The authors are grateful to Brett Gordon for his help and guidance in implementing the value iteration method used in this work. We are also indebted to Rong Guo for preparing the data used in the analysis. Finally, we thank an anonymous online food service company for providing the data used in this research. The first two authors are listed in reverse alphabetical order.
Managing CLV Using the Migration Model Framework: Overcoming the ‘Recency Trap’

ABSTRACT

The migration model of customer lifetime value classifies customers into recency states depending on how long it has been since their previous purchase. Purchase likelihood typically declines as recency increases. As a result, firms face a "recency trap," whereby recency increases for customers who do not purchase in a given period, making it less likely they will purchase in the next period. The goal therefore is to target marketing depending on the customer's recency to prevent the customer from lapsing to such high recency that the customer is essentially lost. We develop a modeling approach to achieving this goal. This requires a model of purchase as a function of recency and marketing efforts, and a dynamic optimization that incorporates these purchase probabilities and the trade-offs in acting now rather than later. In our application we find that purchase likelihood as well as customer response to marketing depend on recency. The results specify how the targeting of email and direct mail should depend on the customer's recency, and show how this would increase firm profits.

Keywords: Database Marketing, Customer Lifetime Value, Optimization, Customer Recency
Managing CLV Using the Migration Model Framework: Overcoming the ‘Recency Trap’

1. INTRODUCTION

Allocating marketing efforts over time in today’s data-intensive, customer-interactive environment poses both opportunities and challenges. The opportunity is the promise of contacting the right customer at the right time using the right marketing instrument. The challenge is in the dynamics. The marketing decisions we make today may affect what we wish to do tomorrow. For example, if there is email “saturation”, today’s email campaign may render tomorrow’s email campaign less effective. The customer is constantly changing, in a different “state”, over time. This means that today’s consumer is not the same as tomorrow’s. Not taking into account these dynamics can result in mis-targeting and mis-timing of marketing actions.

An important dynamic is the concept of “recency” – how long it has been since the customer’s previous purchase. Recency has been found to be highly correlated with customer purchase, and directly related to customer lifetime value through the “migration model” of CLV (Berger and Nasr 1998; Pfeifer and Carraway 2000; Blattberg, Kim, and Neslin 2008). It is imperative to take into account recency in order to target the right customers at the right time. However, “managing” customer recency has its challenges. For example, a common finding is that higher recency (longer time since previous purchase) is associated with lower purchase likelihood (Bult and Wansbeek 1995; Bitran and Mondschein 1996; Fader, Hardie, and Lee 2005). As a result, firms face a “recency trap”: When customers do not purchase in a given period, this increases their recency, which makes it less likely they will purchase in the next period and thereby will transition to a still higher recency state, where they are even less likely to
purchase, etc. The result is that the customer is drifting away from the company and the lifetime value of the customer is decreasing.

Confronted with the recency trap, should the firm turn up its marketing efforts for high recency customers, or give up and let the customer lapse into oblivion? If the firm turns up its marketing efforts for high recency customers, is it wasting its money and even desensitizing the customer to future efforts? On top of this, the firm has multiple marketing instruments at its disposal. Which ones should it use and when?

The purpose of this paper is to devise a procedure that prescribes what marketing efforts should be targeted to which customers at which time, exploiting the relationship between recency and purchase and its link to CLV. We estimate a purchase model that is a function of recency, marketing, and their interaction, as well as other marketing dynamics including carryover and saturation. We then use an infinite horizon dynamic program to derive the optimal decision policy for two marketing instruments – in our application, email and direct mail promotions.

Our paper aims to contribute to the burgeoning literature on what Blattberg, Kim, and Neslin (2008) call “optimal contact models.” The theme of these models is that CLV is something to be managed, not merely measured. They rely on a customer response model and a dynamic optimization, although they differ in the marketing variables considered, the method of optimization, and the particular phenomena included in the response model. Our paper is unique in its use of recency as a conceptual foundation, and the combination of issues it addresses: (1) consideration of recency/marketing interactions, (2) consideration of marketing carryover and saturation, (3) consideration of more than one marketing instrument, and (4) use of the value iteration approach to solving dynamic programs.
We apply our approach to a meal preparation service provider whose key marketing tools are email and direct mail promotions. Our customer marketing response function shows that recency is related negatively to purchase probabilities, setting up the recency trap. We also find that marketing interacts with recency and is subject to carryover and saturation. The nature of these effects differs for email and direct mail. Direct mail has much more carryover and interacts positively with recency. Email is particularly subject to saturation effects although it has zero distribution cost. Our optimization balances these considerations by accounting explicitly for customer migration between recency states. Our application suggests four key findings: (1) the firm currently is underutilizing both email and direct mail, (2) more budget should be allocated to direct mail than email, (3) marketing efforts should generally increase as customer recency increases, whereas the firm’s current policy does not target in this way, and (4) we predict that implementation of our procedure would increase CLV by $175-$200, depending on the recency state of the customer.

We proceed to review the literature in more detail. Then we discuss our model, beginning with a detailed illustration of the recency trap, and a description of our response model and optimization. Next we describe the data for our application, and finally, the application itself. We close with a discussion of implications for researchers and practitioners.

II. LITERATURE REVIEW

II.1 Customer Recency and Customer Lifetime Value

Customer recency has long been recognized as a key concept in CRM, comprising, along with its cousins, purchase frequency and monetary value, the RFM framework that has been used
for years as a segmentation tool by direct marketers (Blattberg, Kim, and Neslin 2008, Chapter 12). It was therefore quite natural for predictive modelers to incorporate recency in their efforts to predict customer behavior. Bult and Wansbeek (1995), Bitran and Mondschein (1996), and Fader, Hardie, and Lee (2005) find a negative association between recency and purchase likelihood. These findings reinforce the common belief that “Consistently, the most recent buyers out-perform all others” (Miglautsch 2002, p. 319), and that “Many direct marketers believe that the negative relationship is a law” (Blattberg, Kim, and Neslin 2008, p. 325).

Blattberg et al. note however that the relationship between recency and purchase likelihood may differ by category. For example, Khan, Lewis and Singh (2009) find for an online grocery retailer that the relationship is positive at first, peaks at about four weeks, and then declines. This still begets a recency trap, because once the customer has not purchased in four weeks, he or she tends to transition to higher recency and lower purchase probabilities. Some researchers have found, within the range of their data, a positive relationship between recency and purchase (e.g., Gönül and Shi 1998, and Gönül, Kim, and Shi 2000 for a durable goods cataloger, and Van den Poel and Leunis (1998) for financial services). These findings may be due to a long purchase cycle; with a long enough data history, high recency would mean lower purchase likelihood. For example, a customer may replace a television every five years, but if those five years pass by and the customer has not purchased from the company, it is likely the customer has purchased from a competitor and the hence the probability of purchasing from the focal company would decline with higher recency.

In summary, while there are exceptions, the common finding is that higher recency means lower purchase likelihood. There is empirical evidence (Khan et al., 2009), and it is reasonable to believe, that even if the relationship is not negative at first, it becomes negative in
the long run. This begets the recency trap. Our procedure does not require the negative relationship; it works for any relationship between recency and purchase. However, we focus on the negative relationship and the resultant recency trap because the negative relationship appears to be most common.

A key breakthrough was to relate recency to customer lifetime value. Important contributions include Berger and Nasr (1998) and Pfeifer and Carraway (2000). These contributions viewed customer lifetime value as a Markov chain with recency serving as the state variable. Customers transition from one recency state to another depending on whether they purchase or not. If the customer purchases, the customer is placed in recency state 1, meaning “just purchased”. If the customer purchases in period 1 but not in period 2, the customer transitions to recency state 2, meaning that at the outset of period 3, the customer last purchased two periods ago, in period 1. Berger and Nasr show the details for calculating CLV using this framework, and Pfeifer and Carraway provide general formulas using matrix algebra.

II.2  

Customer Response to Email and Direct Mail

Many predictive models find that email and direct mail affect purchase likelihood. The evidence regarding email is more recent and less definitive. An important recent paper by Drèze and Bonfrer (2008) found that the scheduling of email solicitations could affect consumer response in terms of customer retention as well as the customer’s tendency to open and click on an email message. This is perhaps related to the traditional effects found with regard to advertising, namely carryover and saturation. Carryover means that a marketing activity in period $t$ has an impact on customer response in period $t+1$. This may be due to the customer remembering the message for more than one period, or simply due to a delay between the
reception of the message and the opportunity to act upon it. Pauwels and Neslin (2008) find evidence of carryover. Saturation means that marketing in period $t$ reduces the impact of marketing in period $t+1$. This could occur due to clutter, or that the customer anticipates that no new information is provided in period $t+1$ because the customer has recently heard from the company. Ansari, Mela, and Neslin (2008) find evidence of saturation.

