Overcoming the ‘Recency Trap’ in Customer Relationship Management

Scott A. Neslin

Albert Wesley Frey Professor of Marketing

Tuck School of Business, 100 Tuck Drive, Dartmouth College, Hanover, NH, 03755

PH: 603-646-2841, FAX: 603-646-1308

Scott.Neslin@Dartmouth.edu

Gail Ayala Taylor

Visiting Associate Professor

Tuck School of Business, 100 Tuck Drive, Dartmouth College, Hanover, NH, 03755

PH: 603-646-0533, FAX: 603-646-1308

Gail.A.Taylor@Dartmouth.edu

Kimberly D. Grantham

Senior Lecturer

Terry College of Business, 310 Herty Drive, University of Georgia, Athens, GA, 30602

PH: 706-542-3766, FAX: 706-542-3835

kdg@UGA.edu

Kimberly R. McNeil

Associate Professor

School of Business and Economics, Quiester Craig Hall, North Carolina A&T State University,

Greensboro, NC, 27411

PH: 336-334-7656, FAX: 336-256-2275

krmcneil@NCAT.edu

December 14, 2011

Keywords: Customer Relationship Management, Customer Lifetime Value, Optimization, Migration Model, Customer Recency

Overcoming the ‘Recency Trap’ in Customer Relationship Management

ABSTRACT

Purchase likelihood typically declines as the length of time since the customer’s previous purchase (“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 even less likely they will purchase in the next period. Eventually the customer is lost to the firm. The goal is to target a firm’s marketing efforts keeping in mind the customer's recency state. We develop and illustrate a modeling approach to achieve this goal. This requires an empirical model that predicts purchase likelihood as a function of recency and marketing, and a dynamic optimization that prescribes the most profitable way to target customers. In our application we find that customers’ purchase likelihoods as well as response to marketing depend on recency. These results are used to show that the targeting of email and direct mail should depend on the customer's recency and that the optimal decision policy enables high recency customers, who currently are virtually worthless to the firm, to become profitable.

Overcoming the ‘Recency Trap’ in Customer Relationship Management

I. INTRODUCTION

Customer relationship management (CRM) places executives in the mindset that customers are “manageable strategic assets of the firm” (Reimann et al. 2010, p. 329). CRM provides firms with the opportunity to contact the right customer at the right time through the right marketing medium. The challenge however is in the *dynamics*. The marketing decisions we make today may affect what we wish to do tomorrow. For example, due to “saturation”, the customer who received a direct mail piece last week may be less receptive to a direct mail piece this week. This suggests that a firm must manage the timing of its direct mail efforts in order to get optimal results. In general, customers are constantly changing over time. Not taking into account these dynamics can result in mis-targeting and mis-timing of marketing actions.

An important dynamic is “recency” – how long it has been since the customer’s previous purchase. Recency, along with its cousins, frequency and monetary value, is one of the pillars of customer management (Hughes 1996; Blattberg et al. 2008). Recency has been found to be highly correlated with customer purchase, and directly related to customer lifetime value (CLV) through the “migration model” of CLV (Berger and Nasr 1998; Pfeifer and Carraway 2000; Blattberg et al. 2008). A common finding is that higher recency (longer time since previous purchase) is associated with lower purchase likelihood (Rhee and McIntyre 2008; 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, which in turn increases their recency, making them even less likely to purchase in the period after that, 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 how marketing efforts should be targeted to which customers at which time in order to effectively manage the recency trap. 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 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 et al. (2008) call “optimal contact models.” The theme of these models is that CLV is something to be *managed*, not merely *measured*. They rely on two key components (1) a customer response model, quantifying customers’ purchase likelihood as a function of marketing and in our case, a detailed array of recency effects, and (2) a dynamic optimization that prescribes the actions the firm should take at a given time for a given customer in order to maximize CLV. Our paper will demonstrate how an optimal contact model can be used to effectively manage the recency trap.

