Incorporating Nonrandom Direct Marketing Activity

into Latent Attrition Models

David A. Schweidel*

George Knox

July 2011

* David A. Schweidel is assistant professor of marketing at the University of Wisconsin-Madison School of Business. George Knox is assistant professor of marketing at Tilburg University. Address all correspondence on this manuscript to David Schweidel (dschweidel@bus.wisc.edu), 4191B Grainger Hall, 975 University Ave., Madison, WI, 53706.
Incorporating Nonrandom Direct Marketing Activity into Latent Attrition Models

Abstract

Latent attrition models serve as the foundation for analyzing transactional customer relationships to derive forecasts of customer behavior and value. Yet, extant models often overlook the effects of direct marketing and how it is individually targeted. We propose a parsimonious model that allows for direct marketing to impact the frequency with which transactions occur, the monetary amount of a transaction, and the latent customer lifetime. Additionally, our modeling framework takes into account the way in which organizations may target direct marketing at the individual level.

Using donation data from a non-profit organization, we find that direct marketing increases the donation rate and the amount donated for active donors. However, direct marketing is a double-edged sword for some donors because it accelerates their attrition. We use our model to examine the financial impact of direct marketing activities, deriving the residual lifetime value (RLV) associated with alternative direct marketing policies and the incremental RLV as a measure of the effectiveness of an additional piece of direct marketing. We examine how these metrics relate to summaries of past donation activity (recency, frequency, and average donation amount) and discuss potential shortcomings of basing direct marketing decisions on such heuristics.

Keywords: Latent attrition, Customer relationship management, Simultaneity
1. Introduction

Latent attrition models (i.e., “buy till you die” models) lie at the heart of transactional (i.e., noncontractual) customer base analysis, because attrition is not observed and hence must be inferred (Kumar and Reinartz 2006, p.103). These models can be used to forecast future purchasing for both new and existing customers (e.g., Schmittlein, Morrison and Colombo 1987; Fader, Hardie and Shang 2010) and aid managers in areas such as customer valuation, customer targeting, and resource allocation (e.g., Reinartz and Kumar 2000; 2003).

Whereas early models ignored covariates entirely (e.g., Schmittlein, Morrison and Colombo 1987; Fader, Hardie and Lee 2005a), more recent research has incorporated time-invariant predictors such as demographics (e.g., Singh, Borle and Jain 2009; Abe 2009; Neslin, Rhoads and Wolfson 2009). These covariates may impact both the purchasing rate for active customers and the rate at which these customers become inactive. Though these models represent an important advance, the next step is to consider the effects of time-varying marketing efforts, a necessary tool for firms seeking to manage their relationships with existing customers. While it may seem a straightforward extension, several issues arise when incorporating time-varying marketing activity.

Under the latent attrition framework, there are three distinct ways in which direct marketing activity may impact customer behavior. First, direct marketing may increase the transaction rate of active customers. Second, it may increase the monetary amount of a given transaction. Third, it may affect the rate at which active customers become inactive. Once a customer transitions to the inactive (“death”) state, it is irreversible: he
or she is “lost for good.” Direct marketing may benefit the firm if it reduces attrition and consequently extends a customer’s latent lifetime. However, it may adversely affect the firm if such efforts irritate customers (e.g., Fournier, Dobscha and Mick 1997; Venkatesan and Kumar 2004; van Diepen, Donkers and Franses 2009) and lead them to become inactive more quickly.

By quantifying the impact of direct marketing on these three aspects of behavior, we offer a more comprehensive view of its effect than previously considered. In doing so, our model also distinguishes between marketing effects experienced in the short-term that impact the transaction rate and size for active customers, and the long-term which affect the latent lifetime (e.g., Montoya, Netzer and Jedidi 2010). The overall impact of direct marketing will therefore hinge on the balance of short-term and long-term effects. For example, if marketing efforts increase the donation rate while active and extend the latent lifetime, this will result in an increase in the number of donations. Alternatively, if the impact on the latent lifetime is negative, this can result in more donations over a shorter lifetime, or, donation acceleration.

We also make two methodological contributions to this stream of research. First, we account for the nonrandom nature of direct marketing. This arises because marketers target customers with direct marketing activity (e.g., Manchanda, Rossi and Chintagunta 2004). Though the targeting decision may be partially driven by observable summaries of customer behavior such as the recency and frequency of transactions, it may also be influenced by factors unobservable to the researcher, like expert judgment. We therefore model both customers’ transactional behavior conditional on the direct marketing received, and the firm’s direct marketing activity conditional on customers’ past actions.
Following Donkers et al. (2006), we allow for the possibility that there are unobservable factors that affect both the customer’s transactional behavior and the firm’s marketing efforts toward that customer.

Second, we empirically assess the extent to which the inter-transaction times and transaction amounts are related. While prior latent attrition models have assumed that these components are independent (e.g., Schmittlein and Peterson 1994; Fader, Hardie and Lee 2005b) or conditionally independent (e.g., Borle, Singh and Jain 2008), we consider two distinct ways in which they may be related. Like Borle, Singh and Jain (2008), we account for associations among underlying parameters governing the inter-transaction times and transaction amounts across customers. For example, customers who make frequent purchases may also spend more. We also model correlation at the individual level between the inter-transaction time process and the transaction amount process (e.g., Jen, Chou and Allenby 2009). For example, after a longer than expected inter-transaction time, a customer may also spend more than average. Hence, we account for correlated transaction timing and spending within and across customers.

In sum, we propose a model that allows direct marketing to influence customers through three distinct mechanisms; we allow for associations between transaction timing and spending; and finally we account for the nonrandom targeting of direct marketing. Our model allows us to quantify the revenue impact of many alternative direct marketing strategies, which the firm can weigh against the associated costs. While we develop our modeling framework at the level of the individual customer, decisions regarding direct marketing plans can be evaluated at a higher level of aggregation, such as at the segment level for firms using RFM segmentations. This flexibility may be valuable to
organizations lacking the resources to develop separate direct marketing plans for each customer.