An extreme form of saturation, “supersaturation”, has been conjectured (e.g., Leeflang, Wittink, Wedel, and Naert 2000, p. 68), whereby high levels of marketing in period $t$ mean that high levels of marketing in $t+1$ decrease purchase likelihood in that period. This could be due to customer irritation (Van Diepen, Donkers, and Franses 2009) or information overload – after a surfeit of emails in period $t$, continuing that level in period $t+1$ encourages the customer to collect them in his/her inbox and ignore them all. As a result, the email=>purchase relationship becomes negative in period $t+1$. Van Diepen et al. looked for supersaturation and didn’t find it. However, early field experiments by Ackoff and Emshoff (1975) found evidence of supersaturation, and more recent work by Naik and Piersma (2002) found that cumulative marketing expenditures related negatively to customer goodwill. Naik and Piersma suggested that as a result, optimal marketing policies may involve “pulsing” to avoid overloading the customer.

II.3 Optimal Contact Models

One of the most exciting areas of CRM research is optimal contact models (Blattberg, Kim, and Neslin 2008). Optimal contact models determine what marketing efforts should be expended on which customers at what time. These models are thus dynamic. They integrate inherently dynamic phenomena such as recency into a prescription for an optimal marketing
policy. There are a large variety of optimal contact models, beginning with the pioneering work of Bitran and Mondschein (1996), and continuing with important contributions by Gönül and Shi (1998), Gönül, Kim, and Shi (2000), Elsner, Krafft, and Huchzermeir (2003, 2004), Rust and Verhoef (2005), Simester, Sun, and Tsitsiklis (2006), and Khan, Lewis, and Singh (2009). All these papers make significant contributions to a burgeoning literature on optimal contact models.

Many of these papers focus on catalog mailings. The catalog industry is a major innovator in CRM, so the emphasis of these papers on catalog applications is not surprising. Rust and Verhoef (2005), and Khan, Lewis, and Singh (2009) focus on multiple marketing activities. Rust and Verhoef consider direct mail and a customer relationship magazine; Khan et al. consider discount coupons, loyalty rewards, and free shipping. Consideration of multiple marketing instruments is important because it is realistic and makes the analysis more complex. Though challenging, this approach addresses a key issue: allocation of marketing investment.

The basic components of an optimal contact model are (1) a customer response model, i.e., a predictive model of how customers respond to marketing, and (2) a method for optimization. Previous papers have used a variety of response models, including hazard models (Khan et al.), RFM categorizations (Bitran and Mondschein), and decision trees (Simester et al.). The optimization usually employs dynamic programming. Dynamic programming is necessary because “forward-looking” is crucial – the actions the firm takes with the customer today may influence what actions it may want to take with them in the future. Dynamic programming methods include infinite horizon models (Simester et al.), rolling horizon models (Neslin, Novak, Baker, and Hoffman 2009), and finite horizon (Khan et al.). Khan et al. note that all these approaches have their advantages and disadvantages. Finite horizon optimization can run into end-game issues, whereby marketing efforts may be distorted at the end of the time horizon ($T$)
because there are no explicit costs or benefits in time $T+1$. On the other hand, infinite horizon methods can be computationally cumbersome. In our paper, we use an infinite horizon dynamic program solved using value iteration, which is relatively simple to program.

II.4 Unique Contributions of This Paper

Our paper is an optimal contact model that is unique in the combination of issues we address:

- Use of recency as a foundation of the model.
- Consideration of interactions between recency and marketing response.
- Derivation of joint policies for multiple marketing instruments – email and direct mail.
- Consideration of saturation and carryover effects of marketing.
- Use of an infinite horizon optimization utilizing value iteration (Judd 1998).

Our overall objective is to develop an optimal contact model that is “complete on the important issues” (Little 1970), yet relatively simple and managerially relevant. Our emphasis on recency stems from the multitude of studies that have shown the importance of this variable, its link to customer lifetime value, and the phenomenon of the recency trap. We also believe it is important to consider the rich set of phenomena that govern customer response to marketing actions, including saturation, carryover, and interactions with recency. Many managers must coordinate and allocate funds between multiple marketing instruments; in this case we consider email and direct mail. Finally, the use of an infinite horizon optimization provides the benefits of considering the long term, unbridled by end-game effects, and the solution mechanism – value iteration – is computationally nontrivial but certainly feasible.
Probably the two most closely related papers to ours are Rust and Verhoef (2005) and Khan, Lewis, and Singh (2009), because they both deal with multiple marketing instruments. Compared to Rust and Verhoef, we emphasize the role of recency, we consider interactions between email and recency as well as saturation and carryover effects, and perform an infinite horizon optimization. Compared to Khan et al., we consider saturation and carryover effects and perform an infinite horizon optimization. Also, while Khan et al. include recency and find interactions between recency and marketing response, we place more emphasis on recency as a foundation for our modeling framework.

III. MODELING FRAMEWORK

Our modeling framework consists of three elements: (1) recognition of the role of recency in determining customer purchase, and the possibility of a recency trap, (2) a logistic regression model of customer purchase that focuses heavily on recency, and (3) a dynamic programming optimization that recognizes recency as an important characterization of the customer at any point in time. These three elements enable us to formulate a model that determines the targeting, timing, and total quantity of marketing efforts, as well as the relative allocation of funds spent on different marketing efforts (in this case, email and direct mail).

III.1 The Role of Recency and the Recency Trap

Figure 1 highlights the key phenomenon at work— that recency is highly associated with purchase likelihood. In this case, based on descriptive statistics from our application, the relationship is negative, similar to that found in previously cited research. Figure 1 shows the effect is particularly pronounced, with customers who have just purchased (recency state 1)
having a 23.1% chance of repurchasing the next period, whereas customers who have not purchased for five months (recency state 5) have only a 4.6% chance of purchasing.

The ramifications of the recency/purchase relationship for CLV is shown vividly in Table 1. Table 1 uses the migration model of CLV to calculate the probability a customer acquired in period 1 will be in various recency states at all future points in time. For example, by the end of period 7, there is a 9% chance the customer will be in recency state 5, i.e., the last purchase was five period ago, in period 3. The recency 1 column in Table 1 is most crucial, because it shows the probability the customer has purchased in each period. The numbers in the top row of Table 1 govern these calculations and are identical to the probabilities shown in Figure 1. They are the conditional probabilities the customer will make a purchase in the current period, given his or her recency state $S$ ($ProbPurch(S)$). In Table 1, the customer migrates to state 1 (just purchased) with probability $ProbPurchase(S)$. However, with probability $1 - ProbPurchase(S)$, the customer migrates to a higher recency state, $S+1$, creating the recency trap.

Table 1 shows how the recency trap plays out. The dominant tendency for the newly acquired customer is to make an initial purchase and then not purchase for several periods, sliding to recency state $\geq 20$. Sometimes the customer in a high recency state makes a purchase. For example, even a customer who has not purchased in 13 periods has a 1.2% chance of purchasing in the current period. But clearly the company is losing its hold on the customer. In the long term, since recency state $\geq 20$ is not absorbing, there is still a 1% chance a customer in that state will purchase, but 89% of the time the customer will be in recency state $\geq 20$, virtually lost to the firm.
Figure 1 and Table 1 pose the managerial problem in vivid terms – devise a targeted marketing strategy that will arrest the drifting away of a newly acquired customer. This strategy will depend on customer purchase probabilities, since they drive the recency trap. We will now estimate these probabilities as a function of marketing. This will enable us to maximize CLV within the migration model framework.

III.2 Logistic Response Model of Purchase Probability

The logistic model of purchase is a simple response model that has been used in numerous applications (e.g., see Neslin et al. 2006). The dependent variable of interest is:

- $\text{Purchase}_{it}$ – a dummy variable equal to 1 if customer $i$ purchases in period $t$; 0 if not.

We will include the following explanatory variables for predicting this dependent variable:

- $\text{Email}_t$ and $\text{Dmail}_t$ – the marketing efforts expended by the firm in period $t$, in our case, either email or direct mail offers.
- $\text{Recency}_{it}$ – the recency state of customer $i$ in period $t$.
- $\text{Recency}_{it}^2$ and $\text{Recency}_{it}^3$ – these variables capture the possibility that the relationship between recency and purchase is non-linear, beyond the inherent nonlinearities included in a logistic regression.
- $\text{Email}_t \times \text{Recency}_{it}$ – the interaction between Email and recency; a significant coefficient means that customers in different recency states respond differently to email offers.
- $\text{Dmail}_t \times \text{Recency}_{it}$ – the interaction between direct mail offers and recency
- $\text{Email}_t \times \text{Recency}_{it}^2$, $\text{Email}_t \times \text{Recency}_{it}^3$, $\text{Dmail}_t \times \text{Recency}_{it}^2$, and $\text{Dmail}_t \times \text{Recency}_{it}^3$ – these variables capture possible nonlinear interactions between recency and current...
marketing efforts. For example, it may be that customers in the middle recency states (e.g., 5-10) respond more readily to marketing solicitations.