We apply our approach to a meal preparation service provider whose key marketing tools are email and direct mail promotions. Our customer marketing response model shows that recency is related negatively to purchase probability, 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 while accounting for customer migration between recency states. Our application suggests four key findings: (1) marketing efforts should increase as customer recency increases, whereas the firm's current policy does not target in this way, (2) the firm currently is underutilizing both email and direct mail, (3) the firm should allocate more budget to direct mail than email, and (4) we predict that implementation of our procedure would increase CLV by \$175-\$200, with particularly important increases for customers with high recency.

We proceed to review the literature in more detail. Then we discuss a detailed illustration of the recency trap, followed by 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

Several researchers have included recency in models that predict customer behavior. Bult and Wansbeek (1995), Bitran and Mondschein (1996), Fader et al. (2005), and Rhee and McIntyre (2008) 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 et al. 2008, p. 325).

Blattberg et al. (2008) 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 presumably result in lower purchase likelihood. For example, a customer may replace a television every five years so purchase likelihood might increase as a function of recency up to roughly five years. 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 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 incorporate the relationship between recency and purchase likelihood into calculations of customer lifetime value (Berger and Nasr 1998; Pfeifer and

Carraway 2000). This work viewed CLV as a Markov chain with recency serving as a “state variable”, i.e. a variable that changes over time and influences the customer’s likelihood of purchasing. Customers move 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 very end of period 2 (or 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 paper by Drèze and Bonfrer (2008) found that the *scheduling* of email solicitations could affect 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 purchase likelihood 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: the customer anticipates that no new information is provided in period $t+1$ because the customer has

recently received a marketing message from the company. Ansari et al. (2008) and Rhee and McIntyre (2008) find evidence of saturation.

An extreme form of saturation, “supersaturation”, has been conjectured (e.g., Leeflang et al. 2000, p. 68), whereby high levels of marketing in previous periods result in current period marketing actually *decreasing* purchase likelihood. This could be due to customer irritation (Van Diepen et al. 2009) or information overload – after a surfeit of emails, additional emails merely encourage the customer to collect them in his/her inbox and ignore them all. As a result, the email => purchase relationship becomes negative. Van Diepen et al. (2009) 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.

II.3 Optimal Contact Models

One of the most exciting areas of CRM research is optimal contact models (Blattberg et al. 2008). Optimal contact models determine what marketing efforts should be expended on which customers at what time. 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 et al. (2000), Elsner et al. (2003, 2004), Rust and Verhoef (2005), Simester et al. (2006), Venkatesan et al. (2007) and Khan et al. (2009).

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), Venkatesan et al. (2007), and Khan et al. (2009) focus on multiple marketing activities. Rust and Verhoef consider direct mail and a customer relationship magazine; Venkatesan et al. consider sales calls and direct marketing; 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 relevant.

As noted earlier, 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. 2009), RFM categorizations (Bitran and Mondschein 1996), and decision trees (Simester et al. 2006). 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 her or him in the future. Dynamic programming methods include infinite horizon models (Simester et al.), rolling horizon models (Neslin et al. 2009), and finite horizon models (Khan et al.). Khan et al. note these approaches each have advantages and disadvantages. Finite horizon optimization can run into end-game issues, distorting marketing efforts 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 Contribution of This Paper

In summary, previous research suggests there is a recency trap. The contribution of this paper is to demonstrate how to target the right marketing to the right customers at the right time

to maximize customer lifetime value in this environment. Our overall approach is to focus on this managerial issue, and develop an optimal contact model that is “complete on the important issues” (Little 1970), yet relatively simple and managerially relevant. Our paper is unique in its focus on the recency trap and how to manage it. This is manifested in our customer marketing response model through the inclusion of a main effect of recency, interactions between recency and marketing, and carryover and saturation effects of marketing. We will show how our dynamic optimization explicitly takes into account the customer’s current recency state, as well as her or his future recency state depending on which marketing action is taken.