The remainder of this paper proceeds as follows. We next describe the data used in our analysis. We then develop our modeling framework and present the empirical findings. Based on our results, we consider a broad range of possible direct marketing strategies that vary in terms of both volume and timing. We assess their financial impact by calculating an individual’s residual lifetime value (RLV) under different direct marketing policies. We further investigate how RLV is related to commonly-used heuristics such as RFM variables and characteristics of the mailing strategies such as pulsing vs. blitz strategies (e.g., Mahajan and Muller 1986). We also derive another measure of financial impact, the incremental RLV associated with sending one more mailing, and show how it varies across RFM segments. We conclude with a discussion of the limitations of the current research and avenues for future work.

2. Data

Our data come from a non-profit organization in the United States and were provided by the Direct Marketing Educational Foundation (DMEF) for its Lifetime Value and Customer Equity Modeling Competition in 2008.¹ The dataset consists of 21,166 donors who made their first donation to the organization during the first half of 2002. It also contains the donation and mailing history with each donor through the end of August 2006. As the organization sends its direct marking at the start of each month, we analyze

¹ As our data are from a non-profit organization, in detailing our model and discussing the results, we use the terms “donor” and “organization” in place of “customer” and “firm,” respectively.
donation behavior on a monthly basis to coincide with the organization’s direct marketing activity.²

We randomly selected 20% of the donors from the dataset, resulting in a sample of 4,234 donors. We used the first 40 months of data (January 2002 – April 2005) to calibrate the model and use the remaining 16 months (May 2005 – August 2006) to validate it. Subsequent to a donor’s initial contribution, we track each donor’s repeat donation behavior. During the 56 month observation period, the average number of repeat donations was 1.54; 50% of donors did not make any repeat donations, 17% made a single repeat donation, 10% made two repeat donations, and 23% made more than two repeat donations. The average number of repeat donations made by donors who make at least one repeat donation was 3.07; this suggests that the distribution is highly skewed. The average repeat donation amount across all individuals and repeat donation occasions was $38.69. Among donors who made a single repeat donation, the average was $36.01, while the average donation amount among those making multiple repeat donations was slightly higher at $39.03.

To further understand donor behavior over the course of our observation period, we investigate donation behavior and direct marketing activity. First, we examine the total number of repeat donations and total dollar amount over months in Figure 1.

² While we analyze the data at the monthly level, our analysis could be applied at a more granular (i.e., daily or weekly) or coarse (i.e., quarterly or annually) level.
Figure 1 shows that the number of donations and the total donation amounts track closely.

Two other aspects of donation behavior become evident from Figure 1. First, donation behavior peaks in the months of November, December, and January throughout the observation period. Second, it appears that the number of donations and the total donation amounts diminish as we move further into the observation period. This can be seen most clearly by examining the peaks in donation activity during the winter months, comparing the last two peaks to the first two peaks.

One explanation for the decline is latent attrition. This is supported by the observation that time since last donation increases over the observation period. For example, as of January 2003, the average number of months that had elapsed since a donor’s last donation was 8.68 months. This increased to 15.98 months by January 2004 and 23.77 months by January 2005.
Next, we examine mailing activity during the 40-month calibration period. We calculate the total number of mailings received during the calibration period, the number of runs (continuous sequences of mailings or non-mailings), and the largest number of consecutive months over which donors received mailings. Summary statistics of these measures are presented in Table 1.

<table>
<thead>
<tr>
<th></th>
<th>Mean (s.d.)</th>
<th>95% Interval</th>
</tr>
</thead>
<tbody>
<tr>
<td>Number of mailings</td>
<td>20.66 (4.28)</td>
<td>(12,28)</td>
</tr>
<tr>
<td>Number of runs</td>
<td>22.91 (3.59)</td>
<td>(13,29)</td>
</tr>
<tr>
<td>Longest stretch of consecutive months with mailings</td>
<td>5.26 (1.46)</td>
<td>(4,9)</td>
</tr>
</tbody>
</table>

Table 1. Summary Statistics of Mailings

While donors on average receive mailings approximately every second month, there is considerable variation. Based on the number of runs observed and the longest stretch of consecutive months over which mailings are received, we also observe variation in the temporal pattern with which donors receive mailings.

To examine patterns in mailing activity at the aggregate level, Figure 2 shows the sample proportion receiving direct marketing each month subsequent to their first donation to the organization. Like donation activity, direct marketing declines over time, consistent with an organization curbing its efforts in response to donors who have likely become inactive.
The most noticeable shift in mailing activity occurs at the beginning of 2005. With the exception of a single peak in mailing activity that occurs in July 2005, less than half of the donors in our sample receive mailings each month during 2005 and 2006. To illustrate this change, consider mailings in January 2005, which were sent to 42% of donors in our sample. This is in sharp contrast to mailings in January 2004, which were sent to 98% of donors in our sample. Among those who donated recently, within one year of January 2005 (21% of our sample), 93% of those donors received a mailing, while only 28% of donors who did not make a contribution within the last year received a mailing.

Hence, evidence suggests that mailing decisions are responsive to donors’ past activity. To account for this, we simultaneously model an individual’s donation activity conditional on the organization’s direct marketing efforts and the organization’s direct...
marketing conditional on the individual’s past actions. We detail our modeling framework in the next section.

3. Model Development

We begin by developing a joint model of donation incidence and amount that governs donor behavior while active, conditional on direct marketing. Next, we account for latent attrition, and the firm’s mailing decisions, which we associate with donor behavior through a latent class approach (e.g., Donkers et al. 2006).