- Email_{t-1} and Dmail_{t-1} – these lagged variables represent carryover effects of marketing. I.e., an offer received in period $t-1$ may have an impact on purchasing in period $t$.

- Email_t \times Email_{t-1} and Dmail_t \times Dmail_{t-1} – these terms represent potential saturation effects. E.g., a negative coefficient for Email_t \times Email_{t-1} means that large email efforts in the previous period render the email efforts in the current period less effective. It is possible of course that the coefficient could be positive, which would represent synergistic effects of prolonged campaigns.

- Month_t – the month pertaining to the particular customer observation (January, February, etc.). We use a dummy variable for each month, since month is the unit of observation.

- First_Amt_t – this is a variable to control for inherent cross-customer differences in preference for the firm. It equals the amount the customer spent on the first purchase when he or she was acquired. We expect the coefficient for this variable to be positive, because customers who start off by making a large purchase are probably very sure they like the product and are therefore likely to purchase on an ongoing basis (see Fader, Hardie, and Jerath 2007).

Collecting these variables into an $n \times k$ matrix $X_{it}$, where $n$ is the number of customer/period observations, and $k$ is the number of explanatory variables described above, the logistic regression model is:

\[
Prob(Purchase_{it} = 1) = \frac{1}{(1+exp^{-X_{it} \beta})}
\]
III.3 Dynamic Program

Once we have estimated equation (1), we know how customers respond to marketing efforts, and can derive a policy that will maximize the lifetime value of the customer. The lifetime value of customer $i$ ($CLV_i$) can be expressed as:

$$CLV_i = \sum_{t=0}^{\infty} \pi_{it}(D|S) \times \delta^t$$

where:

- $\pi_{it}(D|S) = \text{Profit contributed by customer } i \text{ in period } t, \text{ given the customer is in state } \text{“S” in that period and the marketing decision } D \text{ is made with respect to that customer.}$ The decision $D$ in our case will be how much emailing and direct mailing to expend on that customer. The “state variables” that define the customer are those that affect current profitability and change over time. In our case, recency will be a state variable, as well as previous email/direct mail efforts, and month.
- $\delta = \text{discount factor, e.g., } 0.995 \text{ on a monthly basis means that profits achieved one year from the present are worth } 94\% (0.995^{12}) \text{ of what they are worth today.}$

Equation (2) emphasizes that the lifetime value of the customer is not a static number – it is an objective to be managed through marketing efforts. These efforts in turn depend on the state of the customer at that time. The challenge is to find the decision policy $(D|S)$ that maximizes $CLV$, taking into account that current actions may place the customer in a different state in the next period, which affects our optimal decision in that period. For example, if the logistic regression finds saturation effects, but otherwise, customers in high recency states are more likely to respond to marketing, it may be optimal not to market to the customer in the
current period, but put this off to the next period. This factor of course may be counter-balanced
by the lower “baseline” purchase probabilities inherent in higher recency states.

Methods that derive the marketing policy to optimize the dynamic program specified by
equation (2) typically work with the “value function”, \( V_{it}(S) \), the maximum expected long-term
profit we can gain from a customer given the customer is in state \( S \) at time \( t \). Value functions
have intuitive interpretations in a customer management environment – they represent the
lifetime value of a customer who starts in state \( S \). The key relationship derived in dynamic
programming theory is that the value function in period \( t \) equals the expected profit we derive
from finding the decision that maximizes the current period profit of the customer plus what we
henceforth expect to gain (on a discounted basis) from the customer, given the decision we’ve
made in period \( t \). In equation form:

\[
(3) \quad V_{it}(S) = \max_{D|S}\{E(\pi_{it}(S)) + \delta E(V_{it+1}(S))\}
\]

Equation (3) presents the customer management viewpoint that we should do what we can to
maximize current period expected profits, but do so in light of the future profits we can expect to
make because of the actions we take in the current period.

In our case, the expected future profits, represented by \( E(V_{it+1}(S)) \) take on a simple
form, because the customer either will purchase or not purchase. We therefore have:

\[
(4) \quad V_{it}(S) = \max_{D|S}\{E(\pi_{it}(S))
+ \delta[ProbPurch(S)_{t} \times E(V_{it+1}(S')) + (1 - ProbPurch(S)_{t}) \times E(V_{it+1}(S''))]\}
\]
ProbPurch(S), will be calculated using our logistic purchase model and will depend on what state the customer is in. The expected future value functions now are conditioned on the customer being in different states, $S'$ and $S''$. For example, if the customer purchases in period $t$, we know the customer will be in recency state 1 in period $t+1$, by definition. So $S'=1$. However, say the customer is currently in recency state 5 (it has been 5 periods since the last period), and doesn’t purchase in period $t$, then the customer will shift to recency state 6, so $S''=6$. The optimal decision to make in period $t+1$ will differ depending on whether recency equals 1 or 6, and so we are accounting for this in our expression for the value function.

Equation (4) integrates the migration model of CLV with finding the marketing policy that maximizes CLV, and shows how the optimization manages the recency trap. Recall from our earlier discussion of Table 1, the customer migrates to recency state 1 with probability $\text{ProbPurch}(S)$, and moves to a higher recency state with probability $1 - \text{ProbPurch}(S)$. In equation (4), we can see how we are taking this into consideration. We take the action that maximizes current period profit, plus what we intend to do if the customer buys (and migrates to recency state 1), which happens with probability $\text{ProbPurch}(S)_t$, as well as what we intend to do if the customer does not buy (and migrates to a higher recency state), which happens with probability $(1 - \text{ProbPurch}(S))_t$. The optimization model thus integrates the migration model of CLV and optimal targeted marketing while addressing the recency trap.

We have not yet specified the profit function $E(\pi_{tt}(S))$. This function involves application-specific costs, etc., and so we describe it fully in Section V (Application). We also describe in Section V the method we use to derive the optimal profit function $D|S$, and the state variables we use besides recency to describe the current status of the customer.
The data for our application come from a meal preparation service provider. Customers log on to the company’s website and order the meals they will assemble during their visit to the service establishment or the meals they wish to pick up that have been pre-assembled. Ordering is primarily done online after the customer has logged onto the company website, so customer-level data are easily collected. The data span 25 months, October, 2006 through November, 2008. We have data on 4121 customers who made an initial purchase. These customers made a total of 4260 additional purchases, an average of one additional purchase per customer (consistent with the data in Table 1). These are the purchases we model, the ones that occur after the customer has been acquired. Customers are acquired at different times, so that on average we observe the customer for 15.0332 months. This means in total we have $4121 \times 15.0332 = 61,952$ customer-month observations available for the logistic regression purchase model.

The two chief marketing instruments used by the firm were email and direct mail promotions. These promotions varied in form, but the “bottom” line was that they all offered a discount on purchased merchandise, during certain periods of time. For example, an email could alert the customer that a promotion was in effect during a specified three-week period. This provided us a means to quantify promotion. In particular, we created monthly email and direct mail variables so that a month-long promotion would assume a value of one. A value of 0.75 associated with an email sent in a particular month means that the email announced a promotion that was available for three weeks. This procedure yielded monthly email and direct mail variables, representing how many months worth of promotion were announced by those
communications. We found that the average email variable, for example, was 0.67 (see Table 2), meaning the average email-communicated promotion was in effect for a little less than three weeks during the month it was announced. The values for the email variable ranged from 0 to 2.5, with an average of 0.67, while the values of the direct mail variable ranged from 0 to 2.9, with an average of 1.52. Values greater than one are possible because there may have been more than one email or direct mail campaign in a given month. Table 2 describes these and other variable used in the model, and provides descriptive statistics.

[Table 2 Goes Here]

V. APPLICATION

V.1 Logistic Regression of Purchase Probability

We estimate equation (1) in stages, adding variables to demonstrate the impact of email and direct mail, the role of recency, and derive a final model. Table 3 shows the results.

[Table 3 Goes Here]

The base model includes just recency as well as the monthly dummies and the First_Amt control variable. The recency variables – linear, squared, and cubed terms – are all highly significant as expected given Figure 1. The First_Amt variable is highly significant, and six of the 11 monthly dummies are significant at the 5% level.