Our paper extends important research in this area. 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. Overall, because of our focus on the recency trap, we model recency, carryover, and saturation effects more thoroughly.

III. MODELING FRAMEWORK

Our modeling framework consists of three elements: (1) quantification of the role of recency in determining customer purchase, (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—recency is negatively associated with purchase likelihood. Figure 1 is calculated using actual raw data from our application. 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 last purchased five months ago (recency state 5) have only a 4.6% chance of purchasing.

[Figure 1 Goes Here]

The ramification 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 (see Blattberg et al. 2008). For example, by the end of period 7, there is a 9% chance that a customer acquired in period 1 will be in recency state 5, i.e., meaning that the customer's last purchase was five periods ago. The "Recency 1" column in Table 1 is most crucial, because it shows the probability the customer purchases each period after acquisition. The numbers in the top row of Table 1 govern these calculations and are taken from the probabilities shown in Figure 1. They represent the conditional probability the customer will 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 Goes Here]

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 ≥ 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 hold of the customer.

Figure 1 and Table 1 pose the managerial problem in vivid terms – devise a profitable 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*.

III.2 Logistic Response Model of Purchase Likelihood

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 is:

- 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:

- Email_t and Dmail_t – the marketing efforts expended by the firm in period t , in our case, either email or direct mail offers.
- Recency_{it} – the recency state of customer i in period t , i.e., how long it has been since the customer's previous purchase.
- Recency_{it}^2 and Recency_{it}^3 – these variables capture the possibility that the relationship between recency and purchase is non-linear, beyond the inherent nonlinearity 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.

- $Email_t \times Recency_{it}^2$, $Email_t \times Recency_{it}^3$, $Dmail_t \times Recency_{it}^2$, and $Dmail_t \times Recency_{it}^3$ – 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. That is, 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.
- $First_Amt_i$ – 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 et al. 2007). Given Fader et al.'s results (pp. 61 (Figure 1) and 64), this variable is associated with both frequency and monetary value (the F and M in RFM).

Collecting these variables (plus a column of ones for the intercept) into a $k + 1$ vector X_{it} , where k is the number of explanatory variables described above, yields the following logistic regression model:

$$Prob(Purchase_{it} = 1) = \frac{1}{1 + \exp^{-X_{it}\beta}} \quad (1)$$

where β is a $k + 1$ vector representing the impact of each of the k explanatory variables on probability of purchase.

III.3 Optimization

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

$$CLV_i = \sum_{t=0}^{\infty} \pi_{it}(D|S) \times \delta^t \quad (2)$$

where:

- $\pi_{it}(D|S)$ = Profit contributed by customer i in period t , given the customer is in state S in that period and marketing decision D 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’s state are those that affect current profitability and change over time. In our case, recency, previous email/direct mail efforts, and month are the state variables. The notation “ $(D|S)$ ” should be read as “the marketing decision we make for customers who are in state S .”
- δ = discount factor, e.g., 0.995 on a monthly basis means that profits achieved one year from the present are worth 94% (0.995^{12}) of what they are worth today.

Equation (2) emphasizes that CLV is not a static number – it is an objective to be managed through marketing efforts. These efforts should depend on the recency state of the customer. 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 next period. For example, if the customer has recently been saturated with emails, it may be optimal not to market to the customer in the current period, but instead to wait until the next period when the customer will pay attention to a new email. This must be balanced against the lower “baseline” purchase probability that results from letting the customer lapse to a higher recency state in that next period.

We will use dynamic programming to determine the optimal policy (D|S). This requires computation of the value function, $V_{it}(S)$. The value function has an intuitive interpretation in our context – it represents the lifetime value we expect from a customer i who is in state S at time t , given that we use the *optimal* policy (D|S). The value function in period t can be expressed as the expected profit we obtain from 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 (Judd 1998):

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

Equation (3) presents the customer management viewpoint that we should maximize current period expected profits, but do so in light of the future profits we expect to result from actions we take in the current period. Note that in the current period the customer is in state S , but depending on the actions we take, the customer will be in a different state, S' , in the next period.