Donation Incidence and Amount

As in previous models for customer base analyses, we assume that there are two distinct states that govern donor behavior: an active (“alive”) state and an inactive (“death”) state. In the active state, after coming under observation by the organization, we assume that donor $i$’s decision to make a donation at time $t$ follows a Bernoulli process with a probability of $p_{it}$. If donor $i$’s $j$th donation occurred at time $t_{ij}$, the cumulative distribution for the time until donor $i$’s next donation would be given by:

$$F(t \mid t_{i,j}) = 1 - \prod_{k=t_{i,j}+1}^{t_{i,j}+t} (1 - p_{ik})$$  \hspace{1cm} (1)

We can then derive the probability of a donation occurring at $t_{i,j}+t$, conditional on the donor not having made a contribution by $t_{i,j}+t-1$, as:

$$f(t \mid t_{i,j}) = \frac{F(t \mid t_{i,j}) - F(t-1 \mid t_{i,j})}{1 - F(t-1 \mid t_{i,j})}$$  \hspace{1cm} (2)

If $p_{it}=p_i$ for all $t$, the monthly donation process follows a Bernoulli distribution with probability $p_i$, and the inter-donation times would follow a geometric distribution, the
individual level inter-donation time model assumed by Fader, Hardie and Shang (2010). To allow for $p_{it}$ to vary over time based on the seasonality (i.e., increased donation frequency in the winter months) and the organization’s direct marketing efforts, we specify $p_{it}$ as:

$$p_{it} = \Phi(\gamma_0 + \gamma_1 \cdot DM_{it} + \gamma_2 \cdot Winter_i)$$  

(3)

where $\Phi(x)$ is the c.d.f of the standard normal distribution at $x$; $Winter_i$ is an indicator variable equal to 1 in November, December and January, and is equal to 0 otherwise; and $DM_{it}$ is a stock variable for the amount of direct marketing activity that donor $i$ has received by month $t$ (e.g., Van Diepen, Donkers and Franses 2009). To allow for carryover effects of direct marketing from one month to the next, we make the standard assumption that $DM_{it}$ is specified as:

$$DM_{it} = Direct_{it} + \lambda \cdot DM_{i,t-1}$$  

(4)

where $Direct_{it}$ is an indicator variable equal to 1 if donor $i$ received a mailing at the beginning of month $t$ and 0 otherwise, and $\lambda$ represents the rate of decay, with $\lambda \in (0,1)$.

The coefficient $\gamma_1$ allows for the impact of direct marketing activity on the likelihood of a donation in a given month, with values $\gamma_1 > 0$ suggesting that direct marketing activity increases the likelihood with which donors make a donation in a given month and consequently reduces the expected time until the next donation. The coefficient $\gamma_2$ accounts for variation in donation incidence related to the time of year. Based on Figure 1, we anticipate $\gamma_2 > 0$.

Conditional on donor $i$ making a donation in month $t$, we assume that the donation amount $A_{it}$ follows a log-normal distribution with parameters $\mu_{it}$ and $\sigma$: 


\[ \mu_a = \beta_0 + \beta_1 \cdot DM_a + \beta_2 \cdot Winter, \]
\[ \ln(A_i) \sim N\left(\mu_a, \sigma^2\right) \tag{5} \]

We denote the corresponding c.d.f. of the amount model as \( G(a) \) and the p.d.f. as \( g(a) \).

We correlate the marginal geometric distribution of inter-donation times and the log-normal distribution of donation amounts using a bivariate Gaussian copula (Danaher and Smith 2011). Let \( t^* = \Phi^{-1}(F(t|t_{ij})) \) and \( a^* = \Phi^{-1}(G(A_{ii})) \), where \( \Phi^{-1}(x) \) is the probit function. We assume that \((t^*, a^*)\) follow a standard bivariate normal distribution with correlation \( \rho \). If \((X, Y)\) are distributed according to a standard bivariate normal distribution with correlation \( \rho \), the conditional distribution of \( X \) given \( Y = y \) is normally distributed with mean \( \rho \cdot y \) and variance \( 1 - \rho^2 \). Conditional on the value \( A_{ii} = a \), \( t^* \) is normally distributed as follows:

\[ t^* \mid a \sim N\left(\rho \cdot \Phi^{-1}(G(a)), 1 - \rho^2\right) \tag{6} \]

Substituting the c.d.f of the log-normal distribution for \( G(a) \) and simplifying, the conditional distribution of \( t^* \) can be written as:

\[ t^* \mid a \sim N\left(\rho \cdot \left(\frac{\ln(a) - \mu_a}{\sigma}\right), 1 - \rho^2\right) \tag{7} \]

The cumulative distribution associated with the conditional distribution of \( F(t|t_{ij}, A_{ii}) \) can then be expressed as:

\[ F(t \mid t_{ij}, A_{ii}) = \Phi\left(\frac{\Phi^{-1}(F(t \mid t_{ij})) - \rho \cdot \left(\frac{\ln(A_{ii}) - \mu_a}{\sigma}\right)}{\sqrt{1 - \rho^2}}\right) \tag{8} \]

When \( \rho = 0 \) and the inter-donation timing and donation amount processes are independent, \( F(t|t_{ij}, A_{ii}) \) in equation (8) is given by \( F(t|t_{ij}) \). Letting \( h(t, a) \) denote the joint likelihood of a
donation occurring \( t \) months since the last donation (at time \( t_{ij} \)) and the donation being for an amount \( a \), conditional on not having observed a donation in \( t-1 \) months, the likelihood is given by:

\[
h(t, a | t_{ij}) = \begin{cases} 
  g(a) \times \left( \frac{F(t | t_{ij}, A_u) - F(t-1 | t_{ij}, A_u)}{1 - F(t-1 | t_{ij})} \right), & a > 0 \\
  \frac{1 - F(t | t_{ij})}{1 - F(t-1 | t_{ij})}, & a = 0 
\end{cases}
\]

(9)

If the donation timing and amount processes are independent \((\rho = 0)\), when a donation occurs, the joint probability \( h(t, a | t_{ij}) \) results from the product of the probability associated with the inter-donation time \( t \) (equation (2)) and the probability associated with the amount donated \( a \), given by \( g(a) \).