Model 2, adds the basic marketing variables: (1) the current period effect (Email and Dmail), (2) the lagged effect (Lagged_Email and Lagged_Dmail), and (3) the interactions between the main and lagged variables, measuring saturation. The addition of these variables “costs” six degrees of freedom, but the likelihood ratio test shown in the bottom three lines of
Table 3 finds that the contribution to fit is statistically significant. The results suggest that carryover and saturation effects are present. Carryover is particularly strong for direct mail; saturation is present for both email and direct mail, significant at the 0.039 level for email albeit only marginally significant for direct mail. Overall the key finding is that the classic marketing effects – current period, carryover, and saturation – are apparent in the data and add to overall fit.

Model 3 adds interactions between marketing and recency. The likelihood ratio test is significant at the 0.014 level, indicating that these interactions add to fit. Specifically, the interaction is not significant for email, but is highly significant (p-value = 0.011) for direct mail. The positive sign means that as customers lapse to higher recency states, they become more receptive to direct mail.

Model 4 adds interactions between marketing and recency-squared. The likelihood ratio test here is significant at the 0.053 level. Model 5 adds interactions between marketing and recency-cubed. This model clearly does not improve fit – the likelihood ratio test has a p-value of 0.232. Given the results of the likelihood ratio test, we decided to use Model 3 (with just the linear reaction between marketing and recency) for our optimization. One could argue that Model 4 could also be used (0.053 is close to p<0.05) but we decided to be conservative and stay with the simpler model. We believe what’s important is the process of model-building, in this case, starting with the “tried and true” conventional marketing effects (current period, carryover, and saturation), and then investigating interactions between recency and marketing.

Figures 2-4 provide graphical illustrations of the effects quantified by the logistic regression. Figure 2 graphs probability of purchase as a function of recency, using the coefficients for recency, recency², and recency³ in Model 3. As expected, the shape of the graph...
is very similar to that shown in Figure 1, calculated from actual data. This says that the results in Figure 1 were not due to a confound with other variables.

[Figure 2 Goes Here]

Figure 3 illustrates the interaction between marketing and recency. Recall from Table 3 that the interaction between recency and marketing was statistically significant and positive for direct mail. This means that direct mail response becomes more pronounced for higher recency states. This is illustrated in Figure 3, which compares response to email and direct mail for customers in different recency states. When recency equals 1, the dotted line, representing direct mail response, is positive but has smaller slope than the solid email line. When recency equals 20, the slope is noticeably steeper, and steeper than the email line. In terms of the coefficients in Table 3, when recency = 20, the email response slope is $0.357 - 20 \times 0.003 = 0.300$, while the direct mail response slope is $0.105 + 20 \times 0.020 = 0.500$. When recency equals 1, the response slope for email is $0.357 - 1 \times 0.003 = 0.354$, while for direct mail it is $0.105 + 1 \times 0.020 = 0.125$.

[Figure 3 Goes Here]

Figure 4 demonstrates saturation effects. These are driven by the negative email × lagged_email and dmail × lagged_dmail coefficients in Table 3. This means that the slope of purchase probability as a function of marketing decreases to the extent that a large level of marketing has been employed in the previous period. These saturation effects are similar to those found by Ansari et al. (2008) as well as Dreze et al. (2009) and represent an additional “cost” to marketing above and beyond distribution or discount costs. For both email and direct mail, the slope of purchase probability versus email/direct mail gets smaller but still positive when lagged email/direct mail = 0 or 1. But at lagged email/direct mail = 2 or 3, the slope actually becomes negative, suggesting supersaturation. This effect is particularly strong for
email. Apparently, when the company is emailing heavily to the customer, the customer becomes so frustrated with the company, or so overloaded with emails, that continued high levels of emailing actually backfire, making the customer less likely to purchase.

In summary, our logistic regression contains (1) a pronounced impact of recency, (2) significant current period, carryover, and saturation effects, and (3) interactions between marketing and recency. We now can appreciate the complexity of the task at hand. For example, the saturation effects present for both email and direct mail suggest that “pulsing” may be optimal, in that if we use a lot of marketing when the customer is in state S, we will be less apt to use marketing in states $S'$ and $S''$, the states that follow depending on whether the customer buys or not. However, direct mail has a particularly high carryover effect, plus it interacts positively with recency, so this may bode for steadily increasing levels of direct mail. There is also the main effect of recency to contend with, whereby baseline purchase probability is decreasing over time, meaning the level from which we attempt to raise purchase probability is becoming lower and lower (the recency trap). How these factors balance out to achieve the optimal policy will be demonstrated in the next section.

V.2 Optimization

V.2.1 State variables. We now use equation (4) to calculate the optimal policy function, $D(|S)$. We will use the method of “Value Iteration” (Judd 1998, pp. 412-413). Value iteration solves for the optimal stationary policy, i.e., the decisions will not depend on the time period per se, but only on the state variables that describe the customer. In our application, we have four state variables:
• Recency: If the customer is in recency state \( r \), the customer moves to recency state 1 if he or she purchases, or state \( \min\{r + 1, \text{Maxrecency}\} \) if he or she does not purchase. That is, if the customer has not purchased for five months and does not purchase in the current period, the customer now has not purchased in six months so is in recency state 6. While in theory, recency could increase indefinitely, for tractability and to ensure not working outside the range of the data, we put a cap on recency, called “Maxrecency”. We use Maxrecency = 20. Once the customer gets to recency state 20 and does not purchase, we consider the customer still in recency state 20. As in Table 1, t state 20 is not absorbing. The customer in that state may still purchase and move to recency state 1.

• Month: There are 12 months in the year. Table 3 shows that month influences purchase probability, and obviously changes from period to period.

• Lagged_Email: Table 3 shows carryover effects of email, and this variable will change period to period, depending on the level of emailing in the previous period. Therefore it is a state variable. Technically, it is a continuous state variable. However, states need to be defined discretely in order to solve the dynamic program. The maximum value for monthly email was close to 3; the minimum was obviously zero. We divided this variable into 30 equal increments (i.e., 0, 0.1, 0.2, etc. up to 3.0). This means that in any period, the customer could be in one of 30 possible lagged_email states.

• Lagged_Dmail: Table 3 shows carryover effects of direct mail, and as for Lagged_Email, we create 30 lagged_direct_mail states.

In summary, recency, month, and two lagged marketing variables describe the customer at any point in time. There are 20 recency states, 12 months, and two 30-level lagged marketing states, so the total number of states is \( 20 \times 12 \times 30 \times 30 = 216,000 \). This means we have 216,000
value functions, each representing the subsequent lifetime value of a customer who starts in state
S and is marketed to optimally according to the decision rule, D|S, derived from value iteration.

V.2.2 Value iteration method. Value iteration is an iterative approach whereby each of
the 216,000 value functions is approximated at each iteration. The procedure terminates when
each of the value functions changes by some small tolerance level, in this case $0.00001. The
procedure was programmed in C and required approximately 15 hours to converge. The
program is available from the authors. The procedure is actually quite simple and can be
outlined as follows:

1. Let $V(S)^w$ = the value function at iteration $w$ for a customer who is in state $S$.
   Eventually, this quantity will converge to the estimated value function $V(S)$.

2. We have two decision variables – email effort and direct mail effort. These each range
   from 0 to 3 (as discussed earlier, 3 is the maximum value of these variables in the data,
   and we wanted to stay within the range of the data). We divide each of these variables
   into 30 increments of 0.1. Therefore, the policy function D|S has can be thought of as a
   vector of two values for each state, consisting of one of 30 possible email decisions and
   one of 30 possible direct mail decisions.

3. Find initial values, $V(S)^0$, for the value function for each state. We did this by
   computing the short-term profit function, $\pi_{it}(S)$ for each of the 900 possible email/direct
   mail combinations, for each state, and taking the maximum as the initial value, $V(S)^0$.

4. Compute the maximum value of $V(S)^w = \max_D\{\pi(S) + \delta(ProbPurch \times V(S)^{w-1} +
   (1 - ProbPurch) \times V(S''')^{w-1})\}$ by trying all 900 possible combinations of email and
direct mail. Call the combination that produces this maximum $D^*(email, dmail)$. 
5. Test whether $V(S)^w - V(S)^{w-1} < 0.00001$ for each state $S$. If this condition holds, the process has converged and the current value of $V(S)^w$ is the value function for customers in state $S$, and the most recently used combination of email and dmail is the optimal policy function $D^*(email, dmail)$. If the condition does not hold for all states $S$, set $V(S)^{w-1} = V(S)^w$ and proceed back to step 4 for another iteration. Note that we have updated the value function because now in step 4, the new value functions we created on the left side of the equation will be on the right side of the equation.