In our case, the expected future profits, $E(V_{it+1}(S'))$, take on a simple form, because the customer either will purchase or not purchase. In particular:

$$E(V_{it+1}(S')) = \delta [ProbPurch(S)_t \times E(V_{it+1}(S' = 1)) + (1 - ProbPurch(S)_t) \times E(V_{it+1}(S' = S + 1))] \quad (4)$$

$ProbPurch(S)_t$ will be calculated using our logistic purchase model and will depend on the customer's current state as well as the marketing actions taken. The expected future value functions depend on which state, S' , the customer transitions to next period. In our case, S' can

either equal 1 or $S+1$. For example, consider the case of the customer who currently is in state $S=5$. If the customer purchases in period t , the customer will be in recency state 1 in period $t+1$, so $S'=1$. However, if the customer doesn't purchase in period t , the customer shifts to recency state $S'=6$, because now it has been six periods since the customer purchased.

Equations (3) and (4) show that we take the action that maximizes current period profit *plus* expected future profits. Future profits depend on whether the customer buys (and migrates to recency state 1), which happens with probability $ProbPurch(S)_t$, or does not buy (and migrates to recency state $S+1$), which happens with probability $(1 - ProbPurch(S)_t)$. The model manages the recency trap by explicitly considering what actions should be taken now, taking into account what happens if the customer does not purchase and slides to a higher recency state.¹

We have not yet specified the profit function $E(\pi_{it}(D|S))$. This function involves application-specific costs, etc., and so we describe it fully in Section V (Application).

IV. DATA

The data for our application are 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 pre-assembled meals they wish to pick up. Ordering is primarily done online, allowing for the easy collection of customer-level data. 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). Customers are acquired

¹ Note that if as in Table 1, the highest recency state is " ≥ 20 ", once ≥ 20 becomes that customer's current state and he or she does not buy, he or she "migrates" to state ≥ 20 , since we are collapsing states 20, 21, 22, etc. into one state, ≥ 20 . I.e., $S+1$ is state ≥ 20 if the current state is ≥ 20 .

at different times; 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 customer response model.

The chief marketing instruments used by the firm were email and direct mail promotions. These promotions 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. We therefore 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 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 these communications. The average email variable 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, 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 overlapping emails or direct mail in a given month. Table 2 describes these and other variables used in the model.

[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, and particularly the role of recency. 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. The interaction is not significant for email, but is significant (p-value = 0.011) for direct mail. The positive sign means that as customers transition 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.

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. From Table 3, the interaction between recency and marketing is 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 direct mail slope is noticeably 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 $\times$ lagged_email and dmail $\times$ lagged_dmail coefficients in Table 3. This means that the slope of purchase probability as a function of marketing decreases if 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 important “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 = 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, 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.

[Figure 4 Goes Here]

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 employ a lot of marketing when the customer is in state S , we will be less apt to use marketing in state S' , the state that follows depending on whether the customer buys or not. However, direct mail has particularly high carryover, plus it interacts positively with recency, so this may bode for steadily increasing levels of direct mail. Of course there is also the negative main effect of recency to contend with, i.e., 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 use equations (3) and (4) to calculate the optimal policy function, $(D|S)$. We use the method of “Value Iteration” (Judd 1998, pp. 412-413). Value

iteration solves for the optimal policy that 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 $\text{Maxrecency} = 20$. Once the customer gets to recency state 20 and does not purchase, we consider the customer still in recency state 20 (state ≥ 20 in terms of Table 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 over time, 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 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. Each of the 216,000 value functions is approximated at each iteration. The procedure terminates when each value function changes by some small tolerance level, in this case we used \$0.00001. The procedure was programmed in C and required approximately 1600 iterations and 15 hours to converge. The program is available from the authors. The procedure is quite straightforward and is outlined in the Appendix.