**Incorporating Latent Attrition in Donation Behavior**

The joint model outlined in equations (1) – (9) considers the link between donation incidence and amount, but it does not allow for the possibility that donors may become inactive. To allow for the possibility that donor \( i \) becomes inactive after the \( t \)th month of observation, we assume that the likelihood with which a donor remains active is given by \( q_{it} \), such that:

\[
q_{it} = \Phi(\alpha_0 + \alpha_1 \cdot DM_{it})
\]

(10)

If \( q_{it} = q_i \), resulting from \( \alpha_1 = 0 \), there is no effect of direct marketing on attrition, and a donor’s latent lifetime follows a geometric distribution, as in Fader, Hardie and Shang (2010).
Let $A_i$ be a vector of length $T$ that contains the amount donated in each month of the observation period by donor $i$, and let $t_i$ denote the last month in which a repeat donation was observed. We denote $f_{iz}$ as the likelihood of donation activity associated with the $z^{th}$ month from which donor $i$ came under observation (i.e., $f_{iz} = h(z - t_i^*, a | t_i^*)$ where $t_i^*$ is the time of the most recent donation prior to $z$). The likelihood function associated with donor $i$’s activities is given by:

$$L(A_i, Direct_i, Winter_i, \alpha, \beta, \gamma, \lambda, \sigma) = \prod_{z=1}^{T} (q_{iz} \cdot f_{iz}) + \sum_{z=t_i}^{T-1} \left[ \prod_{k=1}^{z} (q_{i(k)} \cdot f_{i(k)})(1 - q_{i(z+1)}) \right]$$

(11)

The first term accounts for the sequence of donation decisions when a donor remains active throughout the observation period. The summation allows for the possibility that a donor may become inactive in any month after the last observed donation, and accounts for the sequence of donation decisions observed until the donor becomes inactive. If the donor was observed to make a donation in the final month of the observation period, he must be active at time $T$ and hence could not have become inactive.

**Accounting for the Organization’s Mailing Process and Heterogeneity**

In addition to modeling the individual’s donation behavior, we also must account for the organization’s targeting of direct marketing activity. Consistent with the industry and extant direct marketing literature (e.g., Gönül and Shi 1998), we model the organization’s decision to send marketing to a donor in a given month as a function of RFM characteristics. We operationalize recency as the number of months that have elapsed since the last donation, frequency as the number of donations made to date, and the monetary amount as the logarithm of the average amount that has been donated. In
addition to these RFM measures, we also include a seasonal variable, as the organization’s mailing efforts may be higher during the winter months, as donation activity tends to be higher at that time of year. We assume that the decision to send donor $i$ a mailing in month $z$ ($\text{Direct}_{iz}$) follows a Bernoulli process, with the probability of sending a mailing given by $m_{iz}$:

$$m_{iz} = \Phi(\kappa_0 + \kappa_1 \cdot \text{Recency}_{iz} + \kappa_2 \cdot \text{Frequency}_{iz} + \kappa_3 \cdot \text{Monetary}_{iz} + \kappa_4 \cdot \text{Winter}_{iz})$$ (12)

The values for the Recency, Frequency, and Monetary at each point in time are calculated based on the donation behavior observed until that point in time. The likelihood associated with the sequence of mailings that donor $i$ receives can then be expressed as:

$$M(\text{Direct}_i \mid \kappa, \text{Winter}, \text{Recency}, \text{Frequency}, \text{Monetary}) = \prod_{z=1}^{T} \left( m_{iz} \right)^{\text{Direct}_i} \left( 1 - m_{iz} \right)^{1-\text{Direct}_i}$$ (13)

While RFM variables are expected to reflect the organization’s mailing activities, we do not know the exact rules of the organization’s mailing decisions. Moreover, expert judgment or other criteria may supersede pre-specified rules (e.g., Donkers et al. 2006). To allow for unobserved characteristics of donors to impact both the donation process at the individual-level and the organization’s marketing efforts toward a given individual, we employ a latent class framework in which both the donation process and the organization’s mailing decisions are class-specific (Donkers et al. 2006). That is, we assume that the vectors $\alpha$, $\beta$, $\gamma$, and $\kappa$, as well as the scalars $\sigma$, $\lambda$, and $\rho$ are class-specific. In doing so, we establish an association between the donation process and the mailing process across donors. Letting the vector $\pi$ denote the size of each latent class and the superscript $s$ on the parameters vector for latent class $s$, the joint individual-level log-likelihood is:
The approach of jointly modeling the donation process and direct marketing activity using a latent class model to link the two model components is consistent with Donkers et al. (2006), who assume that the parameters governing the donation process and the organization’s solicitation decisions are class-specific. While we employ a latent class approach to establish an association between the model parameters across donors, this approach is also consistent with Manchanda et al. (2004) and van Diepen, Donkers and Franses (2009) who assume that the process governing a firm’s marketing activity to a customer is related to the process governing the customer’s behavior, with the association established through a relationship among the parameters of the two processes.