The process required 1600 iterations to converge and approximately 15 hours of computing time.

### V.2.3 Profit function

Implementing the algorithm described above requires specification of the current period profit function. For our application, that function was expressed as follows:

\[
E[\pi_{it}(S)] = M \times ProbPurch(Email_{it}, Dmail_{it}, S) - DISTE \times Email_{it} - DISTD \times Dmail_{it} - DISCE \times ProbPurch(Email_{it}, Dmail_{it}, S) \times \min\{Email_{it}, 1\} - DISCD \times ProbPurch(D, S) \times \min\{Dmail_{it}, 1\}
\]

where:

- $\pi_{it}(S)$ = Net Profit contributed by customer $i$ in time $t$.
- $M$ = Gross profit contribution if customer makes a purchase.
- $Email_{it}$ = Level of emailing targeted at customer $i$ in time $t$.
- $Dmail_{it}$ = Level of direct mailing targeted at customer $i$ in time $t$. 
• $\text{ProbPurch}(Email_{it}, Dmail_{it}, S) = $ Probability customer $i$ purchases in time $t$ if customer is in state $S$ at that time and receives marketing equal to $Email_{it}$ and $Dmail_{it}$.

• $DISTE = $ Distribution cost per unit of Emailing effort.

• $DISTD = $ Distribution cost per unit of direct mailing effort.

• $DISCE = $ Average price discount when customer buys under an Email promotion.

• $DISCD = $ Average price discount when customer buys under a direct mail promotion.

The first term in equation (5) represents the expected positive contribution, equal to the average contribution ($M$) multiplied times the probability the customer makes a purchase. This probability depends on what state the customer is in, plus the level of emailing and direct mailing the customer receives. Information provided by the firm in this application suggested that $M = $71.93. Purchase probability was calculated using the estimated logistic regression model. The next two terms represent distribution costs, e.g., mailing a direct mail piece. Information provided by the firm was that $DISTE = $0 and $DISTD = $0.40. The final two terms reflect the expected price discount when the customer responds to an email or a direct mail. Calculations using customer purchase records suggested $DISCE = $10.70 while $DISCD = $6.16. The use of the “min” function in the final two terms reflects the empirical fact that no customers purchased more than once a month in the data. For example, the “min” function assures that the customer never can gain more than $DISCD$ when purchasing under a direct mail promotion, and if $DISCD < 1$, this means that the direct mail promotion lasted less than a month, so we assume the customer’s chance of receiving the discount was proportional to how long the direct mail promotion was in effect.

V.2.4 Optimization results. Figure 5 shows the optimal email and direct mail policies as functions of recency, and compares them with the company’s current policy. Recall we have
216,000 possible states. To assess the relationship between email/direct mail policies and recency, we conduct four regressions: one for each of the two marketing instruments (email/direct mail) and for both the optimal and current policies. For the optimal policy, the dependent variable is the optimal level of email/direct mail. For the current policy, we use the current data, at the customer/time level, and use the actual level of email/direct mail used. The explanatory variables in both cases are the state variables: recency (19 dummy variables), month (11 dummies), lagged email, and lagged direct mail (scaled from 0 to 3 in increments of 0.1). We do this for both the optimal policy and for the raw data. Figure 5 displays the recency state dummies.

Figure 5 leads to the following conclusions:

- Optimal levels of direct mail are higher than optimal levels of email. This makes sense in that (1) emailing has higher saturation effects (Figure 4), (2) direct mail has much stronger carryover effects, and (3) emailing yields larger discounts off regular price.

- Optimal levels of both email and direct mail generally increase with recency. This reinforces the theme that marketing should do its best to arrest the progression of the customer to higher recency states (Table 3, Figures 1 and 2).

- We see some signs of “pulsing” in the email policy. For recency levels 13-18, high levels of email when the customer is in state $r$ are followed by low levels of email if the customer does not purchase and therefore progresses to state $r+1$. This is probably due to the saturation effects shown in Figure 4. If the customer does not purchase and moves to state $r+1$, it becomes unprofitable to follow up with additional emailing which will just
be ignored due to saturation. It is better to wait to see if the customer drifts further, to state \( r+2 \), and then expand emailing when the customer is more receptive to it.

- When the customer is in state 20, direct mailing falls off while emailing increases. We interpret this to be the result of strong carryover for direct mail. Part of the attraction in direct mail is the carryover effect, which means that high levels of direct mail ensure the customer will be more likely to purchase in the next period even if the customer does not purchase and drifts to a higher recency state. However, when the customer gets to state 20, that additional insurance benefit no longer is in play, since if the customer does not purchase when in state 20, he or she stays in state 20.

- The optimal policy suggests the firm should be spending more on both email and direct mail, compared to their current policy.

- The company’s current policy is not to target based on recency (the relationships between recency and email/direct mail distribution are basically flat).

Figure 6 shows a revealing picture of what would be gained by following the optimal policy. It displays the lifetime value of the customer, given various recency states. For the optimal policy, these are merely the average value functions after controlling for our other state variables (see footnote to Figure 6). For the current policy, these values were calculated using simulation of the current policy over the lifetime of the customer. As can be seen, CLV decreases markedly as a function of recency – even with the optimal policy, a high recency customer just is not as profitable in the long run as a low recency customer. But the difference between current practice CLV and optimal CLV is clear, on the order of $150-$200 per customer. The results for high recency customers are particularly salient: these customers are
currently virtually worthless to the firm, but our optimization suggests that with proper marketing, they would be worth roughly $150-$175. Together with Figure 6, this suggests the firm is now giving up too soon on these customers.

[Figure 6 Goes Here]

VI. SUMMARY AND AVENUES FOR FUTURE RESEARCH

We have developed and demonstrated an approach to deriving optimal policies for managing customer value, guided by the migration model of customer lifetime value. The approach consists of three key elements: (1) focus on customer recency and the related customer migration model of CLV, (2) estimation of a customer-level marketing response function that includes several recency phenomena as well as marketing carryover and saturation, (3) use of a dynamic program utilizes the estimated response function to derive a customer-specific optimal policy for utilizing two marketing tools – in this case, email and direct marketing.

The method integrates the purchase probabilities that drive the migration model of CLV with optimizing CLV. The key is that Step 2 estimates these purchase probabilities as a function of marketing; this in turn means that the optimization in Step 3 can find the targeted marketing strategy that maximizes CLV within the migration model framework. Equation (4) shows this analytically.

Our paper can be seen as an advocacy for recency and the migration model of CLV, but recency is not the only phenomenon to be factored into optimal customer-targeted marketing programs. Marketing carryover and saturation play a crucial role. The need to keep track of these variables increased the complexity of the optimization – as not only recency but recent
marketing efforts also became state variables – but our application shows that incorporating these factors is feasible.

Our application serves as an interesting case study. This company was truly falling victim to the recency trap, as shown in Table 1. Their current marketing program was underfunded and did not expend the *additional* efforts needed as customers moved to higher recency states. Our prescribed policy called for increasing efforts as customers drifted away, but this is clearly a function of the particular response function and costs involved. One could imagine, for example, that higher recency groups might become significantly *less* responsive to marketing, whereby beyond a point, when recency becomes just too high, it no longer becomes worth it and the firm lets the customer drift away.

The key implications of our work for researchers are: (1) Recency and the migration model of customer lifetime value are key tools that merit increased attention in customer management models. (2) Recency and marketing response can interact, reinforcing Khan, Lewis, and Singh (2009). This needs to be thoroughly incorporated in order to prescribe the optimal marketing policy. (3) In fact, several response phenomena – interactions, carryover, and saturation all need to be factored into an optimal targeted marketing policy. (4) Value iteration is a valuable and practical tool for deriving infinite horizon policies.

The key implications of our work for managers are: (1) Recency and migration model diagnostics such as shown in Table 1 and Figure 1 should constantly be monitored by firms. It is possible that in a given circumstance, companies will not be at the mercy of the recency trap. But the accumulated evidence, including this paper, suggests that this is a key phenomenon. (2) The tools to derive an optimal CLV marketing policy are feasible for practical implementation. The driving methods used in this work were logistic regression – very well known to companies
– and the value iteration solution of a dynamic program, an iterative method that can be easily programmed. (3) The optimization derives specific customer recommendations for a targeted one-to-one marketing effort. But the approach also contributes important general strategic guidance – in this case (i) increase marketing efforts, and (ii) recency is a crucial criterion for targeting marketing efforts. (4) Optimization can have a large impact on CLV. Our results suggest that in this case, customer value would increase by hundreds of dollars, per customer, and customers who heretofore were worth virtually $0 to the company could be converted to customers worth roughly $150 on average. (5) Finally, this work reinforces the emerging view that customer lifetime value is something to be managed, not merely measured. Certainly, CLV is valuable in a measurement in itself, for example in managing customer acquisition. But a key challenge is to derive a set of marketing policies that will maximize CLV.