V.2.3 Profit function. Implementing the algorithm above requires specification of the current period profit function. For our application that function is expressed as:

$$\begin{aligned}
 E[\pi_{it}(S)] = & M \times ProbPurch(Email_{it}, Dmail_{it}, S) - DISTE \times Email_{it} - DISTD \times \\
 & Dmail_{it} - DISCE \times ProbPurch(Email_{it}, Dmail_{it}, S) \times \min\{Email_{it}, 1\} - \\
 & DISCD \times ProbPurch(Email_{it}, Dmail_{it}, S) \times \min\{Dmail_{it}, 1\}
 \end{aligned} \tag{5}$$

where:

- $\pi_{it}(S)$ = Net profit contributed by customer i in time t .
- M = Gross profit contribution if customer i 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 .

- $ProbPurch(Email_{it}, Dmail_{it}, S) =$ Probability customer i purchases in time t if customer is in state S at that time and receives marketing equal to $Email_{it}$ and $Dmail_{it}$.²
- $DISTE =$ 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 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. This assures that the customer never can gain more than $DISCD$ when purchasing under a direct mail promotion.

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

² Note that to calculate the purchase likelihood, we need the value of First_Amt, which varies across customers. We used the average value, \$116, in the optimization, so the results pertain to the optimal policy for the average customer.

(email/direct mail) and one 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 by the meal preparation establishment. 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). Figure 5 displays the estimated recency state dummies.

[Figure 5 Goes Here]

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 and so is more costly.
- 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 optimal 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. One might assert it is obvious to increase email because email has zero distribution cost. But email has two costs: First is the discount it advertises, which is expended on baseline sales as well as any incremental sales stimulated by the email. Second is the saturation effects of email (Figure 4) that show that too much emailing in period t decreases its effectiveness in period $t+1$. So the cost of emailing in period t is that it makes emails less effective in period $t+1$.
- The company's current policy is not to target based on recency (the relationships between recency and email/direct mail distribution are basically flat).³

Figure 6 shows a revealing picture of what would be gained by following the optimal policy. It displays CLV as a function of recency. For the optimal policy, these are 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 a simulation of the current policy over the

³ In fact, discussion with the firm's management suggested that the company was not currently targeting email or direct mail efforts in any way, i.e., they were not using previous purchase, etc., to target marketing. If they had been, this would have created an endogeneity that we would have had to handle in our estimation of the logistic customer response function (see Rhee and McIntyre 2008).

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 is not as profitable 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 5, 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 managing the recency trap that maximizes 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 that 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 optimization explicitly accounts for the recency trap by determining what marketing decision to make given the customer's current recency state, taking into account the ramifications for which recency state the customer is likely to transition to next period. This is shown explicitly in equations (3) and (4).

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 calls for increasing efforts as customers drift away. Note this is a function of the particular response function we estimated. One could imagine, for example, that higher recency groups might become significantly *less* responsive to marketing, whereby beyond a point, when recency becomes too high, it no longer becomes worth it and the firm lets the customer drift away.

The implications of our work for researchers are: (1) Recency and the migration model of CLV are key tools that merit increased attention in customer management models. (2) Recency and marketing response can interact, reinforcing Khan et al. (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.