Though we account for the association between donation activity and the organization’s marketing activity using an approach similar to Donkers et al. (2006) and van Diepen, Donkers and Franses (2009), there are important differences in our empirical context and model. First, Donkers et al. (2006) and van Diepen, Donkers and Franses (2009) model an individual’s decision to respond to a particular mailing. In our data, however, we observe individuals making donations months after they have received direct marketing and in months when they received no direct marketing. Attributing a donation to a particular solicitation may ignore the effects of any direct marketing received between the time at which the solicitation was sent and the time at which the donation was made. Rather than model the decision to make a donation in response to a particular mailing, we model the monthly incidence decision in a manner consistent with prior literature (e.g., Fader, Hardie and Shang 2010), and allow for the impact of current

\[
LL(A_i, Direct_i | Winter, \alpha, \beta, \kappa, \sigma, \lambda, \rho) = \\
\log \left( \sum_{s=1}^{S} \left[ \pi_i L(A_i | Direct_i, Winter, \alpha^s, \beta^s, \kappa^s, \sigma^s, \lambda^s, \rho^s) \right] \times M(Direct_i | \kappa^s, Winter, Recency, Frequency, Monetary) \right) 
\]
and prior direct marketing efforts on incidence, yielding a model that is generalizable to other direct marketing contexts.

The second difference is the inclusion of latent attrition. Over time, customers are more likely to have transitioned to an inactive state, at which point direct marketing has no impact on their behavior.\(^3\) Hence, latent attrition will contribute to the decreasing impact of direct marketing’s effectiveness over time. By allowing direct marketing to impact the decision to contribute while a donor is active and the likelihood of becoming inactive, our modeling framework provides a means by which we can assess the increase in revenue associated with direct marketing activity.

4. **Empirical Analysis**

We estimate a series of models using maximum likelihood estimation. We begin with the full model, as described in equation (1) – (14). As we go from two to three latent classes, the size of the smallest latent class is less than 3%. Given the small size of the third segment, to demonstrate the applicability of our modeling framework, we estimate models using two latent classes. We then consider a model that ignores the impact of direct marketing on donation behavior \((\alpha_j=\beta_j=\gamma_j=0)\).\(^4\) Next, we assume that the model governing the organization’s mailing decisions is the same across all latent classes \((\kappa^s=\kappa)\) for all \(s\). Under this nested model, the organization’s mailing decision is the same for all

---

\(^3\) To ensure that our modeling approach is consistent with the data, in the next section we test the predictive ability of our latent attrition model against a model similar to Donkers et al. (2006) that does not incorporate latent attrition.

\(^4\) This specification still assumes that mailing decisions are associated with donation behavior, as we estimate the joint likelihood function in equation (14). We considered a model that ignores the effects of direct marketing and assumes the mailing process is the same for all donors, and we found that this model has poorer forecasting performance.
donors (i.e., a single mailing process) and, as such, is independent of the donation process conditional on a donor’s recency, frequency and average donation amount.

To gauge the empirical relevance of latent attrition, in addition to the aforementioned benchmark models, we estimate an alternative latent class model similar to Donkers et al. (2006), with RFM variables as predictors both in the organization’s mailing decision and the donor’s behavior. Under this specification, we assume that marginal distribution for monthly donation incidence follows a Bernoulli distribution and that the marginal distribution for donation amount follows a log-normal distribution. For the incidence model, we employ a probit regression using Recency, Frequency and Monetary as predictors, along with DM and Winter. We use the same predictors for the donation amount model, and assume that the error terms associated with the incidence and amount components are correlated, resulting in a class-specific monthly Tobit II model that is jointly estimated with the mailing model described in equations (12) – (13). This model does not incorporate latent attrition and explains temporal variation in donation activity using the observed covariates.

4.1. Model Results

We compare model performance during the calibration period using the BIC for each of the model specifications. To assess the predictive ability of the modeling framework, we estimate the root mean square error (RMSE) of the total donations received in each month of the forecasting period for the sample used to estimate our model (“Calibration Sample” in Table 2). To further assess forecasting performance, we extract a new 20% random sample from the database that does not overlap our calibration
sample (“Holdout Sample”) and estimate the RMSE of total donations received each month during the forecasting period using the parameter estimates from the calibration sample. We present these model fit statistics in Table 2.

<table>
<thead>
<tr>
<th>Model</th>
<th>BIC</th>
<th>Calibration Sample Forecasting RMSE</th>
<th>Holdout Sample Forecasting RMSE</th>
</tr>
</thead>
<tbody>
<tr>
<td>Full model</td>
<td>232907</td>
<td>$2093</td>
<td>$2788</td>
</tr>
<tr>
<td>No direct marketing</td>
<td>234991</td>
<td>$2941</td>
<td>$2909</td>
</tr>
<tr>
<td>Single mailing process</td>
<td>234510</td>
<td>$2178</td>
<td>$3033</td>
</tr>
<tr>
<td>RFM model</td>
<td>236592</td>
<td>$3533</td>
<td>$4517</td>
</tr>
</tbody>
</table>

Table 2. Model Performance

Since the full model outperforms the three benchmark models, we proceed to describe these results and employ them in our subsequent analyses. To further ensure that the parameters of the full model are estimated accurately, we simulated 100 datasets and estimated the full model on each dataset. All parameter estimates from our model were close to the true values. The RMSE ranges from .0001 to .60, with an average of .054. The average bias ranges from -.043 to .12. This provides evidence that bias and standard errors are low.

We present a tracking plot of total donations made each month during the holdout period in Figure 3 and see that the full model captures the general pattern in donations.
Next, we examine the parameters governing donation behavior in each of the latent classes, which are presented in Table 3.

