\right)\ldots\left(\alpha_n + \sum_{t=2}^{T_n} x_t - 1\right)\Gamma\left(\alpha_n\right)$,

and for $\sum_{t=2}^{T_n} x_t \geq 0, \Gamma\left(\alpha_n + \sum_{t=2}^{T_n} x_t\right) = \Gamma\left(\alpha_n\right)$, we can further simplify L_n in (C5) as follows:

$$L_n = W_n \left[\frac{1}{\left(\beta_n + T_n - 1\right)^{\alpha_n + \sum_{t=2}^{T_n} x_t}} - \frac{Y_R}{\left(\beta_n + T_n\right)^{\alpha_n + \sum_{t=2}^{T_n} x_t}} - \frac{Y_L}{\left(\beta_n + T_n\right)^{\alpha_n + \sum_{t=2}^{T_n} x_t}} + \frac{Y_{RL}}{\left(\beta_n + T_n + 1\right)^{\alpha_n + \sum_{t=2}^{T_n} x_t}} \right],$$

where

$$W_n = \frac{\left(\alpha_n\left(\alpha_n + 1\right)\ldots\left(\alpha_n + \sum_{t=2}^{T_n} x_t - 1\right)\right)\beta_n^{\alpha_n}}{\left(x_2 !\right)\ldots\left(x_{T_n} !\right)}. \quad \text{[Q.E.D]}$$

APPENDIX D

We derive equation (25) of the manuscript here. Let us assume that we are computing the expected profitability of the customer's next purchase over a finite number of mailings M_n.

$$\pi_{nj}^{M_n} = \sum_{x=1}^{M_n} [P_n - xc_n](1-p_n)^{x-1} p_n$$
$$= (P_n - c_n)p_n + (P_n - 2c_n)p_n(1-p_n) + (P_n - 3c_n)p_n(1-p_n)^2 + \ldots + (P_n - M_n c)p_n(1-p_n)^{M_n - 1} \tag{D1}$$

$$(1-p_n)\pi_{nj}^{M_n} = (P_n - c_n)p_n(1-p_n) + (P_n - 2c_n)p_n(1-p_n)^2 + (P_n - 3c_n)p_n(1-p_n)^3$$
$$+ \ldots + (P_n - (M_n - 1)c_n)p_n(1-p_n)^{M_n - 1} + (P_n - M_n c_n)p_n(1-p_n)^{M_n} \tag{D2}$$

Subtracting (D2) from (D1), we obtain

$$p_n \pi_{nj}^{M_n} = (P_n - c_n)p_n + (-c_n)p_n(1-p_n) + (-c_n)p_n(1-p_n)^2 + \ldots + (-c_n)p_n(1-p_n)^{M_n - 1} - (P_n - M_n c_n)p_n(1-p_n)^{M_n}$$
$$= (P_n - c_n)p_n + (-c_n)p_n \cdot \frac{(1-p_n) + (1-p_n)^{M_n - 1}}{1-(1-p_n)} - (P_n - M_n c_n)p_n(1-p_n)^{M_n}$$
$$= (P_n - c_n)p_n + (-c_n)(1-p_n) + (-c_n)(1-p_n)^{M_n - 1} - (P_n - M_n c_n)p_n(1-p_n)^{M_n}$$

$$\Rightarrow \pi_{nj}^{M_n} = \frac{P_n \cdot p_n - c_n}{p_n} + \frac{(-c_n)(1-p_n)^{M_n - 1}}{p_n} - (P_n - M_n c_n)(1-p_n)^{M_n} \tag{D3}$$

When $M_n \to \infty$, (D3) becomes:

$$\pi_{nj} = \sum_{x=1}^{\infty} [P_n - xc_n](1-p_n)^{x-1} p_n = \frac{P_n \cdot p_n - c_n}{p_n} \quad \text{[Q.E.D]}$$

[Q.E.D]

APPENDIX E

The derivation of equation (31) is given as follows:

$$
\begin{aligned}
L_n &= \int_0^1 p_n^{T_n} \left(1-p_n\right)^{l_n+M_{T_n-1}+r_n-T_n} g_B\left(p_n \mid a_n, b_n\right) dp_n \\
&= \int_0^1 p_n^{T_n} \left(1-p_n\right)^{l_n+M_{T_n-1}+r_n-T_n} \frac{\Gamma\left(a_n+b_n\right)}{\Gamma\left(a_n\right)\Gamma\left(b_n\right)} p_n^{a_n-1} \left(1-p_n\right)^{b_n-1} dp_n \\
&= \frac{\Gamma\left(a_n+b_n\right)}{\Gamma\left(a_n\right)\Gamma\left(b_n\right)} \int_0^1 \left(p_n^{a_n+T_n-1} \left(1-p_n\right)^{b_n+l_n+M_{T_n-1}+r_n-T_n-1} dp_n \right) \\
&= \frac{\Gamma\left(a_n+b_n\right)}{\Gamma\left(a_n\right)\Gamma\left(b_n\right)} \int_0^1 \left(\begin{array}{l} \dfrac{\Gamma\left(a_n+b_n+l_n+M_{T_n-1}+r_n\right)}{\Gamma\left(a_n+T_n\right)\Gamma\left(b_n+l_n+M_{T_n-1}+r_n-T_n\right)} \\ \cdot p_n^{a_n+T_n-1} \left(1-p_n\right)^{b_n+l_n+M_{T_n-1}+r_n-T_n-1} dp_n \dfrac{\Gamma\left(a_n+T_n\right)\Gamma\left(b_n+l_n+M_{T_n-1}+r_n-T_n\right)}{\Gamma\left(a_n+b_n+l_n+M_{T_n-1}+r_n\right)} \end{array} \right) \\
&= \frac{\Gamma\left(a_n+b_n\right)}{\Gamma\left(a_n\right)\Gamma\left(b_n\right)} \frac{\Gamma\left(a_n+T_n\right)\Gamma\left(b_n+l_n+M_{T_n-1}+r_n-T_n\right)}{\Gamma\left(a_n+b_n+l_n+M_{T_n-1}+r_n\right)}
\end{aligned}
$$