While we believe this paper has covered and addressed several key issues in managing customer lifetime value, there are of course many challenges ahead. These include: (1) Models of purchase quantity could be included in the approach. In our case, we used an average customer contribution in our profit function. However, purchase quantity could be influenced by marketing, and in fact previous purchase quantity could serve as a state variable for the optimization. In our “defense”, previous research has indeed found that purchase incidence is more malleable to marketing efforts than purchase quantity in a CRM setting (e.g., Ansari, Mela, and Neslin 2008), but this still would be an interesting area of future work. (2) While the logistic regression model includes an implicit interaction between email and direct mail, we did not model this explicitly, in order to keep the model as simple as possible. This would be quite feasible within our framework because it would not expand the state space required to solve the dynamic program. (3) Our work is highly suggestive of the gains to be had by managing
customer recency effectively. However, the efficacy of the approach should be demonstrated in a field test, which would provide convincing evidence. We indeed encourage future researchers to undertake these important improvements over our current paper.
REFERENCES


Table 1
The Customer Migration Model, Decreasing Purchase Probabilities as Function of Recency, and The Recency Trap

<table>
<thead>
<tr>
<th>Prob(Purchase</th>
<th>Recency) =</th>
<th>0.231</th>
<th>0.144</th>
<th>0.094</th>
<th>0.059</th>
<th>0.046</th>
<th>0.036</th>
<th>0.025</th>
<th>0.026</th>
<th>0.023</th>
<th>0.018</th>
<th>0.019</th>
<th>0.016</th>
<th>0.012</th>
<th>0.007</th>
<th>0.005</th>
<th>0.010</th>
<th>0.006</th>
<th>0.007</th>
<th>0.006</th>
<th>0.005</th>
</tr>
</thead>
<tbody>
<tr>
<td>Recency State (Periods since last purchase):</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
<td>5</td>
<td>6</td>
<td>7</td>
<td>8</td>
<td>9</td>
<td>10</td>
<td>11</td>
<td>12</td>
<td>13</td>
<td>14</td>
<td>15</td>
<td>16</td>
<td>17</td>
<td>18</td>
<td>19</td>
<td>≥ 20</td>
<td></td>
</tr>
<tr>
<td>Period: 1</td>
<td>1.00</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>0.23</td>
<td>0.77</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>0.16</td>
<td>0.18</td>
<td>0.66</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>4</td>
<td>0.12</td>
<td>0.12</td>
<td>0.16</td>
<td>0.60</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>5</td>
<td>0.10</td>
<td>0.10</td>
<td>0.11</td>
<td>0.14</td>
<td>0.56</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>6</td>
<td>0.08</td>
<td>0.07</td>
<td>0.08</td>
<td>0.10</td>
<td>0.13</td>
<td>0.54</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>7</td>
<td>0.07</td>
<td>0.06</td>
<td>0.06</td>
<td>0.07</td>
<td>0.09</td>
<td>0.12</td>
<td>0.52</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>8</td>
<td>0.06</td>
<td>0.05</td>
<td>0.05</td>
<td>0.06</td>
<td>0.07</td>
<td>0.12</td>
<td>0.51</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>9</td>
<td>0.05</td>
<td>0.04</td>
<td>0.05</td>
<td>0.05</td>
<td>0.07</td>
<td>0.08</td>
<td>0.12</td>
<td>0.49</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>10</td>
<td>0.05</td>
<td>0.04</td>
<td>0.04</td>
<td>0.04</td>
<td>0.05</td>
<td>0.04</td>
<td>0.06</td>
<td>0.08</td>
<td>0.11</td>
<td>0.48</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>11</td>
<td>0.04</td>
<td>0.04</td>
<td>0.03</td>
<td>0.04</td>
<td>0.04</td>
<td>0.04</td>
<td>0.05</td>
<td>0.06</td>
<td>0.08</td>
<td>0.11</td>
<td>0.47</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>12</td>
<td>0.04</td>
<td>0.03</td>
<td>0.03</td>
<td>0.03</td>
<td>0.04</td>
<td>0.04</td>
<td>0.05</td>
<td>0.06</td>
<td>0.08</td>
<td>0.11</td>
<td>0.46</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>13</td>
<td>0.04</td>
<td>0.03</td>
<td>0.03</td>
<td>0.03</td>
<td>0.04</td>
<td>0.04</td>
<td>0.04</td>
<td>0.05</td>
<td>0.06</td>
<td>0.08</td>
<td>0.11</td>
<td>0.46</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>14</td>
<td>0.03</td>
<td>0.03</td>
<td>0.03</td>
<td>0.03</td>
<td>0.03</td>
<td>0.03</td>
<td>0.03</td>
<td>0.04</td>
<td>0.05</td>
<td>0.06</td>
<td>0.07</td>
<td>0.11</td>
<td>0.45</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>15</td>
<td>0.03</td>
<td>0.02</td>
<td>0.02</td>
<td>0.02</td>
<td>0.02</td>
<td>0.03</td>
<td>0.03</td>
<td>0.03</td>
<td>0.04</td>
<td>0.05</td>
<td>0.06</td>
<td>0.07</td>
<td>0.10</td>
<td>0.45</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>16</td>
<td>0.02</td>
<td>0.02</td>
<td>0.02</td>
<td>0.02</td>
<td>0.02</td>
<td>0.02</td>
<td>0.02</td>
<td>0.03</td>
<td>0.03</td>
<td>0.03</td>
<td>0.04</td>
<td>0.04</td>
<td>0.06</td>
<td>0.07</td>
<td>0.10</td>
<td>0.45</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>17</td>
<td>0.02</td>
<td>0.02</td>
<td>0.02</td>
<td>0.02</td>
<td>0.02</td>
<td>0.02</td>
<td>0.02</td>
<td>0.03</td>
<td>0.03</td>
<td>0.03</td>
<td>0.04</td>
<td>0.04</td>
<td>0.06</td>
<td>0.07</td>
<td>0.10</td>
<td>0.44</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>18</td>
<td>0.02</td>
<td>0.02</td>
<td>0.02</td>
<td>0.02</td>
<td>0.02</td>
<td>0.02</td>
<td>0.02</td>
<td>0.03</td>
<td>0.04</td>
<td>0.04</td>
<td>0.06</td>
<td>0.07</td>
<td>0.10</td>
<td>0.44</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>19</td>
<td>0.02</td>
<td>0.02</td>
<td>0.02</td>
<td>0.02</td>
<td>0.02</td>
<td>0.02</td>
<td>0.02</td>
<td>0.02</td>
<td>0.03</td>
<td>0.04</td>
<td>0.06</td>
<td>0.07</td>
<td>0.10</td>
<td>0.44</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>20</td>
<td>0.02</td>
<td>0.02</td>
<td>0.02</td>
<td>0.02</td>
<td>0.02</td>
<td>0.02</td>
<td>0.02</td>
<td>0.02</td>
<td>0.02</td>
<td>0.03</td>
<td>0.04</td>
<td>0.06</td>
<td>0.07</td>
<td>0.10</td>
<td>0.43</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>21</td>
<td>0.02</td>
<td>0.01</td>
<td>0.01</td>
<td>0.01</td>
<td>0.01</td>
<td>0.01</td>
<td>0.01</td>
<td>0.02</td>
<td>0.02</td>
<td>0.02</td>
<td>0.02</td>
<td>0.02</td>
<td>0.02</td>
<td>0.03</td>
<td>0.04</td>
<td>0.05</td>
<td>0.07</td>
<td>0.53</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>22</td>
<td>0.02</td>
<td>0.01</td>
<td>0.01</td>
<td>0.01</td>
<td>0.01</td>
<td>0.01</td>
<td>0.01</td>
<td>0.02</td>
<td>0.02</td>
<td>0.02</td>
<td>0.02</td>
<td>0.02</td>
<td>0.02</td>
<td>0.03</td>
<td>0.03</td>
<td>0.04</td>
<td>0.05</td>
<td>0.60</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>23</td>
<td>0.02</td>
<td>0.01</td>
<td>0.01</td>
<td>0.01</td>
<td>0.01</td>
<td>0.01</td>
<td>0.01</td>
<td>0.02</td>
<td>0.02</td>
<td>0.02</td>
<td>0.02</td>
<td>0.02</td>
<td>0.02</td>
<td>0.03</td>
<td>0.03</td>
<td>0.04</td>
<td>0.04</td>
<td>0.65</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>24</td>
<td>0.01</td>
<td>0.01</td>
<td>0.01</td>
<td>0.01</td>
<td>0.01</td>
<td>0.01</td>
<td>0.01</td>
<td>0.02</td>
<td>0.02</td>
<td>0.02</td>
<td>0.02</td>
<td>0.02</td>
<td>0.02</td>
<td>0.03</td>
<td>0.03</td>
<td>0.04</td>
<td>0.69</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>25</td>
<td>0.01</td>
<td>0.01</td>
<td>0.01</td>
<td>0.01</td>
<td>0.01</td>
<td>0.01</td>
<td>0.01</td>
<td>0.01</td>
<td>0.02</td>
<td>0.02</td>
<td>0.02</td>
<td>0.02</td>
<td>0.02</td>
<td>0.03</td>
<td>0.03</td>
<td>0.04</td>
<td>0.72</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>...</td>
<td>...</td>
<td>...</td>
<td>...</td>
<td>...</td>
<td>...</td>
<td>...</td>
<td>...</td>
<td>...</td>
<td>...</td>
<td>...</td>
<td>...</td>
<td>...</td>
<td>...</td>
<td>...</td>
<td>...</td>
<td>...</td>
<td>...</td>
<td>...</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>50</td>
<td>0.01</td>
<td>0.01</td>
<td>0.01</td>
<td>0.01</td>
<td>0.01</td>
<td>0.01</td>
<td>0.01</td>
<td>0.01</td>
<td>0.01</td>
<td>0.01</td>
<td>0.01</td>
<td>0.01</td>
<td>0.01</td>
<td>0.01</td>
<td>0.01</td>
<td>0.01</td>
<td>0.01</td>
<td>0.01</td>
<td>0.89</td>
<td></td>
<td></td>
</tr>
<tr>
<td>51</td>
<td>0.01</td>
<td>0.01</td>
<td>0.01</td>
<td>0.01</td>
<td>0.01</td>
<td>0.01</td>
<td>0.01</td>
<td>0.01</td>
<td>0.01</td>
<td>0.01</td>
<td>0.01</td>
<td>0.01</td>
<td>0.01</td>
<td>0.01</td>
<td>0.01</td>
<td>0.01</td>
<td>0.01</td>
<td>0.01</td>
<td>0.89</td>
<td></td>
<td></td>
</tr>
<tr>
<td>52</td>
<td>0.01</td>
<td>0.01</td>
<td>0.01</td>
<td>0.01</td>
<td>0.01</td>
<td>0.01</td>
<td>0.01</td>
<td>0.01</td>
<td>0.01</td>
<td>0.01</td>
<td>0.01</td>
<td>0.01</td>
<td>0.01</td>
<td>0.01</td>
<td>0.01</td>
<td>0.01</td>
<td>0.01</td>
<td>0.01</td>
<td>0.89</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