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Description</th>
<th>Class 1</th>
<th>Class 2</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\lambda$</td>
<td>Direct marketing decay</td>
<td>.66</td>
<td>.80</td>
</tr>
<tr>
<td>$\gamma_0$</td>
<td>Incidence - intercept</td>
<td>-2.31</td>
<td>-2.37</td>
</tr>
<tr>
<td>$\gamma_1$</td>
<td>Incidence - DM</td>
<td>.45</td>
<td>.22</td>
</tr>
<tr>
<td>$\gamma_2$</td>
<td>Incidence - Winter</td>
<td>.04</td>
<td>.31</td>
</tr>
<tr>
<td>$\beta_0$</td>
<td>Amount - intercept</td>
<td>2.87</td>
<td>2.92</td>
</tr>
<tr>
<td>$\beta_1$</td>
<td>Amount - DM</td>
<td>.00</td>
<td>.20</td>
</tr>
<tr>
<td>$\beta_3$</td>
<td>Amount - Winter</td>
<td>.05</td>
<td>.08</td>
</tr>
<tr>
<td>$\sigma$</td>
<td>Amount – shape parameter</td>
<td>.47</td>
<td>1.22</td>
</tr>
<tr>
<td>$\rho$</td>
<td>Correlation</td>
<td>.07</td>
<td>.15</td>
</tr>
<tr>
<td>$\alpha_0$</td>
<td>Latent lifetime – intercept</td>
<td>6.28</td>
<td>.99</td>
</tr>
<tr>
<td>$\alpha_1$</td>
<td>Latent lifetime - DM</td>
<td>-1.54</td>
<td>.27</td>
</tr>
<tr>
<td>$\kappa_0$</td>
<td>Mailing decision - intercept</td>
<td>-.10</td>
<td>.37</td>
</tr>
<tr>
<td>$\kappa_1$</td>
<td>Mailing decision – Recency</td>
<td>-.01</td>
<td>-.04</td>
</tr>
<tr>
<td>$\kappa_3$</td>
<td>Mailing decision – Frequency</td>
<td>.11</td>
<td>.00</td>
</tr>
<tr>
<td>$\kappa_4$</td>
<td>Mailing decision – Monetary</td>
<td>.00</td>
<td>.00</td>
</tr>
<tr>
<td>$\kappa_5$</td>
<td>Mailing decision – Winter</td>
<td>2.65</td>
<td>1.19</td>
</tr>
<tr>
<td>$\pi$</td>
<td>Class Size</td>
<td>.27</td>
<td>.73</td>
</tr>
</tbody>
</table>

Table 3. Donation Behavior Parameter Estimates

Examining the parameters governing the incidence of donation, we observe some variation across latent classes. In the smaller latent class (Class 1), we observe a higher baseline probability with which donations occur, as reflected by $\gamma_0=-2.31$ (s.e.=.01)
compared to the intercept for the larger class of $\gamma_0=-2.37$ (s.e.=.01). In both classes the likelihood of donations is higher during the winter months, though the effect is larger in Class 2 ($\gamma_2=.31$, s.e.=.02) compared to Class 1 ($\gamma_2=.04$, s.e.=.02). Direct marketing increases the likelihood of donations in a given month in both classes; the impact of direct marketing is larger in Class 1 ($\gamma_1=.45$, s.e.=.01) than in Class 2 ($\gamma_1=.22$, s.e.=.01). While the impact of direct marketing on donation incidence is greater in Class 1, it also decays at a faster rate in this class, as reflected by the parameter $\lambda$.

Next, we consider the effects of covariates on amount contributed when a donation occurs. In Class 1, donation amounts are higher in the winter months ($\beta_2=.05$, s.e.=.02). The effect of winter months for individuals in Class 2, while directionally the same as Class 1, is not significant at the 95% level ($\beta_2=.08$, s.e.=.04). Examining the impact of direct marketing on the donation amount, the results are also mixed. While direct marketing increases the donation amount from individuals in Class 2 ($\beta_1=.20$, s.e.=.01), it does not impact the donation amount from individuals in Class 1. The extent of variation in donation amounts, reflected by the shape parameter $\sigma$, also varies across classes, with individuals in Class 1 displaying less variation than individuals in Class 2.

In both latent classes, we find evidence of positive correlation in the inter-donation time and donation amount processes (i.e., longer interdonation times are associated with larger donations). The extent of correlation is larger in Class 2 ($\rho=.15$, s.e.=.02) compared to Class 1 ($\rho=.07$, s.e.=.02). This is consistent with the findings from Jen, Chou and Allenby (2009), who describe the positive correlation as compensating behavior.
There are stark differences in latent attrition across the two classes. In Class 1, donors are most likely to remain active from one month to the next, as reflected by the intercept $\alpha_0$ ($\alpha_0=6.28$, s.e.=.07). However, the amount of direct marketing that donors receive diminishes their latent lifetimes ($\alpha_1=-1.54$, s.e.=.02). Hence direct marketing both increases the donation likelihood in a given month and shortens the time horizon over which donors in Class 1 make donations. This finding is consistent with direct marketing irritating donors. Thus, while direct marketing is an effective tool in the short term, reducing the time between donations while donors remain active, it accelerates latent attrition. Because Class 1 donors have a high baseline likelihood of remaining active, it may still be worthwhile to send such donors high volumes of direct marketing. As we will demonstrate in the next section, our modeling framework can be used to address such decisions, taking into account both the short-term and long-term effects of direct marketing.

A different picture emerges for donors in Class 2. These donors have a higher baseline probability with which they become inactive ($\alpha_0=.99$, s.e.=.01). Additionally, the likelihood with which they become inactive diminishes with direct marketing ($\alpha_1=.27$, s.e.=.01). Though those donors in Class 1, accounting for 27% of the donor base, are more prone to become inactive at high levels of direct marketing, the majority of donors (those in Class 2) will have longer latent lifetimes. Thus, for donors in Class 2, direct marketing not only reduces the time between donations when donors are active and the amount of donations, but it also extends the latent lifetime over which time donations are made.
The differences across latent classes underscore the importance of accounting for heterogeneity in donors’ response to direct marketing. On all three dimensions of behavior that we consider, the majority (Class 2) are positively impacted. In contrast, there is a segment of donors (Class 1) whose latent lifetime over which donations are made may be reduced by direct marketing. For these donors, the organization must be cognizant of the risks of high levels of direct marketing, as it may ultimately reduce the value of these donors. The organizations’ mailing activities appear consistent with this.