[Q.E.D]

REFERENCES

Allenby, G. M., Leone, R. P., Jen, L. 2005. "A Dynamic Model of Purchase Timing with Application to Direct Marketing," *Journal of the American Statistical Association*, 94, 446, 365-374.

Boatwright, P., Borle, S., Kadane, J. B. 2003. "A Model of the Joint Distribution of Purchase Quantity and Timing," *Journal of the American Statistical Association*, 98, 463, 564-572.

Bodapati, A., Gupta, S. 2004. "A Direct Approach to Predicting Discretized Response in Target Marketing," *Journal of Marketing Research*, 41, 1, 73-85.

Borle, S., Boatwright, P., Kadane, J. B., Nunes, J. C., Schmueli, G. 2005. "The Effect of Product Assortment Changes on Customer Retention," *Marketing Science*, 24, 4, 616-622.

Buchanan, B. S., Morrison, D. G. 1988. "Stochastic Modeling of List Falloffs with Implications for Repeat Mailings," *Journal of Direct Marketing*, 2, 3, 7-14.

Fader, P. S., Hardie, B. G. S., Lee, K. L. 2005a. "RFM and CLV: Using Iso-Value Curves for Customer Base Analysis," *Journal of Marketing Research*, 42, 4, 415-430.

Fader, P. S., Hardie, B. G. S., Lee, K. L. 2005b. "Counting Your Customers the Easy Way: An Alternative to the Pareto/NBD Model," *Marketing Science*, 24, 2, 275-284.

Gonul, F. F., Kim, B. D., Shi, M. 2000. "Mailing Smarter to Catalog Customers," *Journal of Interactive Marketing*, 9, 2, 20-31.

Gonul, F. F., Hofstede, F. T. 2006. "How to Compute Optimal Catalog Mailing Decisions," *Marketing Science*, 25, 1, 65-74.

Gupta, S., Lehmann, D. R., Stuart, J. A. 2004. "Valuing Customers," *Journal of Marketing Research*, 41, 1, 7-18.

Gupta, S., Morrison, D. G. 1991. "Estimating Heterogeneity in Consumers' Purchase Rates," *Marketing Science*, 10, 3, 264-269.

Iyengar, R., Ansari, A., Gupta, S. 2003. "Leveraging Information Across Categories," *Quantitative Marketing and Economics*, 1, 4, 425-465.

Li, S., Sun, B., Wilcox, R. T. 2005. "Cross-Selling Sequentially Ordered Products: An Application to Consumer Banking Services," *Journal of Marketing Research*, 42, 2, 233-239.

Morrison, D. G., Schmittlein, D. C. 1981. "Predicting Future Random Events Based on Past Performance," *Management Science*, 27, 9, 1006-1023.

Rao, V. R., Steckel, J. H. 1995. "Selecting, Evaluating and Updating Prospects in Direct Mail Marketing," *Journal of Direct Marketing*, 9, 2, 20-31.

Reichcheld, F. F. 1996. *The Loyalty Effect: The Hidden Force Behind Growth, Profits and Lasting Value*, Boston: Harvard Business School Press.

Rossi, P. E., McCulloch, R. E., Allenby, G. M. 1996. "On the Value of Household Purchase History Information in Target Marketing," *Marketing Science*, 15, 4, 321-340.

Rust, R. T., Lemon, K. N., Zeithaml, V. A. 2004. "Return on Marketing: Using Customer Equity to Focus Marketing Strategy," *Journal of Marketing Research*, 41, 1, 31-45.

Seetharaman, P. B., Chintagunta, P. K. 2003. "The Proportional Hazard Model for Purchase Timing: A Comparison of Alternative Specifications," *Journal of Business and Economic Statistics*, 21, 3, 368-382.

Ter Hofstede, F., Wedel, M. 1998. "A Monte Carlo Study of Time Aggregation in Continuous Time and Discrete Time Parametric Hazard Models," *Economics Letters*, 58, 2, 149-156.

Thomas, J. S., Blattberg, R. C., Fox, E. J. 2004. "Recapturing Lost Customers," *Journal of Marketing Research*, 41, 1, 31-45.

Venkatesan, R., Kumar, V. 2004. "Customer Lifetime Value Framework for Customer Selection and Resource Allocation Strategy," *Journal of Marketing*, 68, 4, 106-125.