* Recency state represents the number of periods since the previous purchase. The customer is acquired in period 1. Cell entries represent the probability the customer will be in each state in each time period. Recency column 1 represents the probability a customer will purchase in each period.
<table>
<thead>
<tr>
<th>Variable</th>
<th>Description</th>
<th>Mean</th>
<th>Std. Dev.</th>
<th>Min.</th>
<th>Max</th>
</tr>
</thead>
<tbody>
<tr>
<td>Recency_{ht}</td>
<td>Periods since last purchase by household ( h ), with “1” signifying the purchase was made in month ( t ).</td>
<td>7.659</td>
<td>5.621</td>
<td>1</td>
<td>20</td>
</tr>
<tr>
<td>First_Amt_{ht}</td>
<td>Amount spent on first purchase by household ( h ).</td>
<td>116.264</td>
<td>58.068</td>
<td>-110**</td>
<td>480</td>
</tr>
<tr>
<td>Email_{t}</td>
<td>Level of Email marketing activity in month ( t ), scaled so that one email campaign lasting one month would be scored as 1. The variable therefore can be interpreted as number of months worth of email campaigning in month ( t ).</td>
<td>0.672</td>
<td>0.689</td>
<td>0</td>
<td>2.548</td>
</tr>
<tr>
<td>Dmail</td>
<td>Level of direct marketing activity in month ( t ), scaled so that one direct mail campaign lasting one month would be scored as 1. The variable therefore can be interpreted as number of months worth of direct mail campaigning in month ( t ).</td>
<td>1.524</td>
<td>0.805</td>
<td>0</td>
<td>2.933</td>
</tr>
<tr>
<td>Purchase_{cht}</td>
<td>1 if household ( h ) purchased in month ( t ); 0 otherwise.</td>
<td>0.069</td>
<td>0.253</td>
<td>0</td>
<td>1</td>
</tr>
</tbody>
</table>