Looking at the baseline incidence with which direct marketing is sent, we observe that donors in Class 2 ($\kappa_0=0.37, \text{s.e.}=0.01$) are more likely to receive a mailing in a given month compared to donors in Class 1 ($\kappa_0=-0.10, \text{s.e.}=0.01$). This directional difference is consistent with the way in which donors react to direct marketing, as there is only a potential downside to sending mailings to donors in Class 1.

For both latent classes, the likelihood with which mailings are sent decreases as the time since the last donation increases (Class 1: $\kappa_1=-0.01, \text{s.e.}=0.001$; Class 2: $\kappa_1=-0.04, \text{s.e.}=0.001$), which is consistent with organizations pruning their mailing lists if donors have not contributed recently. Frequent donors in Class 1 are more likely to receive mailings ($\kappa_2=0.11, \text{s.e.}=0.002$). Though directionally consistent, the impact of frequency on the mailing decision for donors in Class 2 is not significant at the 95% level ($\kappa_2=0.00, \text{s.e.}=0.002$). The organization is also more prone to send mailings in the winter months (Class 1: $\kappa_4=2.65, \text{s.e.}=0.09$; Class 2: $\kappa_4=1.19, \text{s.e.}=0.01$).

4.2. Residual Lifetime Value as a Measure of Marketing Impact
We next turn our attention to expectations of future behavior and evaluating the effectiveness of the organization’s direct marketing. Fader, Hardie and Shang (2010) derive an expression for the discounted expected residual transactions, which takes into account not only the likelihood that an individual is still active, but also the frequency with which transactions occur while the individual is active. By incorporating the effects of direct marketing into the frequency of transactions while a donor is active, the size of donations, and the donor’s latent lifetime, a donor’s RLV can be estimated as a function of a given direct marketing strategy. The organization can then estimate the increase in a donor’s RLV attributable to its direct marketing efforts, which can then be weighed against the cost of these activities.

We consider different direct marketing strategies during the 12 months following our calibration period and examine the impact of these policies on RLV of our calibration sample. As a baseline scenario, we conducted 1,000 simulations of donation behavior over a 50-year period with no direct marketing activity, assuming a 15% annual discount rate, after which time the present value of future contributions is negligible. By varying direct marketing activity during the 12-month period and comparing the resulting RLV to that of our baseline scenario, we can assess the financial impact of alternative policies.

In our first comparison of direct marketing policies, we hold the number of mailings fixed and vary the timing with which direct marketing is sent. We begin by assuming that mailings are sent in alternating months during the 12-month period, mirroring the mailing activity observed during our calibration period as shown in Figure 2. We consider an alternative scenario in which mailings are sent only in the first half of the 12-month period, as well as a scenario in which mailings are only sent in the second
half of the 12-month period. Compared to the pulsing strategy of sending mailings in alternating months, these policies more closely resemble a blitz strategy (Mahajan and Muller 1986). As a middle ground, we consider a policy in which mailings are sent during the 1st and 3rd quarters of the 12-month period, as well as a policy with mailings sent in the 2nd and 4th quarters.\(^5\)

The impact of direct marketing differs between the two latent classes. As latent class membership is unobserved and inferred from past behavior, we examine how estimates of RLV vary with summary measures of donors’ past activity. Mirroring the operations of many direct marketing organizations, we derive RFM segments with two levels for each variable. We define “recent” donors as those who have made a contribution within one year of the end of the calibration period, “frequent” donors as those who have made at least two repeat donations, and “high amount” donors as those whose average contribution exceeded the mean average contribution (across donors). In Table 4, we show the average increase in RLV (across donors in each RFM segment) for each scenario compared to the baseline scenario (i.e., no direct marketing activity for the 12 month period).

<table>
<thead>
<tr>
<th>Recency</th>
<th>Frequency</th>
<th>Monetary</th>
<th>Alternating</th>
<th>1st Half</th>
<th>2nd Half</th>
<th>1st &amp; 3rd Quarters</th>
<th>2nd and 4th Quarters</th>
</tr>
</thead>
<tbody>
<tr>
<td>Recent</td>
<td>Frequent</td>
<td>High</td>
<td>$20.74</td>
<td>$34.54</td>
<td>$11.94</td>
<td>$26.06</td>
<td>$14.21</td>
</tr>
<tr>
<td>Recent</td>
<td>Frequent</td>
<td>Low</td>
<td>$9.89</td>
<td>$12.33</td>
<td>$8.98</td>
<td>$11.17</td>
<td>$8.76</td>
</tr>
<tr>
<td>Recent</td>
<td>Infrequent</td>
<td>High</td>
<td>$15.48</td>
<td>$26.27</td>
<td>$8.94</td>
<td>$19.66</td>
<td>$11.01</td>
</tr>
<tr>
<td>Recent</td>
<td>Infrequent</td>
<td>Low</td>
<td>$10.91</td>
<td>$16.14</td>
<td>$8.48</td>
<td>$12.97</td>
<td>$8.92</td>
</tr>
<tr>
<td>Nonrecent</td>
<td>Frequent</td>
<td>High</td>
<td>$10.33</td>
<td>$17.19</td>
<td>$6.01</td>
<td>$13.17</td>
<td>$7.09</td>
</tr>
<tr>
<td>Nonrecent</td>
<td>Frequent</td>
<td>Low</td>
<td>$6.77</td>
<td>$8.86</td>
<td>$6.02</td>
<td>$7.78</td>
<td>$5.90</td>
</tr>
<tr>
<td>Nonrecent</td>
<td>Infrequent</td>
<td>High</td>
<td>$2.47</td>
<td>$4.14</td>
<td>$1.45</td>
<td>$3.04</td>
<td>$1.72</td>
</tr>
<tr>
<td>Nonrecent</td>
<td>Infrequent</td>
<td>Low</td>
<td>$1.66</td>
<td>$2.34</td>
<td>$1.34</td>
<td>$1.97</td>
<td>$1.36</td>
</tr>
</tbody>
</table>

Table 4. Increase in Residual Lifetime Value from Alternative Direct Marketing Policies

---

\(^5\) So that we can draw comparisons to our baseline scenario with no direct marketing, in each of these scenarios, we assume no direct marketing activity after the 12-month period.
While each of these policies entails sending 6 mailings to donors during the 12-month period, we see a wide range in the expected increase in RLV for all of the RFM segments. Among these policies, sending mailings in the first half of the period yields the largest increase, followed by sending mailings in the first and third quarters. Meanwhile, sending mailings only in the second half yields the smallest expected increase. Moreover, the two best of these policies both yield a larger expected increase compared to sending mailings in alternating months; this suggests a benefit to sending mailings earlier and in bursts rather than sending mailings in alternating months.