The implications of our work for managers are: (1) Recency and migration model diagnostics 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 this is a key phenomenon. (2) The tools to derive an optimal CLV marketing policy are feasible for practical implementation. Logistic regression is very well-known to companies and the value iteration solution of a dynamic

program can be easily programmed. (3) The optimization derives recommendations for targeted one-to-one marketing. But the approach also contributes important 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*. 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 many opportunities ahead. These include: (1) 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. 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 et al. 2008). (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

1. Ackoff, Russell L. and James R. Emshoff (1975), "Advertising Research at Anheuser-Busch, Inc (1963-68)," *Sloan Management Review*, 16 (2), 1-15.
2. Ansari, Asim, Carl F. Mela, and Scott A. Neslin (2008), "Customer Channel Migration," *Journal of Marketing Research*, 45 (1), 60-76.
3. Berger, Paul D. and Nada I. Nasr (1998), "Customer Lifetime Value: Marketing Models and Applications," *Journal of Interactive Marketing (John Wiley & Sons)*, 12 (1), 17-30.
4. Bitran, Gabriel R. and Susana V. Mondschein (1996), "Mailing Decisions in the Catalog Sales Industry," *Management Science*, 42 (9), 1364-81.
5. Blattberg, Robert C., Byung-Do Kim, and Scott A. Neslin (2008), "Database Marketing Analyzing and Managing Customers." New York: Springer.
6. Bult, Jan Roelf and Tom Wansbeek (1995), "Optimal Selection for Direct Mail," *Marketing Science*, 14 (4), 378-94.
7. Dreze, Xavier and Andre Bonfrer (2008), "An Empirical Investigation of the Impact of Communication Timing on Customer Equity," *Journal of Interactive Marketing*, 22 (1), 36-50.
8. Elsner, Ralf, Manfred Krafft, and Arnd Huchzermeier (2004), "The 2003 ISMS Practice Prize Winner - Optimizing Rhenania's Direct Marketing Business through Dynamic Multilevel Modeling (DMLM) in a Multicatalog-Brand Environment," *Marketing Science*, 23 (2), 192-206.
9. --- (2003), "Optimizing Rhenania's Mail-Order Business through Dynamic Multilevel Modeling (DMLM)," *Interfaces*, 33 (1), 50-66.
10. Fader, Peter S., Bruce G. S. Hardie, and Kinshuk Jerath (2007), "Estimating CLV Using Aggregated Data: The Tuscan Lifestyles Case Revisited," *Journal of Interactive Marketing*, 21 (3), 55-71.
11. Fader, Peter S., Bruce G. S. Hardie, and Ka Lok Lee (2005), "RFM and CLV: Using Iso-Value Curves for Customer Base Analysis," *Journal of Marketing Research*, 42 (4), 415-30.
12. Gönül, Füsün, Byung-Do Kim, and Mengze Shi (2000), "Mailing Smarter to Catalog Customers," *Journal of Interactive Marketing*, 14 (2), 2-16.
13. Gönül, Füsün and Meng Ze Shi (1998), "Optimal Mailing of Catalogs: A New Methodology Using Estimable Structural Dynamic Programming Models," *Management Science*, 44 (9),

1249-62.