* Based on \( n = 61,952 \) household-week observations.
** There was one customer outlier with a negative value for First_Amt. The rest of the values were above zero. We decided to leave this customer in the data, although this had virtually no influence on the results.
<table>
<thead>
<tr>
<th>Variable</th>
<th>Base Model</th>
<th>Model 2</th>
<th>Model 3</th>
<th>Model 4</th>
<th>Model 5</th>
</tr>
</thead>
<tbody>
<tr>
<td>Intercept</td>
<td>-1.105</td>
<td>-1.774</td>
<td>-1.720</td>
<td>-1.626</td>
<td>-1.623</td>
</tr>
<tr>
<td>P-val</td>
<td>&lt;.001</td>
<td>&lt;.001</td>
<td>&lt;.001</td>
<td>&lt;.001</td>
<td>&lt;.001</td>
</tr>
<tr>
<td>Recency</td>
<td>-0.714</td>
<td>-0.715</td>
<td>-0.741</td>
<td>-0.790</td>
<td>-0.791</td>
</tr>
<tr>
<td>P-val</td>
<td>&lt;.001</td>
<td>&lt;.001</td>
<td>&lt;.001</td>
<td>&lt;.001</td>
<td>&lt;.001</td>
</tr>
<tr>
<td>Recency²</td>
<td>0.050</td>
<td>0.049</td>
<td>0.049</td>
<td>0.052</td>
<td>0.053</td>
</tr>
<tr>
<td>P-val</td>
<td>&lt;.001</td>
<td>&lt;.001</td>
<td>&lt;.001</td>
<td>&lt;.001</td>
<td>&lt;.001</td>
</tr>
<tr>
<td>Recency³</td>
<td>-0.00134</td>
<td>-0.00130</td>
<td>-0.00131</td>
<td>-0.00123</td>
<td>-0.00131</td>
</tr>
<tr>
<td>P-val</td>
<td>&lt;.001</td>
<td>&lt;.001</td>
<td>&lt;.001</td>
<td>&lt;.001</td>
<td>&lt;.001</td>
</tr>
<tr>
<td>Email</td>
<td>0.375</td>
<td>0.357</td>
<td>0.364</td>
<td>0.364</td>
<td>0.364</td>
</tr>
<tr>
<td>P-val</td>
<td>0.007</td>
<td>0.030</td>
<td>0.014</td>
<td>0.014</td>
<td>0.014</td>
</tr>
<tr>
<td>Lagged Email</td>
<td>0.095</td>
<td>0.092</td>
<td>0.086</td>
<td>0.086</td>
<td>0.086</td>
</tr>
<tr>
<td>P-val</td>
<td>0.441</td>
<td>0.455</td>
<td>0.487</td>
<td>0.487</td>
<td>0.487</td>
</tr>
<tr>
<td>Email×Lagged Email</td>
<td>-0.247</td>
<td>-0.242</td>
<td>-0.221</td>
<td>-0.221</td>
<td>-0.221</td>
</tr>
<tr>
<td>P-val</td>
<td>0.039</td>
<td>0.043</td>
<td>0.067</td>
<td>0.067</td>
<td>0.067</td>
</tr>
<tr>
<td>Email×Recency</td>
<td>-0.003</td>
<td>-0.026</td>
<td>-0.270</td>
<td>-0.270</td>
<td>-0.270</td>
</tr>
<tr>
<td>P-val</td>
<td>0.717</td>
<td>0.270</td>
<td>0.045</td>
<td>0.045</td>
<td>0.045</td>
</tr>
<tr>
<td>Email×Recency²</td>
<td>0.00198</td>
<td>0.00198</td>
<td>0.00198</td>
<td>0.00198</td>
<td>0.00198</td>
</tr>
<tr>
<td>P-val</td>
<td>0.206</td>
<td>0.206</td>
<td>0.016</td>
<td>0.016</td>
<td>0.016</td>
</tr>
<tr>
<td>Email×Recency³</td>
<td>-0.00056</td>
<td>-0.00056</td>
<td>-0.00056</td>
<td>-0.00056</td>
<td>-0.00056</td>
</tr>
<tr>
<td>P-val</td>
<td>0.091</td>
<td>0.091</td>
<td>0.091</td>
<td>0.091</td>
<td>0.091</td>
</tr>
<tr>
<td>Dmail</td>
<td>0.140</td>
<td>0.140</td>
<td>0.140</td>
<td>0.140</td>
<td>0.140</td>
</tr>
<tr>
<td>P-val</td>
<td>0.191</td>
<td>0.191</td>
<td>0.191</td>
<td>0.191</td>
<td>0.191</td>
</tr>
<tr>
<td>Lagged Dmail</td>
<td>0.375</td>
<td>0.375</td>
<td>0.375</td>
<td>0.375</td>
<td>0.375</td>
</tr>
<tr>
<td>P-val</td>
<td>0.001</td>
<td>0.001</td>
<td>0.001</td>
<td>0.001</td>
<td>0.001</td>
</tr>
<tr>
<td>Dmail×Lagged Dmail</td>
<td>0.090</td>
<td>0.116</td>
<td>0.116</td>
<td>0.116</td>
<td>0.116</td>
</tr>
<tr>
<td>P-val</td>
<td>0.001</td>
<td>0.001</td>
<td>0.001</td>
<td>0.001</td>
<td>0.001</td>
</tr>
<tr>
<td>Dmail×Recency</td>
<td>0.020</td>
<td>0.020</td>
<td>0.020</td>
<td>0.020</td>
<td>0.020</td>
</tr>
<tr>
<td>P-val</td>
<td>0.011</td>
<td>0.011</td>
<td>0.011</td>
<td>0.011</td>
<td>0.011</td>
</tr>
<tr>
<td>Dmail×Recency²</td>
<td>-0.004</td>
<td>-0.004</td>
<td>-0.004</td>
<td>-0.004</td>
<td>-0.004</td>
</tr>
<tr>
<td>P-val</td>
<td>0.016</td>
<td>0.016</td>
<td>0.016</td>
<td>0.016</td>
<td>0.016</td>
</tr>
<tr>
<td>Dmail×Recency³</td>
<td>-0.0003</td>
<td>-0.0003</td>
<td>-0.0003</td>
<td>-0.0003</td>
<td>-0.0003</td>
</tr>
<tr>
<td>P-val</td>
<td>0.291</td>
<td>0.291</td>
<td>0.291</td>
<td>0.291</td>
<td>0.291</td>
</tr>
<tr>
<td>First Amt</td>
<td>0.00349</td>
<td>0.00349</td>
<td>0.00349</td>
<td>0.00349</td>
<td>0.00349</td>
</tr>
<tr>
<td>P-val</td>
<td>&lt;.001</td>
<td>&lt;.001</td>
<td>&lt;.001</td>
<td>&lt;.001</td>
<td>&lt;.001</td>
</tr>
<tr>
<td>Jun</td>
<td>0.103</td>
<td>0.067</td>
<td>0.077</td>
<td>0.085</td>
<td>0.087</td>
</tr>
<tr>
<td>P-val</td>
<td>0.208</td>
<td>0.461</td>
<td>0.385</td>
<td>0.353</td>
<td>0.338</td>
</tr>
<tr>
<td>Jul</td>
<td>0.006</td>
<td>-0.195</td>
<td>-0.163</td>
<td>-0.146</td>
<td>-0.142</td>
</tr>
<tr>
<td>P-val</td>
<td>0.947</td>
<td>0.125</td>
<td>0.201</td>
<td>0.252</td>
<td>0.267</td>
</tr>
<tr>
<td>Aug</td>
<td>0.239</td>
<td>0.223</td>
<td>0.242</td>
<td>0.230</td>
<td>0.229</td>
</tr>
<tr>
<td>P-val</td>
<td>0.003</td>
<td>0.205</td>
<td>0.176</td>
<td>0.197</td>
<td>0.201</td>
</tr>
<tr>
<td>Sep</td>
<td>0.548</td>
<td>0.495</td>
<td>0.504</td>
<td>0.489</td>
<td>0.488</td>
</tr>
<tr>
<td>P-val</td>
<td>&lt;.001</td>
<td>0.002</td>
<td>0.002</td>
<td>0.003</td>
<td>0.003</td>
</tr>
<tr>
<td>Oct</td>
<td>0.183</td>
<td>0.183</td>
<td>0.183</td>
<td>0.183</td>
<td>0.183</td>
</tr>
<tr>
<td>P-val</td>
<td>0.021</td>
<td>0.021</td>
<td>0.021</td>
<td>0.021</td>
<td>0.021</td>
</tr>
<tr>
<td>Nov</td>
<td>-0.101</td>
<td>-0.167</td>
<td>-0.162</td>
<td>-0.169</td>
<td>-0.166</td>
</tr>
<tr>
<td>P-val</td>
<td>0.222</td>
<td>0.092</td>
<td>0.103</td>
<td>0.088</td>
<td>0.094</td>
</tr>
<tr>
<td>Dec</td>
<td>-0.120</td>
<td>-0.210</td>
<td>-0.191</td>
<td>-0.179</td>
<td>-0.184</td>
</tr>
<tr>
<td>P-val</td>
<td>0.188</td>
<td>0.055</td>
<td>0.080</td>
<td>0.102</td>
<td>0.094</td>
</tr>
<tr>
<td>Jan</td>
<td>-0.113</td>
<td>0.066</td>
<td>0.002</td>
<td>-0.007</td>
<td>-0.011</td>
</tr>
<tr>
<td>P-val</td>
<td>0.211</td>
<td>0.955</td>
<td>0.985</td>
<td>0.950</td>
<td>0.923</td>
</tr>
<tr>
<td>Feb</td>
<td>0.333</td>
<td>0.536</td>
<td>0.534</td>
<td>0.524</td>
<td>0.520</td>
</tr>
<tr>
<td>P-val</td>
<td>0.007</td>
<td>&lt;.001</td>
<td>&lt;.001</td>
<td>&lt;.001</td>
<td>&lt;.001</td>
</tr>
<tr>
<td>Mar</td>
<td>0.243</td>
<td>0.221</td>
<td>0.244</td>
<td>0.256</td>
<td>0.254</td>
</tr>
<tr>
<td>P-val</td>
<td>0.002</td>
<td>0.025</td>
<td>0.014</td>
<td>0.010</td>
<td>0.011</td>
</tr>
<tr>
<td>Apr</td>
<td>-0.093</td>
<td>0.140</td>
<td>0.130</td>
<td>0.124</td>
<td>0.122</td>
</tr>
<tr>
<td>P-val</td>
<td>0.263</td>
<td>0.230</td>
<td>0.269</td>
<td>0.291</td>
<td>0.301</td>
</tr>
<tr>
<td>N</td>
<td>61,952</td>
<td>61,952</td>
<td>61,952</td>
<td>61,952</td>
<td>61,952</td>
</tr>
<tr>
<td>-2log likelihood</td>
<td>25925.793</td>
<td>25887.240</td>
<td>25878.750</td>
<td>25872.857</td>
<td>25869.933</td>
</tr>
<tr>
<td>Incremental Log LL</td>
<td>38.553</td>
<td>8.490</td>
<td>5.893</td>
<td>2.924</td>
<td></td>
</tr>
<tr>
<td>Incremental P-value</td>
<td>&lt;.001</td>
<td>0.014</td>
<td>0.053</td>
<td>0.232</td>
<td></td>
</tr>
</tbody>
</table>
Figure 1
Purchase Frequency vs. Recency Calculated Directly from the Data *

* Descriptive statistics based on 61,952 customer/month observations.
Figure 2
Probability of Purchase as a Function of Recency Calculated from the Model

* Calculation assumes no marketing effort, i.e., Email and Dmail = 0, and lagged Email and Dmail = 0; Month = 0. The shape of the curve is unaffected by changes in these assumptions.
Figure 3
Probability of Purchase Response to Email and Dmail for Different Recency States

Probability of Purchase When Recency = 1

Probability of Purchase When Recency = 20
Figure 4
Probability of Purchase Response to Email and Dmail Depending on Previous Email and Dmail – Illustrating Saturation Effects

Saturation Effects for Email

Saturation Effects for Dmail
Figure 5
Optimal and Actual Email/Dmail Policies as Function of Recency

Email Policies

Dmail Policies
Figure 6
Customer Lifetime Value: Optimal vs. Current Policies

* Graph is based on regression of state-specific value functions vs. recency, month, and LastEmail/Dmail. Graphed numbers use month=0, LastEmail=0, and LastDmail=0 as base cases. Changing these bases would change the level of the graphs slightly but the general trends and difference between optimal and current policy would remain roughly the same.