As the policy in which direct marketing is sent for the first half of the 12-month period yields the highest average increase in RLV for each of RFM segment in Table 4, we next examine a set of mailing policies in which mailings are sent continually during the 12-month period until the organization decides to cease mailing to donors. We simulated donation behavior under 12 different scenarios, with mailings sent in the first month, the first two months, the first three months, and so forth. By taking the difference in RLV between the scenario with mailings sent the first \( t-1 \) months and the scenario with mailings sent the first \( t \) months, we can derive the incremental RLV associated with sending one more mailing (e.g., Braun and Schweidel 2011; Schweidel, Bradlow and Fader 2011), which can serve as a measure of impact for the last mailing sent. The average incremental RLV for 12 mailings is presented, by RFM segment, in Figure 4.
Across the RFM segments, there is an initial increase in the incremental contribution to RLV, which peaks during the winter months (months 7-9 of the 12-month period), followed by a subsequent general decline. This suggests that there can be initially increasing returns to direct marketing activity as direct marketing in the current period may build upon efforts in prior periods. As direct marketing may only impact those donors who are still active, though, donors’ increasing likelihood of having become inactive, as well as discounting of future revenue, contribute to the reduced levels of incremental RLV in later periods.

We find the highest incremental RLV associated with marketing to recent donors who have a high average donation amount; the incremental RLV for frequent donors ($4.52 – $8.47) is larger than that of infrequent donors ($2.92 – $7.69). Though recent and frequent donors with a low average donation amount may still be active, the incremental RLV associated with mailing to these donors ($1.58 – $2.24) during the 12-
month period is less than the incremental RLV of mailing to recent, infrequent donors with low average amounts ($2.00 – $3.86) and nonrecent, frequent donors with high average amounts ($2.27 – $3.93) in most months.

As an additional illustration, we consider the incremental RLV under a model in which mailings are sent in alternating months in Figure 5.

![Figure 5. Incremental RLV of Mailings in Alternating Months](image)

Under this mailing policy, we do not observe the initial increasing returns expected we saw when mailings are sent in consecutive months. Instead, we find that the incremental RLV is flat or declining early on, with the exception of the winter months in which donation activity occurs at a higher level, followed by steeper declines in later months. As in Figure 4, we observe a segment of non-recent donors whose incremental RLV from receiving additional mailings is on par with or greater than the incremental RLV of some recent donors.

As Figures 4 and 5 illustrate, on the basis of the additional revenue expected from sending direct marketing, recent donors are not necessarily the most promising targets.
Some donors are likely to contribute in the future, potentially without receiving a high volume of direct marketing activity. On the other hand, others may be unlikely to contribute unless they are prompted by direct marketing activity. Rather than relying on heuristics such as sending mailings only to those who have contributed within a specified number of months, deriving the incremental RLV associated with sending an additional mailing and the increase in RLV associated with a given mailing policy allows the organization to quantify the benefits of direct marketing and compare this to their costs. Using heuristics such as recency may result in the organization ignoring donors for whom its direct marketing efforts may make a substantial difference in the amount contributed while exerting efforts on those individuals who may have contributed without as much prompting.

**CONCLUSION**

Though customer base analysis remains a central part of CRM research and customer valuation, many models do not consider the firms’ marketing activities. Our analysis incorporates time-varying covariates into a latent attrition framework and allows for marketing to influence the rate at which individuals make contributions to the non-profit organization, the amount of the contributions, and the length of the latent lifetime over which they make contributions. We also account for within-donor correlation between inter-donation times and donation amounts, as well as the nonrandom nature of the organization’s mailings.
Though we find that direct marketing reduces the time between donations, we find mixed results as far as its impact on the latent lifetime. Whereas direct marketing increases the latent lifetime for most donors, consistent with strengthening the relationship between the individual and the organization, there are individuals who appear irritated by direct marketing and whose latent lifetimes are shortened by it. By accounting for direct marketing’s impact on three dimensions that comprise transactional activity (latent lifetime, transaction frequency, and transaction amount), we demonstrate how the latent attrition framework can be used to estimate the RLV of donors under alternative direct marketing plans and the incremental RLV associated with sending an additional mailing, a metric that can be used for managerial decisions including mail scheduling and customer selection. As we illustrate, the framework can be used for assessing the increased transactional activity associated with alternative direct marketing policies, enabling the organization to tailor them to individuals or segments.

There are a number of directions in which the current research can be extended. First, though we do not consider the content of mailings, if the right data exists, our framework could be used to examine which message is most effective for which individuals, enabling the organization to customize the appeals sent to different individuals based on their response to past mailings. Future research should examine the role of other, potentially complementary forms of marketing, such as coupling direct marketing efforts with electronic communications, personal contact with development offices, mass marketing, and social media strategies. Though some tactics may reduce the time between transactions, other forms may influence the duration of donors’ (or customers’) latent relationship. Such a nuanced understanding is key not only in
estimating customer value, but also in estimating the effectiveness of marketing actions so that resources may be allocated appropriately based both on both the impact and cost of different marketing decisions.
REFERENCES