14. Hughes, Arthur M. (1996), *The Complete Database Marketer*, Revised, New York: McGraw-Hill.
15. Judd, Kenneth L. (1998), *Numerical Methods in Economics*. Cambridge, Mass.: MIT Press.
16. Khan, Romana, Michael Lewis, and Vishal Singh (2009), "Dynamic Customer Management and the Value of One-to-One Marketing," *Marketing Science*, 28 (6), 1063-79.
17. Leeflang, Peter S. H., Dick R. Wittink, Michel Wedel, and Philippe A. Naert (2000), *Building Models for Marketing Decisions*. Boston; Dordrecht; London: Kluwer.
18. Little, John D. C. (1970), "Models and Managers - Concept of a Decision Calculus," *Management Science Series B-Application*, 16 (8), B466-B85.
19. Miglautsch, John (2002), "Application of RFM Principles: What to Do with 1-1-1 Customers?," *Journal of Database Marketing*, 9 (4), 319.
20. Naik, Prasad A. and Nanda Piersma (2002), *Understanding the Role of Marketing Communications in Direct Marketing*. Rotterdam: Econometric Institute.
21. Neslin, Scott A., Sunil Gupta, Wagner Kamakura, Junxiang Lu, and Charlotte H. Mason (2006), "Defection Detection: Measuring and Understanding the Predictive Accuracy of Customer Churn Models," *Journal of Marketing Research*, 43 (2), 204-11.
22. Neslin, Scott A., Thomas P. Novak, Kenneth R. Baker, and Donna L. Hoffman (2009), "An Optimal Contact Model for Maximizing Online Panel Response Rates," *Management Science*, 55 (5), 727-37.
23. Pauwels, Koen and Scott A. Neslin (2008), "Building with Bricks and Mortar: The Revenue Impact of Opening Physical Stores in Multichannel Environment." working paper, Hanover, NH: Tuck School of Business, Dartmouth College.
24. Pfeifer, Phillip E. and Robert L. Carraway (2000), "Modeling Customer Relationships as Markov Chains," *Journal of Interactive Marketing (John Wiley & Sons)*, 14 (2), 43-55.
25. Reimann, Martin, Oliver Schilke, and Jacquelyn S. Thomas (2010), "Customer Relationship Management and Firm Performance: The Mediating Role of Business Strategy," *Journal of the Academy of Marketing Science*, 38 (3), 326-346.
26. Rhee, Subom, and Shelby McIntyre (2008), "Including the Effects of Prior and Recent Contact Effort in a Customer Scoring Model for Database Marketing," *Journal of the Academy of Marketing Science*, 36 (4), 538-551.
27. Rust, Roland T. and Peter C. Verhoef (2005), "Optimizing the Marketing Interventions Mix

- in Intermediate-Term CRM," *Marketing Science*, 24 (3), 477-89.
28. Simester, Duncan I., Peng Sun, and John N. Tsitsiklis (2006), "Dynamic Catalog Mailing Policies," *Management Science*, 52 (5), 683-96.
 29. Van den Poel, Dirk and Joseph Leunis (1998), "Database Marketing Modeling for Financial Services Using Hazard Rate Models," *International Review of Retail, Distribution & Consumer Research*, 8 (2), 243-257.
 30. Van Diepen, Merel, Bas Donkers, and Philip Hans Franses (2009), "Does Irritation Induced by Charitable Direct Mailings Reduce Donations?," *International Journal of Research in Marketing*, 26 (3), 180-88.
 31. Venkatesan, Rajkumar, V. Kumar, and Timothy Bohling (2007), "Optimal Customer Relationship Management Using Bayesian Decision Theory: An Application for Customer Selection," *Journal of Marketing Research*, 44 (4), 579-594.

Table 1
The Recency Trap: The Progression of the Customer to Higher and Less Profitable Recency States

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

* 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 2
Key Variable Definitions and Descriptive Statistics *

Variable	Description	Mean	Std. Dev.	Min.	Max
Recency _{ht}	Periods since last purchase by household <i>h</i> , with “1” signifying the purchase was made in month <i>t</i> .	7.659	5.621	1	20
First_Amt _h	Amount spent on first purchase by household <i>h</i> .	116.264	58.068	-110**	480
Email _t	Level of Email marketing activity in month <i>t</i> , 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 <i>t</i> .	0.672	0.689	0	2.548
Dmail	Level of direct marketing activity in month <i>t</i> , 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 <i>t</i> .	1.524	0.805	0	2.933
Purchase _{ht}	= 1 if household <i>h</i> purchased in month <i>t</i> ; 0 otherwise.	0.069	0.253	0	1

* 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 3
Logistic Regression Results

	Base Model		Model 2		Model 3		Model 4		Model 5	
Variable	Coef	P-val	Coef	P-val	Coef	P-val	Coef	P-val	Coef	P-val
Intercept	-1.105	<.001	-1.774	<.001	-1.720	<.001	-1.626	<.001	-1.623	<.001
Recency	-0.714	<.001	-0.715	<.001	-0.741	<.001	-0.790	<.001	-0.791	<.001
Recency ²	0.050	<.001	0.049	<.001	0.049	<.001	0.052	<.001	0.053	<.001
Recency ³	-0.00134	<.001	-0.00130	<.001	-0.00131	<.001	-0.00123	<.001	-0.00131	0.002
Email			0.375	0.007	0.357	0.001	0.364	0.014	0.453	0.0004
Lagged_Email			0.095	0.441	0.092	0.455	0.086	0.487	0.085	0.491
Email×Lagged_Email			-0.247	0.039	-0.242	0.043	-0.221	0.067	-0.219	0.068
Email×Recency					-0.003	0.717	-0.026	0.270	-0.107	0.045
Email×Recency ²							0.00198	0.206	0.016	0.057
Email×Recency ³									-0.00056	0.091
Dmail			0.140	0.191	0.105	0.335	0.034	0.731	-0.0046	0.970
Lagged_Dmail			0.375	0.001	0.394	0.0003	0.391	0.0003	0.384	0.0004
Dmail×Lagged_Dmail			0.090	0.116	-0.105	0.071	-0.104	0.072	-0.103	0.076
Dmail×Recency					0.020	0.011	0.065	0.001	0.102	0.020
Dmail×Recency ²							-0.004	0.016	-0.103	0.134
Dmail×Recency ³									0.0003	0.291
First_Amt	0.00349	<.001	0.00343	<.001	0.00344	<.001	0.00344	<.001	0.00344	<.001
Jun	0.103	0.208	0.067	0.461	0.077	0.385	0.085	0.353	0.087	0.338
Jul	0.006	0.947	-0.195	0.125	-0.163	0.201	-0.146	0.252	-0.142	0.267
Aug	0.239	0.003	0.223	0.205	0.242	0.176	0.230	0.197	0.229	0.201
Sep	0.548	<.001	0.495	0.002	0.504	0.002	0.489	0.003	0.488	0.003
Oct	0.183	0.021	-0.096	0.469	-0.074	0.579	-0.079	0.554	-0.072	0.590
Nov	-0.101	0.222	-0.167	0.092	-0.162	0.103	-0.169	0.088	-0.166	0.094
Dec	-0.120	0.188	-0.210	0.055	-0.191	0.080	-0.179	0.102	-0.184	0.094
Jan	-0.113	0.211	0.006	0.955	0.002	0.985	-0.007	0.950	-0.011	0.923
Feb	0.333	0.007	0.536	<.001	0.534	<.001	0.524	<.001	0.520	<.001
Mar	0.243	0.002	0.221	0.025	0.244	0.014	0.256	0.010	0.254	0.011
Apr	-0.093	0.263	0.140	0.230	0.130	0.269	0.124	0.291	0.122	0.301
N	61,952		61,952		61,952		61,952		61,952	
-2log likelihood	25925.793		25887.240		25878.750		25872.857		25869.933	
Incremental Log_LL			38.553		8.490		5.893		2.924	
Incremental P-value			<0.001		0.014		0.053		0.232	

Figure 1
Purchase Likelihood vs. Recency Calculated Directly from the Data *

* Descriptive statistics based on 61,952 customer/month observations.

Figure 2
Purchase Likelihood 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
Purchase Likelihood as a Function of Email and Dmail for Different Recency States

Probability of Purchase When Recency = 1

Probability of Purchase When Recency = 20

Figure 4
Purchase Likelihood as a Function of 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.

Appendix

Value Iteration Algorithm Used to Derive the Optimal Policy ($D^*|S$)

1. Let $V(S)^w$ = the value function at iteration w for a customer who is in state 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 ($D|S$) is a set 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, $V(S)^0$. 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. For each state 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 ($D^*|S$).
5. Test whether $V(S)^w - V(S)^{w-1} < 0.00001$ for each state S . If this 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 ($D^*|S$). 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.